

T | e | L | O | S

123
NOVIEMBRE 2023. 9€

ENTREVISTA: VOCES DE LA IA

Fundación
Telefónica

ARTIFICIAL,
SÍ; PERO
¿INTELIGENCIA?
Andrés Ortega



QUÉ DIRÍA
ARISTÓTELES
DE LA IA
Nerca Blanco

ORILLA A,
ORILLA B
**Carmen
Torrijos**

MÁQUINAS
INTELIGENTES
Doreen A. Ríos

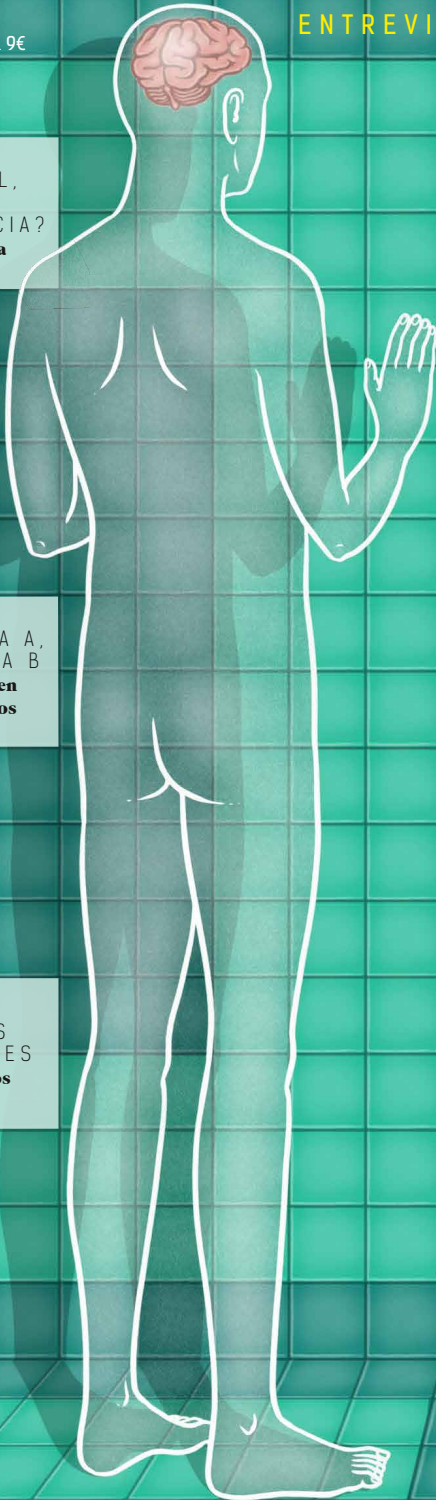
IA: UNA
HERRAMIENTA
PARA CREAR
Rodrigo Fábrega

EL CAMINO
HACIA UNA
IA SEGURA
Y SOSTENIBLE
Richard Benjamins

I
N
T
E
L
I
G
E
N
C
I
A

A
R
T
I
F
I
C
I
A
L

UN PUNTO
DE INFLEXIÓN
EN LA
HUMANIDAD



MUNDO EXPAN DIDO

ENTRE
LO FÍSICO
Y LO VIRTUAL
↗

23 NOV 2023 – 5 MAY 2024

ESPACIO FUNDACIÓN TELEFÓNICA / FUENCARRAL, 3

#EXPOMUNDOEXPANDIDO

ENTRADA GRATUITA

LENGUAJE NATURAL • LENGUAJE FORMAL • IA GENERATIVA
 PROGRAMACIÓN • TECNOLOGÍA • EDUCACIÓN • REGULACIÓN
 INTELIGENCIA ARTIFICIAL GENERATIVA • DERECHOS DE AUTOR
 CONOCIMIENTO • INNOVACIÓN • SUPERVISIÓN • PELÍCULAS
 INTELIGENCIA ARTIFICIAL • INDUSTRIA DEL CINE • CÁMARAS
 ACTORES • CIBERSEGURIDAD • PRIVACIDAD • IA GENERAL
 TECNOLOGÍA ADECUADA • IA EVOLUCIONADA • IA DISEÑADA
 IA DISTOPÍA • ARISTÓTELES FENOLÓGICA • INTELIGENCIA • CRÍTICA
 DATOS • ESTADÍSTICAS • EXPLICACIONES
 ARTE • CREATIVIDAD • AUTORIZACIÓN
 AUTOMATIZACIÓN • MÁQUINAS MENTALES
 MÁQUINAS • T | e | l | o | s COMPUTACIONAL
 CAPITAL HUMANO • N EMPRESARIAL
 DESEMPLEO • JE • IA ÉTICA
 TECNOLOGÍA RESPONSABLE • COMPROMISO SOCIAL
 RESKILLING • GESTIÓN RESPONSABLE • UPSKILLING
 INFORMACIÓN • CRISIS • EXPERTOS HUMANOS • ERA AXIAL
 MÁQUINA Y HUMANO • ESPECIFICIDAD HUMANA • SESGO
 ANTROPOCENTRISMO • CONSCIENCIA • ESTADO MENTAL
 CONDUCTA HUMANA • NEURONA ARTIFICIAL • APRENDIZAJE
 REDES NEURONALES ARTIFICIALES • DOCENTES • SALUD
 EVOLUCIÓN • MACHINE LEARNING • APRENDIZAJE FEDERADO
 EMPLEO • TRABAJO • DISCRIMINACIÓN • EMPLEABILIDAD
 PROPIEDAD INTELECTUAL • CREACIÓN CONTEMPORÁNEA



18



El presente y futuro de la IA

Las principales incógnitas que despierta el uso y desarrollo de la inteligencia artificial, en una visión global con seis expertos.

Índice

-06-
Presentación

-10-
Autor invitado

-17-
Asuntos de comunicación

-27-
Cuaderno
Inteligencia artificial.

-54-
Entrevista

-77-
Análisis

-105-
Experiencias

-115-
Biblioteca
Recursos para ampliar los conocimientos sobre la IA.

-123-
Regulación

-140-
Artefacto
El peligro de las máquinas inteligentes. Por Naief Yehya.

“La inteligencia artificial generativa es una cuestión de valores. Debemos priorizar los derechos humanos por encima de todo. Mientras las máquinas pueden pensar y existir, los humanos pueden pensar, sentir y ser”.

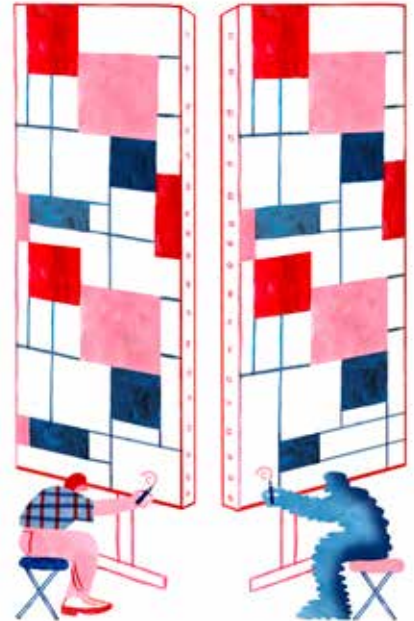
84



78



130



**-10-
CARMEN TORRIJOS**

La especialista en IA aplicada a la comunicación cuenta cómo las tribus del lenguaje natural y formal firmaron la paz.

**-18-
MILLÁN BERZOSA**

El asesor de estrategia digital y gobernanza IT analiza la IA generativa y los derechos de autor.

**-22-
EDUARDO GROJO**

¿Quiere la inteligencia artificial destruir el cine? Se pregunta el director, guionista y productor de cine y publicidad.

**-28-
RICHARD BENJAMINS**

El responsable de IA en Telefónica marca "El camino hacia una IA general segura y sostenible".

**-34-
NEREA BLANCO MARAÑÓN**

La filósofa reflexiona sobre qué diría Aristóteles de la IA.

**-40-
ANDRÉS ORTEGA**

Artificial, sí; pero ¿inteligencia? Un análisis del director del Observatorio de las Ideas.

**-46-
DOREEN A. RÍOS**

La curadora detalla la evolución del uso de la IA en el arte.

**-64-
GERARDO IVÁN TUDURÍ**

El investigador explica cómo construir una mente biodigital.

**-70-
ANDRÉS PEDREÑO**

El experto detalla la relación de España con la (IA) generativa.

**-78-
RODRIGO FÁBREGA**

Cómo potenciar la creatividad con la IA en la educación.

**-84-
ANA SAINZ**

Historia de humanos y algoritmos. La conexión entre progreso económico y social.

**-88-
MANUEL CEBRIÁN**

Cómo la IA puede ayudarnos a navegar por las crisis.

**-94-
CARLOS BLANCO**

¿Qué nos hará humanos? La capacidad de inventarnos a nosotros mismos.

**-100-
ADOLFO CASTILLA**

La consciencia humana ante la irrupción de la inteligencia artificial general (IAG).

**-106-
JUAN CARLOS F. GALINDO**

Desarrollos y miedos de la aplicación de la IA a la salud.

**-124-
JOSÉ VARELA FERRIO**

Quo vadis machina? El impacto de la inteligencia artificial en el empleo.

**-130-
MARTA SUÁREZ-MANSILLA**

¿Es la IA una herramienta para la creación contemporánea o un competidor de los artistas?

T | E | L | O | S 123

Siempre es buen momento para pararse a pensar

Vivimos en una era privilegiada: inmersos en la revolución tecnológica más significativa de la historia de la humanidad, y esta revolución está en su inicio. La convergencia de Internet, la capacidad de cómputo, la web 2.0 y los teléfonos móviles ya supusieron un cambio profundo, pero algo aún más trascendente se está desarrollando.

Mientras la ley de Moore parecía alcanzar sus límites, emergen las redes neuromórficas que emulan el funcionamiento de neuronas y sinapsis en el cerebro humano, lo que amplía nuestra capacidad de procesar información. Además, las tecnologías cuánticas nos abren un horizonte casi sin límites.

La inteligencia artificial (IA) va más allá de los algoritmos y el *big data*, el reconocimiento de voz, de imagen y el aprendizaje profundo. Modelos como ChatGPT, aunque imperfectos, han deslumbrado por su capacidad y su poder.

Mientras nos adaptábamos al impacto de la web 2.0, la web 3 ha surgido con realidades virtuales y aumentadas, la adopción generalizada de *blockchain* y la *tokenización*. Las redes de fibra y el 5G están desplegándose, convirtiéndose en superordenadores omnipresentes. La convergencia de una capacidad de procesamiento y almacenamiento casi ilimitada, redes de comunicación transformadas en superordenadores y la descentralización de la web allanan el camino para la implantación masiva de la IA y la llegada de la inteligencia artificial generativa (IAG).

La IAG, capaz de desarrollar nuevo conocimiento y realizar tareas que estaban aparentemente reservadas a los humanos, es un hito tecnológico que podría superar a la imprenta o la fisión nuclear en su impacto en la humanidad. Aunque su fecha de llegada es incierta, estamos cada vez más cerca de la inteligencia artificial general o una inteligencia artificial casi humana. En la última década, la velocidad de aprendizaje de los modelos de IA ha aumentado significativamente, alimentándolos con grandes cantidades de datos disponibles en Internet.

Hoy en día, la IAG puede superar exámenes médicos y de abogacía en el percentil más alto, escribir buena parte del código de un ingeniero de *software* y desarrollar habilidades complejas hasta crear una falsa realidad para nosotros, los humanos. Distinguir entre texto generado por la IAG y texto humano se ha convertido en una tarea complicada.

La IAG tiene un potencial exponencial para avanzar en campos como la biomedicina, la mitigación del cambio climático y la descomposición del plástico. Sin embargo, también plantea riesgos existenciales, como la creación de armas químicas o cibernéticas, el auge de las *fake news* y desinformación indetectables, e incluso un escenario de pérdida de control sobre la IAG por parte de las empresas que la desarrollan.

La IAG es un llamado a las ciencias sociales, como la sociología, la filosofía, la antropología y el derecho, para definir un nuevo contrato social que garantice que la tecnología sea beneficiosa para todos y que proteja derechos fundamentales como la privacidad, la seguridad y la verdad en un mundo digital. Esta revolución tecnológica nos enfrenta a desafíos éticos, como la elección del coeficiente intelectual de nuestros hijos y la corrección de sesgos en los datos de la IAG.

En el pasado, revoluciones tecnológicas como la industrial, la energía atómica y la investigación bacteriológica trajeron avances, pero también desafíos. La IAG es una cuestión de valores, y debemos priorizar los derechos humanos por encima de todo. Mientras las máquinas pueden pensar y existir, los humanos pueden pensar, sentir y ser.

Es hora de detenernos, reflexionar y ser conscientes de nuestro papel en esta era de cambios sin precedentes.



JUAN MANUEL ZAFRA
Director de TELOS

Ni siquiera las IA más avanzadas son hoy capaces de pensar; no son conscientes, pero pueden desarrollar una enorme capacidad lingüística y creativa. La IA actuales crean una copia comprimida de toda la información que han recogido durante su aprendizaje y su entrenamiento.

Nuestra civilización es el resultado de aplicar la inteligencia humana. **La inteligencia artificial es una consecuencia más de la automatización, sistémica y sistemática**, en la que se encuentra la humanidad desde hace siglos.

Ante la extensión del uso de la IA, debemos considerar cuestiones muy relevantes – **la privacidad, la seguridad, la no discriminación, la transparencia y la responsabilidad social**, entre otras– que afectan a la esencia de nuestras sociedades democráticas.

La inteligencia artificial representa una gran oportunidad para ampliar nuestra inteligencia y expandir nuestras capacidades de razonamiento y creativas para la resolución de problemas complejos.

El progreso de la IA se debe a tres factores clave: los avances algorítmicos, la disponibilidad de grandes cantidades de datos y la capacidad de computación.

Toda herramienta tan poderosa como la inteligencia artificial generativa **puede ser beneficiosa o perjudicial dependiendo de quién la utilice y con qué fines.**

Nos encontramos inmersos sin vuelta atrás en un nuevo paradigma coevolutivo en el que la humanidad y la inteligencia artificial se han embarcado conjuntamente, gestando una interdependencia y cohabitación que exigen respuestas sin demora, **porque la IA no esperará al ser humano.**

Debemos aceptar con humildad esta coevolución en la que nos encontramos, pues marca una nueva era de la evolución humana. Si el mundo del futuro tiene alguna posibilidad de ser mejor lo será gracias a la participación de la inteligencia artificial, pues su uso aumentará las capacidades de la inteligencia humana.

La IA ha de utilizarse de manera responsable y beneficiosa para la sociedad en su conjunto. Para prevenir riesgos urge establecer un marco de regulación que vele por el desarrollo y el uso de la inteligencia artificial.

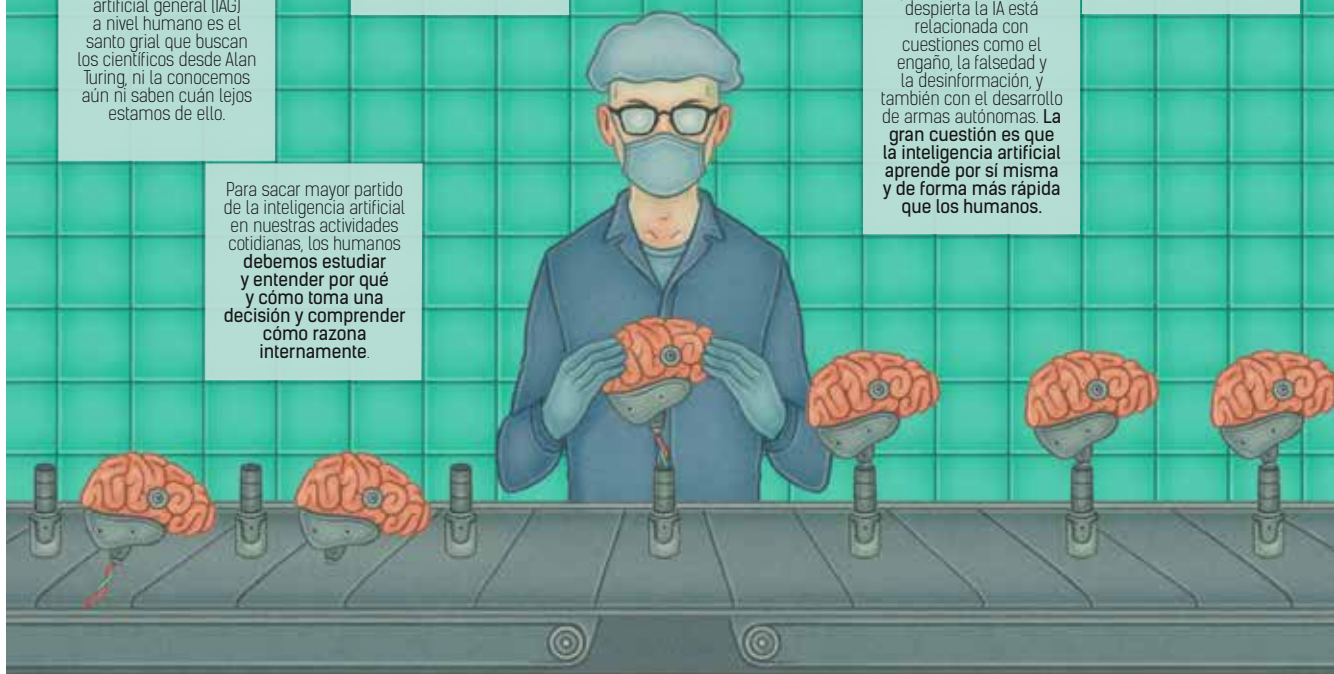
Es una certeza que la IA obligará a reconvertirse a millones de empleados porque sus funciones serán redundantes e innecesarias. Es innegable que también aparecerán nuevos puestos de trabajo como el de ingeniero de instrucciones para hacer trabajar a la inteligencia artificial.

La complejidad de las neuronas biológicas y sus conexiones dista mucho de las neuronas artificiales computacionales y de su capacidad de proceso. Aunque una verdadera inteligencia artificial general (IAG) a nivel humano es el santo grial que buscan los científicos desde Alan Turing, ni la conocemos aún ni saben cuán lejos estamos de ello.

La IA tiene el potencial para darle la vuelta al mercado laboral y a la economía global. Se prevé un cambio de perfiles de puestos de trabajo que afectará a todos los sectores, incluidas las tareas creativas. La formación continua es uno de los grandes retos a los que nos enfrentamos.

Una de las grandes preocupaciones que despierta la IA está relacionada con cuestiones como el engaño, la falsedad y la desinformación, y también con el desarrollo de armas autónomas. **La gran cuestión es que la inteligencia artificial aprende por sí misma y de forma más rápida que los humanos.**

Para sacar mayor partido de la inteligencia artificial en nuestras actividades cotidianas, los humanos **debemos estudiar y entender por qué y cómo toma una decisión y comprender cómo razona internamente.**



INTELIGENCIA ARTIFICIAL UN PUNTO DE INFLEXIÓN EN LA HUMANIDAD

Colaboradores

ILUSTRACIÓN:
JORGE ESTEBAN

CARMEN TORRIJOS

Traductora, filóloga, lingüista computacional y especialista en IA aplicada a la comunicación. Es responsable de inteligencia artificial en Prodigios Volcán, profesora asociada de Minería de Textos en la UC3M, y coautora de *La primavera de la inteligencia artificial* (Catarata, 2023).



MILLÁN BERZOSA

Asesor de Estrategia Digital y Gobernanza IT con perspectiva ética y cultural. Consejero independiente y ejecutivo con liderazgo en *Big Tech*. Doctor en Periodismo (UCM), licenciado en Derecho (UNED), con posgrados en Liderazgo, Marketing y Finanzas. Profesor en excedencia de la Universidad Francisco de Vitoria.



EDUARDO GROJO

Director, guionista y productor de cine y publicidad. Licenciado en Ciencias de la Información por la Universidad Complutense. Miembro de la Academia de Cine. Imparte clases de creatividad publicitaria en el IE Business School y de guion en la UNIR.



DAVID SÁNCHEZ

Es uno de los ilustradores y novelistas gráficos más importante de España. Sus últimas obras son *Los años de internet* (2021) junto con el fundador de WeTransfer, Damian Bradfield, *Un millón de años* (2017), *En otro lugar, un poco más tarde* (2019) y *Fuego de bengala* (2023).



RODRIGO FÁBREGA

Especialista en desarrollo de la creatividad e inteligencia artificial como herramienta cognitiva en educación. Lidera la Fundación Cruzando de innovación educativa. Es doctor de Teoría y política de la educación por PennState University. Actualmente es *Visiting Scholar* del grupo *Lifelong Kindergarten* del Media Lab MIT.



ANA SAINZ

Directora general de la Fundación SERES, una organización nacida hace 14 años que impulsa la integración de los criterios ESG (medioambiente, social y de gobernanza por sus siglas en inglés) en la estrategia de las empresas, con especial foco en el valor de lo social.



MANUEL CEBRIÁN

Investigador distinguido y catedrático de Excelencia Permanente en la Universidad Carlos III de Madrid. Sus afiliaciones anteriores incluyen la Sociedad Max Planck, el Instituto Tecnológico de Massachusetts, la Organización de Investigación Científica e Industrial del Commonwealth (CSIRO) y la Universidad de California en San Diego.

RICHARD BENJAMINS

Chief Responsible AI Officer en Telefónica y Head of AI for Society and Environment. Cofundador del Observatorio del Impacto Social y Ético de la Inteligencia Artificial (OdiseIA) y consejero del CDP (la ONG para el reporting del cambio climático). Copresidente del Consejo Empresarial de la UNESCO por su recomendación sobre la IA y asesor de IA ética para el Vaticano.



CARLOS BLANCO

Profesor titular de la Universidad Pontificia Comillas. Autor de *La integración del conocimiento* (2018).

NEREA BLANCO MARAÑÓN

Licenciada en Filosofía por la Universidad Autónoma de Madrid. Creadora de la plataforma Filoosfers, un proyecto de divulgación filosófica en línea y presencial. Autora del libro *Filosofía entre líneas* y coautora del libro de primero de Bachillerato de Filosofía de la Editorial Santillana.



ADOLFO CASTILLA

Doctor ingeniero del ICAL y catedrático de Economía aplicada. MBA por The Wharton School y máster en Ingeniería de Sistemas e Investigación Operativa por The Moore School. Es presidente de AESPLAN y del Capítulo Español de la World Future Society.

ANDRÉS ORTEGA

Escritor, politólogo y periodista, ha sido director del Departamento de Estudios en la presidencia del Gobierno en dos ocasiones, investigador del Real Instituto Elcano y director del Observatorio de las Ideas. Su último libro, *La imparabla marcha de los robots* (2017), dio pie a que comisionara la exposición *Nosotros, Robots*, de Fundación Telefónica.



JUAN CARLOS F. GALINDO

Periodista, divulgador y comunicador. Ha gestionado la comunicación de empresas tecnológicas como Telefónica, Sherpa.ai, Bit2Me, Genially, INLAB, Divilo y ha escrito en cabeceras como *Dinero*, *Actualidad Económica* o *Muy Interesante*.

DOREEN A. RÍOS

Curadora e investigadora independiente. Su trabajo se enfoca en el arte digital, prácticas postdigitales y nuevas materialidades. Fundadora de [ANTI] MATERIA, plataforma en línea dedicada a la investigación y exposición de arte producido a través de medios digitales.



JOSÉ VARELA FERRÍO

Responsable de Digitalización en el Trabajo de UGT y experto en Derecho en inteligencia artificial. Ha publicado más de 50 estudios y artículos de opinión relacionados con las nuevas tecnologías y el trabajo. Es empleado de Telefónica España desde 1991.

GERARDO IVÁN TUDURÍ

Investigador, escritor y artista visual. Fundador del ADSM (Análisis Digital de Sistemas Mentales). Estudió Medicina en Montevideo y cine en España, especializándose en la creación de ficción colectiva. Trabaja como analista diseñando respuestas del sistema mental al impacto de la inteligencia artificial.



MARTA SUÁREZ-MANSILLA

Abogada especializada en Derecho del Arte, propiedad intelectual, arte y tecnología, y patrimonio cultural, además de gestora cultural. Fundadora de la asesoría jurídica Artworld Law y vicepresidenta de ADA (Asociación de Derecho del Arte).

ANDRÉS PEDREÑO

Catedrático de universidad en Economía Aplicada. Doctor Honoris Causa por la Nottingham Trent University. Rector de la Universidad de Alicante (1993-2000) y CEO de Unersia (2000-2004). Autor de más de 70 libros, entre los que destacan *Europa frente a EE. UU. y China y España en la nube*. *Afrontando los retos en la era de la inteligencia artificial*.



NAIEF YEHYA

Escritor, narrador, periodista y crítico cultural. Autor, entre otros, de *Las cenizas y las cosas* (2018), *Tecnocultura* (2008) y *El cuerpo transformado*. *Cyborgs y nuestra descendencia tecnológica en la realidad en la ciencia ficción* (2001).

...EL DÍA QUE EL
EL LENGUAJE
NATURAL Y EL
FORMAL FIRMARON
LA PAZ...

HISTORIA
Y LENGUAJE
DE PROGRAMACIÓN
A
CUPIDO

lb|67|hd(2)|f.-|s|?|d(3)|f.-|s|
?|d(3)|b|b|67|hd(2)|f.-|s|?|d
(3)|f.-|s|?|d(3)|f.-|s|?|d(3)|b|b|6
7|hd(2)|f.-|s|?|d(3)|f.-|s|?|d(3)|
b|b|67|hd(2)|f.-|s|?|d(3)|f.-|s|
?|d(3)|b|b|67|hd(2)|f.-|s|?|d
3)|f.-|s|?|d(3)|b|b|67|hd(2)|

(HOLA)

HOLA





CARMEN TORRIJOS
Lingüista computacional y
especialista en IA aplicada
a la comunicación

**SOBRE CÓMO
LAS TRIBUS
DEL LENGUAJE
NATURAL
Y FORMAL
FIRMARON
LA PAZ**



**Palabras
clave:**

Lenguaje natural,
lenguaje formal,
IA generativa,
programación,
tecnología,
educación

Orijis A,

La IA generativa trata de allanarnos el camino a la información y la producción, pero en su frenética llegada no ha sido consciente de la existencia de dos tribus bien diferenciadas: la que solo utiliza el lenguaje natural, y la que también puede comunicarse con los sistemas en lenguaje formal.

On how the natural and formal language tribes signed peace
BANK A, BANK B

Generative AI is trying to pave the way to information and production, but in its frantic arrival it has not been aware of the existence of two distinct tribes: the one that only uses natural language, and the one that can also communicate with systems in formal language.

Keywords: Natural language, formal language, generative AI, programming, technology, education.



1 OpenAI: "ChatGPT plug-ins" en openai.com (marzo 2023). Disponible en: <https://openai.com/blog/chatgpt-plugins>

2 Expedia Group: "ChatGPT wrote this press release – no, it didn't, but it can now assist with travel planning in the Expedia app" en *Expedia.com* (abril 2023). Disponible en: <https://www.expediagroup.com/investors/news-and-events/financial-releases/news/news-details/2023/Chatgpt-Wrote-This-Press-Release--No-It-Didnt-But-It-Can-Now-Assist-With-Travel-Planning-In-The-Expedia-App/default.aspx>

3 Hu, K. "ChatGPT sets record for fastest-growing user base - analyst note" en *Reuters* (febrero 2023). Disponible en: <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>

Había una vez un río, con un pueblo en cada orilla. Los habitantes de una orilla, en la que vivía la mayoría de la gente, se comunicaban únicamente en lenguaje natural, aunque en distintos idiomas. Escribían novelas y poemas, pero también leyes, manuales y tratados. En la otra orilla, mucho menos poblada, los habitantes utilizaban a diario este lenguaje, pero también comprendían y utilizaban el lenguaje formal cuando era necesario. Programaban, estudiaban matemáticas y lógica y combinaban ambos lenguajes para vivir, trabajar o incluso hacer bromas. No entendían que los habitantes de la otra orilla no hicieran lo mismo, pero al mismo tiempo disfrutaban de atesorar un lenguaje privado. En la orilla del lenguaje natural, el lenguaje formal era considerado como algo complicado y ajeno, que les generaba una gran pereza. Los adolescentes elegían muy pronto en qué orilla querían vivir y, a partir de ese momento, un río ancho y profundo y muy pocos puentes hacían difícil la comunicación con la otra orilla.

Así, las relaciones entre ambos pueblos llevaban siglos estancadas cuando la inteligencia artificial llegó para instalarse en la región. Para poder desarrollarla, era necesario conformar equipos con profesionales de uno y otro lado del río, que pronto descubrieron que tenían dificultades para entenderse en las cuestiones más básicas del proyecto. La gente de la orilla del lenguaje formal empezó a construir por su cuenta un sistema de codificación que permitía a los nativos del lenguaje natural utilizar sus herramientas y producir más rápido, y así nació algo a lo que llamaron *inteligencia artificial generativa*. Por primera vez pusieron en circulación una moneda de uso común entre los pueblos, a la que llamaron *prompt*. Sin embargo, después de tantos siglos, las fronteras entre las dos formas de comunicación no eran tan fáciles de neutralizar. En ese sistema, aún en pruebas, las capas superpuestas de un lenguaje y otro se confundían y transparentaban.

Orilla A: lenguaje natural

La historia que acabamos de contar es más alegórica que ficticia. La revolución de la IA generativa, que nos permite hacer peticiones a un *chatbot* para asistirnos en tareas que antes nos llevaban mucho más tiempo, como escribir un informe o ilustrar un artículo, no es solo tecnológica. Es también social y cultural, y en el ámbito económico abre nuevos perfiles profesionales, nuevos nichos de mercado y toda una nebulosa de consecuencias y efectos donde por ahora solo vemos dos o tres pasos más allá.

La interacción entre el código informático y el lenguaje se ha revelado en su máxima expresión precisamente en las herramientas de IA generativa


Los complementos o *plug-in* para ChatGPT son un buen ejemplo¹. No muchas marcas los han descubierto todavía, pero los casos de uso siguen creciendo y empiezan a comportarse como las nuevas *app*. El complemento de Expedia², por ejemplo, te permite preparar un viaje con datos reales, como horarios de vuelos, alojamientos disponibles y tarifas. Si estás preparando un viaje a la República Checa, puedes pedirle que te sugiera hoteles para dormir en Praga, que ajuste la búsqueda a tu precio máximo por persona y noche, y que te recomiende los mejores restaurantes típicos en Brno o dónde beber cerveza en Pilsen. También puede ayudarte con la organización de los horarios y el plan de ruta, y aquí no hay riesgo de que nos cuele falsos datos: podemos confiar, porque la información no proviene del modelo de lenguaje, sino del *plug-in*, que nos facilita los enlaces a la web de Expedia para que podamos comprobarlos. Nuestra confianza en los resultados estará, por tanto, en relación directa con nuestra confianza en la marca y no en el sistema de IA. Otro complemento interesante es el de Amazon Gift Chooser, que te ayuda a elegir el mejor regalo siempre y cuando le des información sobre la persona a la que quieres regalar, sus gustos y tu límite de gasto. Con cada idea, añadirá un enlace a Amazon para que puedas comprarlo directamente si te convence, aunque ahora mismo solo trabaja con Amazon EE. UU. y con precios en dólares. El objetivo siempre es el mismo, y no es pequeño: permitir a los usuarios combinar las capacidades de conversación GPT con la oferta de valor de una marca, en una plataforma web que sobrepasó los 100 millones de usuarios solo dos meses después de su lanzamiento³. Todas las peticiones a estos sistemas se realizan únicamente en lenguaje natural, en un nuevo género literario a medio camino entre la instrucción y la descripción al que llamamos *prompt*.

Accedemos a estos complementos de dos maneras, ambas sencillas una vez que estamos en la pista: la primera es dentro del propio ChatGPT, en su versión de pago, habilitando las funcionalidades beta. Cuando selecciona-

mos GPT-4, aparece una pestaña *plug-in* donde podemos encontrar un *plug-in store*, una tienda al más puro estilo Google Play o Apple Store, con un catálogo de extensiones gratuitas que podemos instalar y usar sin más. La segunda opción, más accesible si no tenemos la versión de pago, es en forma de extensión para Google Chrome, que se instalará en nuestro navegador y se activará cuando abramos ChatGPT como una capa adicional. Entre estas últimas se encuentran algunas muy prácticas, como editGPT, que te permite aplicar un control de cambios similar al de Word para saber en qué mejora el *chatbot* tus textos, o Prompt Genius, un administrador de *prompts* destinado a cubrir algunas carencias de funcionalidad en ChatGPT. Este genio incorporado te permite algo fundamental, que es guardar tus *prompts* favoritos en la propia interfaz para no tener que reescribirlos cada vez. Por ejemplo, si has dado con el *prompt* ideal para tus notas de prensa o tus menús semanales, puedes lanzarlos cada vez que lo necesites sin tener que escribirlos de nuevo.

Fue precisamente trasteando con Prompt Genius cuando encontré algo que me llamó la atención: la posibilidad de definir variables dentro de los *prompts*. Tardé un tiempo en entender por qué una funcionalidad que reconocía tan natural y necesaria me resultaba a la vez tan extraña en el contexto, tan estridente y fuera de lugar como un iglú en una playa tropical: en el ejemplo, en medio de una frase en español, había una palabra encerrada entre dobles llaves ({{}}).

Orilla B: lenguaje formal

Empezando por el principio: una variable es un concepto conocido para cualquier estudiante de los cursos más básicos de iniciación a la programación. Es un elemento contenedor que no es fijo como el resto del código, como indica su nombre, sino que su contenido puede cambiar con 

cada ejecución. Ya no ocurre con tanta frecuencia, pero en los primeros tiempos de los envíos masivos por correo electrónico era habitual recibir de vez en cuando un *email* con el mensaje “Te damos la bienvenida <NOMBRE>”. Esta sensación de desencanto al encontrarnos con las tripas mismas de la informática se produce porque estamos ante una variable, el nombre de la persona receptora, una información cambiante en cada mensaje que se extrae de una base de datos en la que cada correo electrónico lleva asociado un nombre de pila, que se insertará automáticamente. La extensión Prompt Genius traslada este concepto de las variables informáticas al *prompt* y permite incluir en él una o más, marcadas en este caso con dobles llaves ({}). Es una manera rústica de incluir una señal gráfica para que el sistema sepa que esa palabra es un comodín que puede adquirir distintos valores. Por ejemplo, imagina que has dado con el *prompt* perfecto para que ChatGPT escriba por ti una guía turística para tus fines de semana de escapada. Has invertido un tiempo en crear ese *prompt*, pero hay contenido que cada vez es diferente, como la ciudad que vas a visitar. Es tan sencillo como dejar la palabra {{ciudad}} entre llaves, y Prompt Genius te preguntará cuál es en cada nueva ejecución, por lo que tendrás el *prompt* preparado en cuestión de segundos. Lo que estamos haciendo no es más que acumular facilitadores adicionales sobre una herramienta ya de por sí facilitadora como es ChatGPT. Nos obsesiona, desde hace varios años, poder poner cualquier contenido que queramos al alcance de una simple petición en nuestro idioma, y caminamos en una trayectoria sin vuelta atrás que nos lleva a allanar más y más el terreno.

Desde luego estamos ante una aproximación ganadora, salvo porque lo que llamamos *nuestro idioma* está mucho más limitado de lo que creemos. El lenguaje natural es el que nos permite expresar nuestras ideas y emociones y crear conexiones sociales, algo que es y va a seguir siendo

prioridad por pura supervivencia. Ahora bien, a diferentes velocidades vamos llegando a momentos evolutivos en los que este sistema de expresión encuentra límites. Un ejemplo es el uso del lenguaje inclusivo en castellano, un idioma que ha visto brotar un conflicto social de una dualidad de género de herencia indoeuropea, y de la ausencia de un género neutro que pudiera rescatarnos⁴. Y aquí, en una sociedad occidental en la que las visiones de progreso valoran el bilingüismo como un tesoro de las personas y los pueblos, batallamos a diario para no tener que traspasar las fronteras entre el lenguaje natural y el formal. Este último lo estamos

utilizando sin darnos cuenta, por ejemplo, cuando indicamos la opcionalidad con una barra (*los/las alumnos/as*), un gesto moderno, seguramente no anterior al siglo XVIII y totalmente ajeno al lenguaje natural, pero útil para expresar dos opciones equivalentes en el mínimo espacio posible. Por un lado, la convención de la barra (/) ha experimentado un fuerte crecimiento desde la llegada de Internet y la mensajería instantánea. En el lenguaje informal por escrito es eficaz en lo que comunica, es breve, resuelve. Y, sin embargo, en un lenguaje más elevado tendemos a evitarla por poco elegante, como en un artículo o en un libro. Ocurre también

con el rechazo generalizado al uso de la arroba (@)⁵ para marcar la inclusividad, una solución escrita también breve y eficaz, pero que hemos descartado de plano y que ya apenas vemos en comunicaciones, publicidad o prensa. ¿Por qué? Es sencillo. En el caso de la barra de opcionalidad, cuando nos la encontramos leyendo en voz alta nos vemos en la obligación de marcarla con una pausa o un gesto leve. En el caso de la arroba, nuestro cerebro colapsa.

Mi sensación de incomodidad al encontrar una palabra entre llaves en medio de un *prompt*, si bien no es verbalizada muy a menudo, es común a gran parte de los hablantes: más allá de las letras y nuestros signos de puntuación habituales, los corchetes y llaves, como los asteriscos o el



4 Álvarez Mellado, Elena: “Lenguaje inclusivo: algunas claves lingüísticas” en *elDiario.es* (julio 2021). Disponible en: https://www.eldiario.es/opinion/zona-critica/lenguaje-inclusivo-claves-linguisticas_129_8021559.html

5 FundéuRAE: “La arroba (@) para el género, uso inadecuado”, en *fundeu.es* (febrero 2011). Disponible en: <https://www.fundeu.es/recomendacion/arroba/>

6 https://es.wikipedia.org/wiki/Ken_Thompson

porcentaje, son símbolos que percibimos como ajenos al lenguaje natural, venidos de otro planeta, de alguna manera no del todo válidos para la comunicación entre las personas. Como si fueran propiedad más bien del mundo de los números, o de las máquinas.

Entre tanto, al otro lado del río, las expresiones regulares llevan décadas siendo una herramienta de inmensa utilidad para los programadores, irreductibles a pesar del tiempo y la llegada de técnicas mucho más avanzadas. Se trata de conjuntos de caracteres que representan de forma simbólica un patrón de búsqueda en un texto, y que en los años 60 el matemático Ken Thompson introdujo entre las funcionalidades de un editor de texto llamado QED⁶. Por ejemplo, imagina que quieres encontrar números de teléfonos móviles españoles en un millón de textos de Internet. Tendrías que indicar en la expresión regular que estás buscando tres conjuntos de tres dígitos, separados por espacios o no, y marcar que el primer dígito del primer conjunto pueda ser tanto un 6 como un 7, porque muchos teléfonos en España empiezan ya por un 7. La expresión podría quedar más o menos así:

```
\b[67]\d{2}[-.\s]?\d{3}[-.\s]?\d{3}\b
```

Nada parecido a la secuencia de nueve números de nuestro teléfono, pero con una eficacia del cien por cien para detectar ese, o cualquier otro. La opcionalidad en este caso se marca con corchetes “[67]”, mientras que la expresión “\d” sirve para indicar que en ese lugar se espera un dígito cualquiera. Entre llaves “{}” indicamos el número de dígitos que queremos encontrar en cada conjunto, junto con otros símbolos para identificar un espacio (“\s”) o marcar el inicio y fin de la regla (“\b”). Estas expresiones son complejas de aprender y de crear, pero una vez dominadas despliegan todo su potencial. Una convención tan arbitraria como nuestras palabras o nuestros fonemas, pero más eficiente, que en la actualidad aún sorprende por cómo sostiene desde las sombras grandes proyectos de *software*.

El puente

Que el grueso de la población digitalizada tuviera un conocimiento básico en expresiones regulares, o que se incluyeran de forma natural como aprendizaje desde el instituto, podría mejorar la eficiencia en la búsqueda de información, pero también promover un mayor entendimiento entre los expertos en programación y el

Nadie es hablante
nativo del lenguaje
formal. Generalizar
el hábito
de aprenderlo
y utilizarlo,
o fomentar
el pensamiento
numérico y lógico,
es una misión
colectiva

resto de disciplinas, además de fomentar el desarrollo de habilidades valiosas, como el pensamiento lógico y sintético. En absoluto se trata de sustituir el lenguaje natural. La motivación es más bien ampliar sus límites democratizando una herramienta que ha probado con creces su valor. La interacción entre el código informático y el lenguaje, hasta ahora en orillas opuestas de un río con muy pocos puentes, se ha revelado en su máxima expresión precisamente en las herramientas de IA generativa. Poco a poco vamos adquiriendo la costumbre de interactuar con ellas sin código (las llamamos *no-code* o *low-code*). Vamos normalizando el texto a texto, texto a imagen, texto a música, texto a vídeo, integramos el *prompting* en nuestras vidas como nuestra nueva interfaz de trabajo. Con los tres o cuatro trucos que ruedan por las redes vamos aprendiendo a pedir a los sistemas lo que necesitamos, sin mayor esfuerzo. Pero es precisamente esta revolución del texto-a-todo la que ha puesto ante nuestros ojos una nueva colisión entre lenguajes. En un camino supuestamente allanado, aparecen diminutos baches donde el militante del lenguaje natural se encuentra con el código sin querer.

La manifestación más reciente de todo esto, y quizá el puente más moderno y mejor trazado hasta la fecha en este río, es el *Code Interpreter* que OpenAI anunció el 6

La revolución de la IA generativa no es solo tecnológica. Es también social y cultural

de julio de 2023⁷. Esta funcionalidad, útil por igual para programadores y no programadores, es una evolución natural de ChatGPT que nos permite subir ficheros, interactuar con ellos y obtener resultados más allá de la conversación escrita. Podemos, por ejemplo, subir un fichero de datos macroeconómicos y obtener gráficas automáticas, o pedirle que nos genere un código QR a partir de un enlace a una web. Podemos pedirle que redacte un informe a partir de unos datos y exportar después el resultado en PDF, o que extraiga el texto que aparece en una imagen. Ahora bien, esta magia revolucionaria sería imposible sin un paso intermedio: la máquina traduce nuestra petición a código en un lenguaje de programación muy concreto, Python, que podemos ver generarse ante nuestros ojos, y es el propio código el que genera el resultado, que después el sistema nos devuelve en un enlace descargable. La razón de que OpenAI haya decidido no hacer el proceso cien por cien transparente para los usuarios no programadores es probablemente maximizar su público, y esto ha generado una coyuntura sin precedentes. Por un lado, los habitantes de la orilla del lenguaje natural tendrán que habituarse a ver molestos fragmentos de código precediendo sus resultados. Por otro, los habitantes de la orilla del código ahora pueden crearlo sin escribirlo, cambiando el esfuerzo de lugar para construir en lenguaje natural el mejor *prompt* posible. Es significativo también que OpenAI decidiera, semanas después del lanzamiento de este módulo, cambiar el nombre de *Code Interpreter*, evidente para una orilla pero no para la otra, a *Advanced Data Analysis*. Mucho más explicativo, mucho más para todo el mundo, mucho más puente.

En esta era de la inteligencia artificial nos encontramos ante tecnologías tan capaces, con un potencial tan brillante, que nos sitúan una y otra vez en un espejo frente a nuestras propias limitaciones. Algunas de ellas son naturales, predeterminadas —nuestra capacidad de cálculo, mínima, a años luz de la artificial—, pero otras son adquiridas o inconscientemente autoimpuestas. Un condicionamiento cultural omnipresente, aunque muy poco evidenciado hasta ahora, ha bloqueado a lo largo de los siglos a buena parte

de la población la inclinación a utilizar el lenguaje formal cuando es necesario. Un instrumento tan humano y útil como la lengua, que descartamos porque cometemos el error de entenderlo como ajeno a la especie.

No sabemos si caminamos hacia un mundo donde todos programaremos o hacia un mundo *no-code* en el que programarán los de siempre, pero, más allá de la incertidumbre, no podemos discutir la utilidad de una mirada híbrida, capaz de comunicarse en ambos lenguajes sin tantos complejos. Quizá es un buen momento para revisar si nuestros itinerarios educativos, tan cuidadosamente contruidos, están actuando como restricciones que han perdido su sentido original. No solo la división tradicional entre ciencias y letras o la elección exclusiva de asignaturas en el bachillerato son parte de un sistema caducado, también los mensajes constantes en los que la gente adulta nos posicionamos —a veces con vergüenza, a veces con un orgullo mal entendido— en una u otra orilla.

Nadie es hablante nativo del lenguaje formal. Generalizar el hábito de aprenderlo y utilizarlo, o fomentar el pensamiento numérico y lógico, es una misión colectiva que va más allá de construir puentes sobre un río, y que es tan capacitadora e integradora como el bilingüismo natural. Consiste en dejar de decir que no todo el mundo vale para todo, y atreverse más a intentarlo. Tiene mucho más que ver con perder el miedo a nadar.

Bibliografía

Álvarez Mellado, E.: "Lenguaje inclusivo: algunas claves lingüísticas" en *ElDiario.es* (julio 2021). Disponible en: https://www.eldiario.es/opinion/zona-critica/lenguaje-inclusivo-claves-linguisticas_129_8021559.html

Expedia Group: "ChatGPT wrote this press release – no, it didn't, but it can now assist with travel planning in the Expedia app" en *Expedia.com* (abril 2023). Disponible en: <https://www.expediagroup.com/investors/news-and-events/financial-releases/news/news-details/2023/Chatgpt-Wrote-This-Press-Release--No-It-Didn-But-It-Can-Now-Assist-With-Travel-Planning-In-The-Expedia-App/default.aspx>

FundéuRAE: "La arroba (@) para el género, uso inadecuado", en *fundeu.es* (febrero 2011). Disponible en: <https://www.fundeu.es/recomendacion/arroba/>

Hu, K.: "ChatGPT sets record for fastest-growing user base - analyst note" en *Reuters* (febrero 2023). Disponible en: <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>

OpenAI: "ChatGPT plug-ins" en *openai.com* (marzo 2023). Disponible en: <https://openai.com/blog/chatgpt-plugins>

Yildirim, S.: "ChatGPT Code Interpreter: how it saved me hours of work", en *Towards Data Science* (julio 2023). Disponible en: <https://towardsdatascience.com/chatgpt-code-interpreter-how-it-saved-me-hours-of-work-3c65a8dfa935#>

⁷ **Yildirim, Soner:** "chatGPT Code Interpreter: how it saved me hours of work", en *Towards Data Science* (julio 2023). Disponible en: <https://towardsdatascience.com/chatgpt-code-interpreter-how-it-saved-me-hours-of-work-3c65a8dfa935#>



La inteligencia artificial permite automatizar la producción de contenido informativo y de entretenimiento y está revolucionando todo el universo de la comunicación y de la creatividad en general. La IA desafía la autenticidad, el aura, pero también ayuda a los creadores humanos a traspasar los límites de la expresión artística y narrativa.

Asuntos de comunicación



MILLÁN BERZOSA
Asesor de estrategia
digital y gobernanza IT con
perspectiva ética y cultural

IA GENERATIVA Y DERECHOS DE AUTOR

Una nueva sinfonía de gobernanza

La inteligencia artificial tiene todo para ser una buena sinfonía. El auge de su vertiente generativa supone una oportunidad de progreso, pero recuerda la necesidad de proteger el conocimiento, modernizar la regulación y defender esquemas viables de derechos de autor.

La inteligencia artificial presenta un universo de aristas, complejo, en crecimiento, pero también de grandes retos por afrontar. Quizá el primero de esos retos sea el de su propio marco de existencia. Hoy, la inteligencia artificial (IA) ni se emociona ni emociona, ni es vulnerable, ni siente, ni sufre. Pero es una sinfonía en desarrollo.

Así que vamos a cerrar los ojos e idear esa sinfonía incipiente, para desgranar la complejidad de este auge de la inteligencia artificial, principalmente la de tipo generativo. Pensemos en los grandes modelos de lenguaje y en el procesamiento del lenguaje natural, de ChatGPT de OpenAI (con o sin Bing), Bard de Google, o promesas por parte de otras plataformas como Amazon, Meta, Anthropic o GitHub CoPilot para código... y la necesidad de un marco regulatorio robusto. Entre múltiples movimientos y melodías que buscan converger en una pieza compacta, vale la pena detenerse en los riesgos y debilidades, también en las oportunidades, que requieren pronta atención y un enfoque de gobernanza.

I. Inicio sinfónico: la promesa del renacimiento tecnológico

La inteligencia artificial, cual movimiento sublime de una sinfonía que nos invita a experimentar emociones y descubrimientos, presenta una primera fase de oportunidades. Es la música del potencial positivo para hacer de esta sociedad una sociedad más sofisticada. Con la posibilidad de creación de contenido desde cero (escritura avanzada y creación de imágenes desde la nada), traducción masiva de contenido en texto o audio, búsqueda mucho más eficaz de recursos, análisis de *software* malicioso, asistentes mucho más intuitivos, desarrollo de código, automatización aplicada a fuerza de ventas, análisis de mercado más fácilmente escalable e, incluso, detección y prevención de ciberataques, gracias a que el procesamiento de ingentes cantidades de *sets* de datos permite señalar vulnerabilidades y patrones sospechosos para generar alertas de manera anticipada.

Ante esta gran composición, la inteligencia artificial merece una regulación de altura, que permita impulsar la innovación de una manera exponencial y lo haga con garantías. Tiene un auténtico poder de transformación e influencia, con oportunidades únicas para los diferentes actores, tanto públicos como privados.

La oportunidad es clara: establecer un marco normativo específico para la inteligencia artificial que permita aportar certidumbre sin apagar el crecimiento y la innovación.

Las iniciativas globales de gobernanza dan esperanza, frente a ideas desconectadas de silos que quizá también puedan reorientarse hacia la multilateralidad y los acuerdos de *partnerships* conjuntos. Lo más esperanzador viene con el “Proceso de Hiroshima” de los líderes del G7¹ para

¹ Katanuma, M. (2023). “El G-7 acuerda establecer el ‘Proceso de Hiroshima’ para gobernar la IA”. Bloomberg, 2023. Disponible en: <https://www.bloomberglinea.com/2023/05/20/el-g-7-acuerda-establecer-el-proceso-de-hiroshima-para-gobernar-la-ia/>



Generative AI and copyright
**A NEW GOVERNANCE
SYMPHONY**

Artificial intelligence has everything to be a good symphony. The rise of its generative side is an opportunity for progress, but reminds us of the need to protect knowledge, modernize regulation and defend viable copyright schemes.

Keywords: generative artificial intelligence, regulation, regulation, knowledge, copyright, innovation, supervision.

ILUSTRACIÓN: JEFF BENEFIT

Palabras clave:
inteligencia artificial generativa, regulación, conocimiento, derechos de autor, innovación, supervisión.

governar la inteligencia artificial, el borrador de Acta de la Ley de Inteligencia Artificial del Parlamento Europeo², la llamada de las Naciones Unidas³ para crear un regulador global de inteligencia artificial, o incluso en Canadá⁴ o Estados Unidos, con el empuje de demócratas y republicanos en este último para que “maximicemos sus beneficios y reduzcamos sus daños⁵”. Hay muchos intereses en juego y muchas posibilidades como la creación de resúmenes del uso de datos o identificar obras completas. Toda sinfonía empieza con un primer movimiento.

II. Segundo movimiento: desarrollo y protección

Para una buena base armónica, un componente fundamental de la inteligencia artificial es la gestión del conocimiento. En esta era del todo digital, es imperativo proteger el legado del saber en respuesta a los riesgos de reproducción y apropiación que avanzan hacia su *crescendo*. Conforme avanza la inteligencia artificial, la vulnerabilidad emergente de la producción intelectual se hace más evidente. Resulta inquietante cómo algunos abogan por que estén al mando empresas distantes, conglomerados con sucursales que otrora soñaron con roles más prominentes.

En nuestra analogía, sin duda es preferible ser un músico en plena ejecución que uno que espera su turno, pero ¿y si adaptamos la música para que las regiones tengan el lugar que les corresponde? Una música en la que la inteligencia artificial tenga un encaje cultural y ético, con reglas que nos permitan desarrollarnos a todos. Un camino hacia la confianza armoniosa desde la gobernanza mejor entendida, la que antepone el interés de las personas y busca la cooperación y el beneficio social.

Lo que ha evolucionado es la tecnología, no la base. La protección de la propiedad intelectual pasó de ser un “privilegio discrecional”⁶ a un “derecho universal” ya durante el siglo XIX, y hasta nuestros días ha navegado *adagios* de aparente calma y *allegros* ciertamente rápidos. Que el conocimiento es una materia sensible es algo fuera de duda, tanto en el derecho anglosajón como en el derecho de origen continental europeo. El *copyright* se ha afinado hacia el concepto de propiedad de autor, y la batalla por la titularidad se libra con actores tan diversos como las grandes corpora-

ciones, las sociedades de gestión de derechos, las agencias recaudadoras, o el sector tecnológico. Una propiedad que parece cada vez más abocada a un valor estimado con metodología de mercado, pero que debe ser cuadrada con la consideración del bien de interés general que supone.

Y en medio de todo, llega la popularización de la inteligencia artificial generativa, capaz de espolear la velocidad y la habilidad de creación de contenido utilizando ingentes bases de datos. Estas bases de datos, compuestas de inicio por elementos únicos, tienden a expandirse con componentes mixtos de origen incierto, y por eso es relevante trabajar en estándares de transparencia más estrictos. Al final, para tener unas especificaciones claras en lenguaje *tech*, que preserven la creación cultural humana, con límites mejor definidos en la minería de datos y la manera en que los automatismos se alimentan de contenido.

III. Tercer movimiento: las promesas se han vuelto desafíos

En este punto, la orquesta muestra signos de inquietud. Es primordial una transparencia en las fuentes empleadas y en los momentos de su uso. Es una armonía posible desde el punto de vista tecnológico, éticamente conveniente y económicamente sostenible. Estamos en el *crescendo* del desafío, con una llamada a empujar, a instar la regulación coordinada, liderada de manera coral desde las regiones sí, pero con la coordinación y el liderazgo de aquellos que ya estén preparados.

A la vez, el mundo corporativo debe trascender de la moda inicial de la inteligencia artificial generativa y del enfoque solo tecnólogo para ahondar en los elementos sociales y éticos, con liderazgo de sus órganos de gobierno, y empapando la cultura corporativa en clave de responsabilidad. Es importante identificar las oportunidades y decidir ir a por ellas con el criterio de buen gobierno y posicionamiento estratégico.

Porque, no nos engañemos, vivimos también tiempos de amenazas y exposición ciertas. Tanto el ciberacoso hacia el que no piensa igual, la desinformación (ya sea casual, *misinformation*, o causal, *disinformation*), y hasta el riesgo de espionaje, van en aumento.

Mientras que una sinfonía puede acelerar hacia un clímax inesperado, la velocidad de crecimiento y evolución de la inteligencia artificial abre la puerta a que actores con intenciones perjudiciales aspiren a apropiarse la batuta. Y que sueñen con hacerlo con rapidez, en busca de un giro de guion, esta vez de caos, de influencia, de destrucción de valor. Los grandes actores tienen más que perder y la gestión del riesgo es clave: si no se adaptan a la melodía en rápida evolución, primeros violines pueden perder su puesto.

Las grandes plataformas por su parte, al igual que un poderoso movimiento en una sinfonía, tienen gran capacidad cuando se trata de ejecución. Sin asemejarse en legitimidad a los entes gubernamentales de sociedades de-

2 Sección de noticias del Parlamento Europeo: “Ley de IA de la UE: primera normativa sobre inteligencia artificial”. www.europarl.europa.eu/news/es/headlines/society/20230601ST093804/ley-de-ia-de-la-ue-primer-normativa-sobre-inteligencia-artificial

3 Sección de noticias de Naciones Unidas: “La inteligencia artificial precisa de una regulación para que beneficie a todos”. UN.org, 2023. Disponible en: <https://news.un.org/es/story/2023/07/1522542>

4 Gobierno de Canadá: Directiva canadiense sobre la toma de decisiones automatizadas. Canada.ca, 2023. Disponible en: <https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592>

5 Osornio, A. (2023). “Regular la IA en EE UU requiere cambios en las leyes de privacidad”. *Wired*, 2023. Disponible en: <https://es.wired.com/articulos/regular-ia-en-ee-uu-requiere-cambios-en-leyes-de-privacidad>

6 Bracha, D. (2016). *Owning Ideas: The Intellectual Origins of American Intellectual Property, 1790-1909*. Cambridge University Press, pág. 34.



La autorregulación voluntaria, aun siendo positiva, es claramente insuficiente

mocráticas, su poder requiere ser moderado para prevenir excesos y evitar un *crescendo* descontrolado.

La propagación de sesgo discriminatorio es un problema real. Como una sinfonía en manos de un músico inexperto, los automatismos, con su inmenso potencial y recorrido, requieren una dirección precisa, buena guía. Es fundamental afinar la inteligencia artificial generativa para evitar desbordar la partitura. La diversidad está en riesgo, precisamente, por la capacidad de multiplicación del sesgo, algo tan humano.

IV. *Allegro moderato* para el renacimiento tecnológico y la cooperación

La inteligencia artificial, cual movimiento sublime de una sinfonía que invita a las emociones, puede ser la música que ayude a hacer de esta una sociedad mejor.

El empuje regulatorio actual es notoriamente mejorable, con exceso de confianza en que los cuerpos legales están preparados, sin ser conscientes de la naturaleza cambiante de la tecnología y que avanza por la vía de los hechos. Los desarrollos en inteligencia artificial avanzan rápidamente, exponencialmente. La autorregulación voluntaria, aun siendo positiva, es claramente insuficiente, y la falta de voluntad política para abordar desafíos es perturbadora.

Un ejemplo positivo es cuando Bruselas hizo de avanzadilla en su día al demandar revisores humanos en las redes sociales. Para que las medidas de control de contenido fueran efectivas y escalables, se necesitaba una cierta automatización, pero también más protección de la privacidad. Podemos construir desde eso. Necesitamos regulación se-

rena pero efectiva, que no tome una eternidad y con mecanismos de flexibilidad que permitan ir al compás adecuado.

Deben existir interlocutores adecuados, que no solo memoricen argumentarios, sino que cuenten genuinamente con preparación y sean personas conectadas. La regulación debe estar guiada por principios claros. Se necesita una visión de conjunto, pero también una especificidad ante cada reto. La agilidad y adaptabilidad son igualmente clave, así como la necesidad de planificación de órganos de control, supervisión y mitigación, centrados tanto en el avance del *software* como en el desarrollo del *hardware*, los tan esenciales *chips*, y la protección del contenido.

Una gobernanza descentralizada, en sintonía con el carácter descentralizado de una inteligencia artificial de nivel superior garantizaría un mayor equilibrio entre control y autonomía. Tenemos ante nosotros una oportunidad sin precedentes para crear órganos de supervisión con un alto grado de independencia, sabios consejos y experiencia, con una vocación genuina de coordinación multilateral entre regiones, en busca de estándares de transparencia sinfónicos, para que cada paso relevante en el mundo de la inteligencia artificial resuene con claridad y sea visible y entendible para todos.

Bibliografía

Bracha, O. (2016). *Owning Ideas: The Intellectual Origins of American Intellectual Property, 1790-1909*. Cambridge University Press, pág. 34.

Katanuma, M. (2023). "El G-7 acuerda establecer el 'Proceso de Hiroshima', para gobernar la IA". Bloomberg, 2023. Disponible en: <https://www.bloomberglinea.com/2023/05/20/el-g-7-acuerda-establecer-el-proceso-de-hiroshima-para-gobernar-la-ia/>

Osornio, A. (2023). "Regular la IA en EE. UU. requiere cambios en las leyes de privacidad". *Wired*, 2023. Disponible en: <https://es.wired.com/articulos/regular-ia-en-ee-uu-requiere-cambios-en-leyes-de-privacidad>



EDUARDO GROJO

Director, guionista y productor de cine y publicidad. Profesor de creatividad publicitaria en el IE Business School.

LOS NUEVOS DEMIURGOS SON LOS CREADORES DE PROMPTS

¿Quiere la IA destruir el cine?

¿Qué está pasando con los contenidos audiovisuales tras los avances en inteligencia artificial? ¿Van a ser capaces las IA de hacer películas sin necesidad de actores ni de técnicos? ¿Y serán buenas esas películas?



Palabras clave:

Películas, inteligencia artificial, industria del cine, cámaras, actores

ChatGPT, escíbeme 1.800 palabras sobre el impacto de la IA en el mundo audiovisual, y cómo el cine está incorporando esta tecnología: “En los últimos años, la inteligencia artificial (IA) ha experimentado un crecimiento exponencial en diversas áreas, revolucionando la forma en que interactuamos con la tecnología y...”.

No lo pude evitar. Pregunté a ChatGPT. Y redactó un texto muy **correcto**, muy impersonal, poco simpático y con algunos datos interesantes, aunque obvios, sobre el impacto de la IA en el audiovisual. Nada más lejos de lo que me apetecía escribir. Para empezar, porque llevo casi 30 años trabajando en esta industria y he vivido —y estoy viviendo— en primera persona el cambio tecnológico que rodea a la producción de cine y contenidos audiovisuales. Me parecía más oportuno ilustrar esa revolución con experiencias vividas, antes que recurrir a la mera recopilación de datos.

En los últimos años, meses, semanas, días, ha cambiado mucho la forma de hacer y de ver las películas. A principios de siglo —me encanta decir principios de siglo para referirme al año 2001, que sigue pareciendo algo futuro—, aún recuerdo, cuando se rodaba en analógico, celuloide, 35mm, y se veía con cierto menosprecio la calidad de una cámara de televisión, esa cámara que nunca tendría la definición de un buen revelado en laboratorio. Pocos años después, ya se hablaba de HD, de rodar en 2k, 4k, 8k... y un gigante como Kodak agonizaba como un T-Rex cualquiera. Tremendo. Pilló a la industria desprevenida.

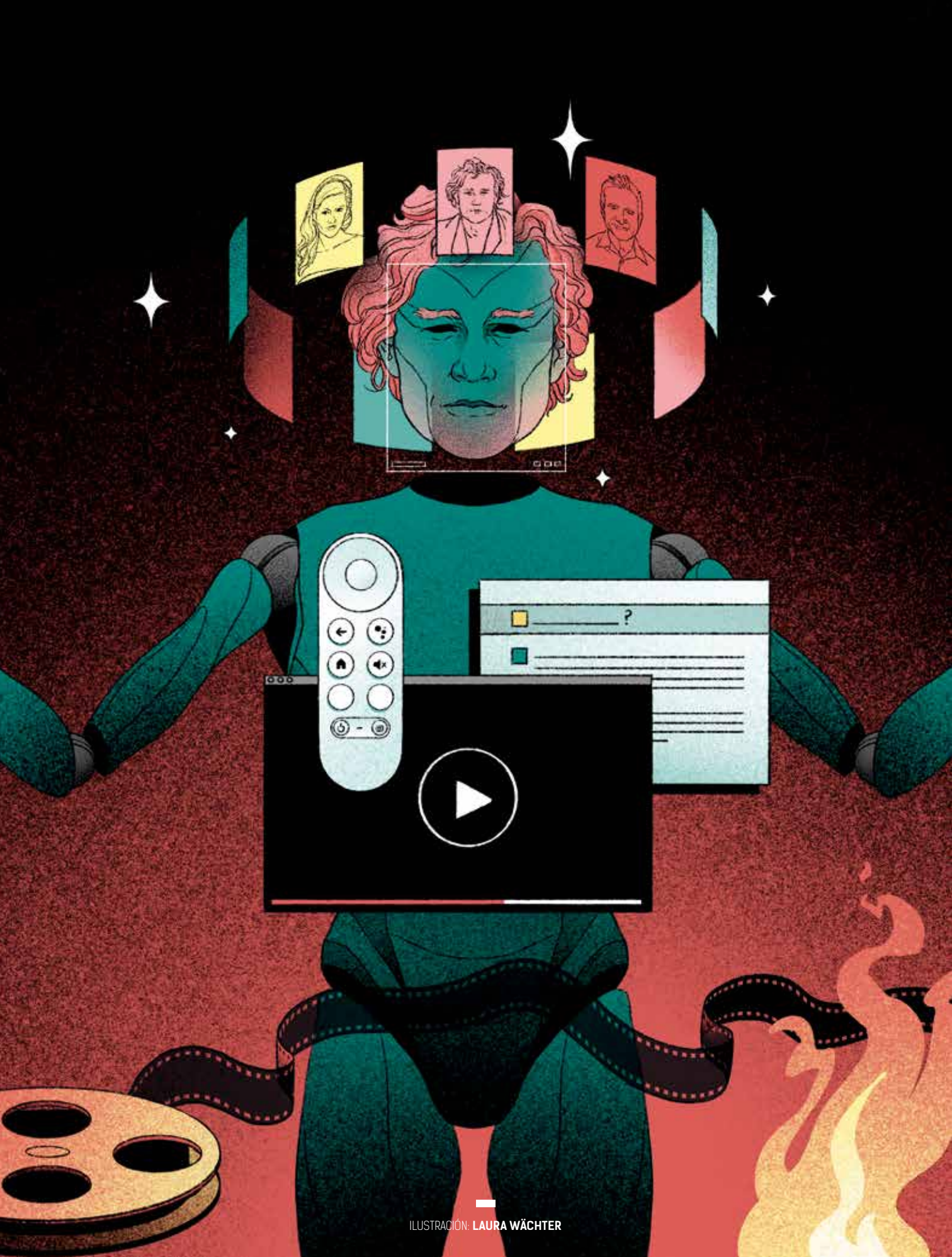
Laboratorios cinematográficos como Fotofilm o Madrid Film tuvieron que cerrar. Distribuidores y exhibidores reformularon sus planteamientos. Esas enormes bobinas con las que se proyectaban las películas se convertían en DCP, discos duros o pinchos que se conectaban directamente a los nuevos proyectores digitales. *Cinema Paradiso* no habría sido lo mismo con un niño enchufando discos externos a un proyector digital.

Muchas salas tuvieron que cerrar porque no les compensaba cambiar de proyector. Sobre todo, porque la gente ya no iba al cine. La calidad de sus televisores invitaba a quedarse en casa a ver las películas. La piratería del cine y series por Internet daba paso a las plataformas —Netflix, HBO, Disney Plus o Amazon— que utilizan IA para conocer nuestros gustos y hacernos recomendaciones. La gente está de nuevo dispuesta a pagar por ver contenidos. Se puede decir que las plataformas acabaron con gran parte de la piratería, pero también con la maravillosa experiencia de ver un *film* en una sala con más gente. No te ríes igual estando solo en tu salón que en una sala con mil personas riendo a la vez. Es otra experiencia. ▶▶▶

The new demiurges are prompts creators
DOES AI WANT TO DESTROY THE CINEMA?

What is happening to audiovisual content in the wake of advances in artificial intelligence? Will AI be able to make movies without the need for actors or technicians? And will those movies be any good?

Keywords: movies, artificial intelligence, film industry, cameras, actors.



Y de experiencias estamos hablando. De repente, casi sin avisar, cambió el paradigma. Se rodaba sin película, se podía ver el plano rodado sin necesidad de esperar a revelar y positivar y se podía hacer *backup* en el momento. Hoy ya no hay miedo por el negativo, del que solo existía un original y que podía dañarse de mil maneras diferentes —una vez tuve que repetir un rodaje porque las latas del negativo, que venían de Londres, habían pasado por los rayos del escáner del aeropuerto; película estropeada, gasto enorme en volver a rodar, en fin, dinero, dinero y dinero...—. Rodar en HD es más barato, ya no hay procesos fotoquímicos en el laboratorio. Pero no nos engañemos, lo más caro de un *film* no es la película. Empezaba una nueva era en la que había que replantear muchas cosas, desde los presupuestos hasta las cámaras con las que se iba a rodar.

Pero vamos con la IA. Hemos visto cómo se rescataban personajes como el Grand Moff Tarkin, en *Rogue One*, interpretado por Peter Cushing en *Star Wars* (1977), y regenerado por ordenador en 2016. Cushing murió en 1994. Se ha rejuvenecido a actores como Robert de Niro en *El irlandés* de Martin Scorsese, y hace poco hemos visto a Harrison Ford haciendo casi de joven Indiana Jones. También hemos leído que Bruce Willis ha “donado” su imagen para que puedan seguir rodando películas con él y sin él, tal y como hiciera Robin Wright en aquella distópica película llamada *El congreso*. Lo único que pedía Robin en el *film* al ofrecer y escanear su imagen era que no se usara para *films* violentos o pornográficos. Hoy eso parece una batalla perdida. Con la IA, cada vez es más fácil utilizar cualquier imagen sin autorización y dotarla de nuevos significados. Se está empezando a usar sin permiso ni límites la imagen de cualquiera para hacer vídeos falsos. Todos los días vemos *deep fakes*¹ de Tom Cruise bailando en TikTok o de Donald Trump declarando cosas insólitas (¿Es Trump un robot?). No hay control. Cualquiera puede hacer esas cosas, incluso con un móvil. Todo se acelera.

Robin Wright no quería que se usara su imagen para el porno. Recordemos que cualquier avance tecnológico rápidamente se sexualiza. En 1896 se rodó el primer *film* erótico, *Le coucher de la mariée* (*Bedtime for the Bride*). El cine se había inventado apenas un año antes, en 1895, cuando los hermanos Lumière proyectaban sus primeros trabajos en el Grand Café de París. Con la llegada de Internet, aparecieron numerosos negocios en torno al porno y también con la realidad virtual (VR). Y ya circulan por las redes vídeos manipulados con IA de actrices famosas, que ven con impotencia cómo se utiliza su imagen sin pudor ni respeto.

Películas como *Her* (de Spike Jonze) o *A.I. Inteligencia Artificial* (de Steven Spielberg) ya hablaban de la relación afectiva entre una IA y los seres humanos. Aunque para mí la película que más está acertando en sus pronósticos del futuro es *WALL · E*, aquel *film* de animación que parecía una ingenua historia en la que un robot que limpiaba el planeta de basura se enamoraba de Eva, otra pequeña robot. En el *film*, toda la humanidad engordaba plácidamente sin nada que hacer, pues todos los trabajos los hacían las máquinas. Nos encaminamos a eso, todos gordos y aburridos escribiendo *prompts*. Y eso también lo harán mejor que nosotros.

Los nuevos creadores

En 2020 estrené mi cortometraje *Conversaciones con un mono*, en el que un director de cine en horas bajas (Nancho Novo) adquiere un robot creado para generar ideas y guiones. El robot IAN (Inteligente Artificial y Narrativo) funciona mediante aprendizaje profundo, y fue en cierto modo el antecesor en la ficción de ChatGPT. En 2019, cuando escribí el guion, se habían creado GPT 1 y GPT 2, y estaba en pañales GPT 3. En esa época, lo más parecido a una IA guionista era un *software* llamado Benjamin que escribió un cortometraje incomprensible en 2016. Para

Con la IA, cada vez es más fácil utilizar cualquier imagen sin autorización y dotarla de nuevos significados

escribir mi guion, me documenté, pregunté y me entrevisté con expertos de todo tipo, desde ingenieros a filósofos. La conclusión a la que llegué fue obvia: esto es imparable.

A nivel de imagen, la creación de universos y personajes mediante IA se ha disparado también. DALL-E 2, Midjourney, Stable Diffusion... Cada semana hay mejoras verdaderamente exponenciales. Se crean mundos nuevos con una calidad increíble. Imágenes sorprendentes y espectaculares generadas en pocos segundos, que cualquier diseñador tardaría horas en crear. En breve, esos nuevos mundos y personajes se podrán animar. Los nuevos demiurgos son los creadores de *prompts*.

Están llegando *softwares* como Gen-2 o Pika Labs para generar vídeos mediante IA. Todavía son muy básicos, pero de aquí a dentro de poco —francamente, no me atrevo a ponerle fecha a eso— veremos películas creadas completamente por IA. ¿Serán buenas? Seguramente al principio no. Pero habrá algún *accidente* (como llamaba el pintor Francis Bacon a ese trazo involuntario que de verdad conseguía emocionarle) que nos sorprenderá a todos. Hasta que ese accidente se convierta en la norma y haya todo tipo de *films* fabulosos para todos los gustos. Los veremos. Sin duda. Recordemos que la IA aprende de sus errores. Y se mejora a sí misma.

En la actualidad, ya hay *softwares* de IA para doblaje, para editar vídeos, para mezclar sonido, para etalonar el color... Será interesante volver a leer este artículo dentro de uno o dos años para ver qué ha pasado.

En el audiovisual, no debemos pensar en la IA como un peligro, sino como una nueva herramienta. O una suma de herramientas. Como sucede con todo, como pasó con el HD desplazando al celuloide, llevará un tiempo acostumbrarnos a estas nuevas tecnologías. El problema puede ser precisamente ese, el tiempo. Cuando nos acostumbramos a un nuevo *software*, rápidamente aparece otro que lo mejora. Y ya está aquí Google Bard, pisándole los talones a ChatGPT. Tenemos que ser muy flexibles y estar muy atentos.

Imposible para las IA

Quiero terminar este artículo contando algo personal, haciendo un juego de asociaciones como hacen los niños, algo que ChatGPT no podría escribir nunca porque el juego está basado únicamente en experiencias propias. Cuando empecé a escribir este artículo, me pregunté qué influencia habían tenido en mí las películas

1 El término *ultrafalso* es una alternativa al anglicismo *deep fake*, que se está empleando sobre todo para referirse a los vídeos que no son reales, pero que lo parecen gracias a una manipulación extrema. Más información en: <https://www.fundeu.es/recomendacion/ultrafalso-alternativa-a-deep-fake>

2 Neuralink es una compañía de neurotecnología especializada en el desarrollo de interfaces cerebro-computadora implantables, fundada por Elon Musk.

con robots o inteligencias artificiales. Me di cuenta de que habían estado ahí desde que era niño, C-3PO y R2-D2 (*Star Wars*), Nexus 6 (*Blade Runner*) o HAL 9000 (*2001: Una odisea del espacio*). Se me ocurrió entonces una forma de asociación para llegar, en cinco pasos, de mi robot guionista IAN a HAL 9000, la IA de Kubrick en *2001: A Space Odyssey*:

- 1.- IAN es un robot que genera ideas y textos para escribir guiones. En su base de datos están todos los libros y guiones que se han escrito. Está conectado a Internet, y se mejora a sí mismo mediante *deep learning*, aprendiendo de sus errores, como hizo AlphaZero para aprender a jugar al ajedrez.
- 2.- El nombre de IAN es un homenaje que quise hacer a Ian Holm, actor que interpretaba al androide Ash en la película *Alien*.
- 3.- La cabeza que se fabricó en *Alien* para la muerte de Ash, aquella cabeza medio humana medio robótica, la construyó Colin Arthur, creador de efectos especiales en *films* como *La Historia Interminable* o *2001* (y además, un gran amigo mío y mentor desde hace muchos años).
- 4.- En 2002, Colin me dejó el disfraz original de uno de los monos del *film 2001* para un anuncio que dirigí para una marca de electrodomésticos inteligentes. Rodamos al mono en diferentes posiciones y los juntamos en posproducción para que parecieran una tribu.
- 5.- Esos mismos monos, efectivamente, comparten *film* con HAL 9000, la IA de *2001*.

Y ya está. De IAN a HAL en cinco pasos. Se cerró el círculo.

No dudo que este juego de asociaciones se pueda hacer en un futuro con los recuerdos de la gente, cuando Elon Musk imponga Neuralink² en nuestras vidas. Por ahora, solo puedo concluir que ese mono que me dejó Colin es lo más cerca que podré estar nunca de Stanley Kubrick. Aunque, quién sabe lo que se podrá hacer en unos años con unos cuantos *prompts*...

Filmografía

- Folman, A. (2013): *El congreso*. Liverpool, Paul Thiltges Distribution, Pandora Film.
- Jonze, S. (2013): *Her*. Annapurna Pictures, Stage 6 Films.
- Kubrick, S. (1968): *2001: Una odisea del espacio*. Metro-Goldwyn-Mayer (MGM).
- Lucas, G. (1977): *La guerra de las galaxias*. 20th Century Fox.
- Scorsese, M. (2019): *El irlandés*. Netflix.
- Scott, R. (1982): *Blade Runner*. Warner Bros.
- Scott, R. (1979): *Alien, el octavo pasajero*. 20th Century Fox, Brandywine Productions.
- Spielberg, S. (2001): *A. I. Inteligencia Artificial*. Warner Bros., DreamWorks.
- Stanton, A. (2008): *WALL-E*. Pixar Animation, Walt Disney Studios, FortyFour Studios.

Cuaderno

INTELIGENCIA ARTIFICIAL



ILUSTRACIONES:
DAVID SÁNCHEZ

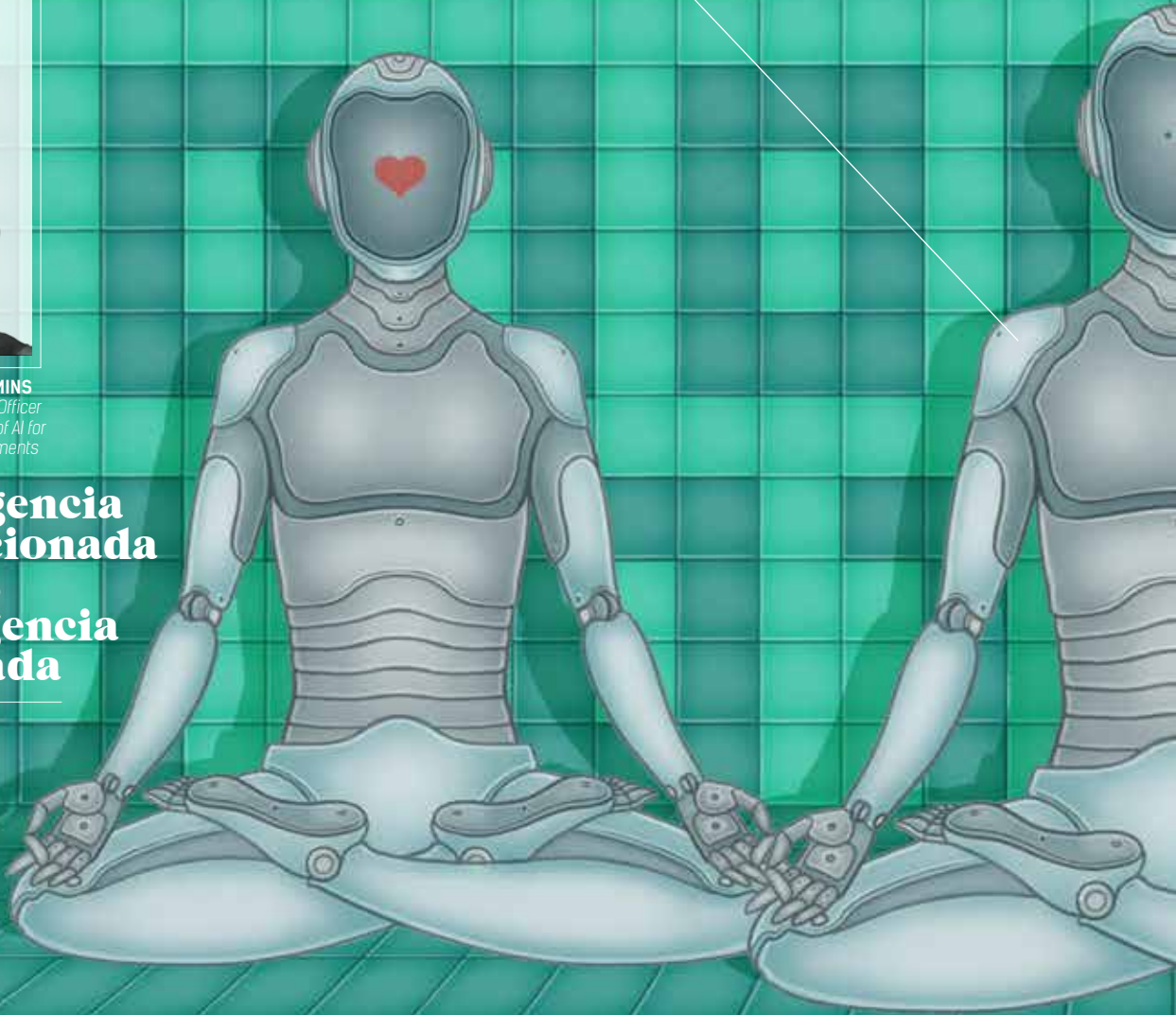


RICHARD BENJAMINS
*Chief Responsible AI Officer
en Telefónica y Head of AI for
Society and Environments*

Inteligencia evolucionada versus inteligencia diseñada

**Palabras
clave:**

IA general,
IA generativa,
IA evolucionada,
IA diseñada,
IA distopía.



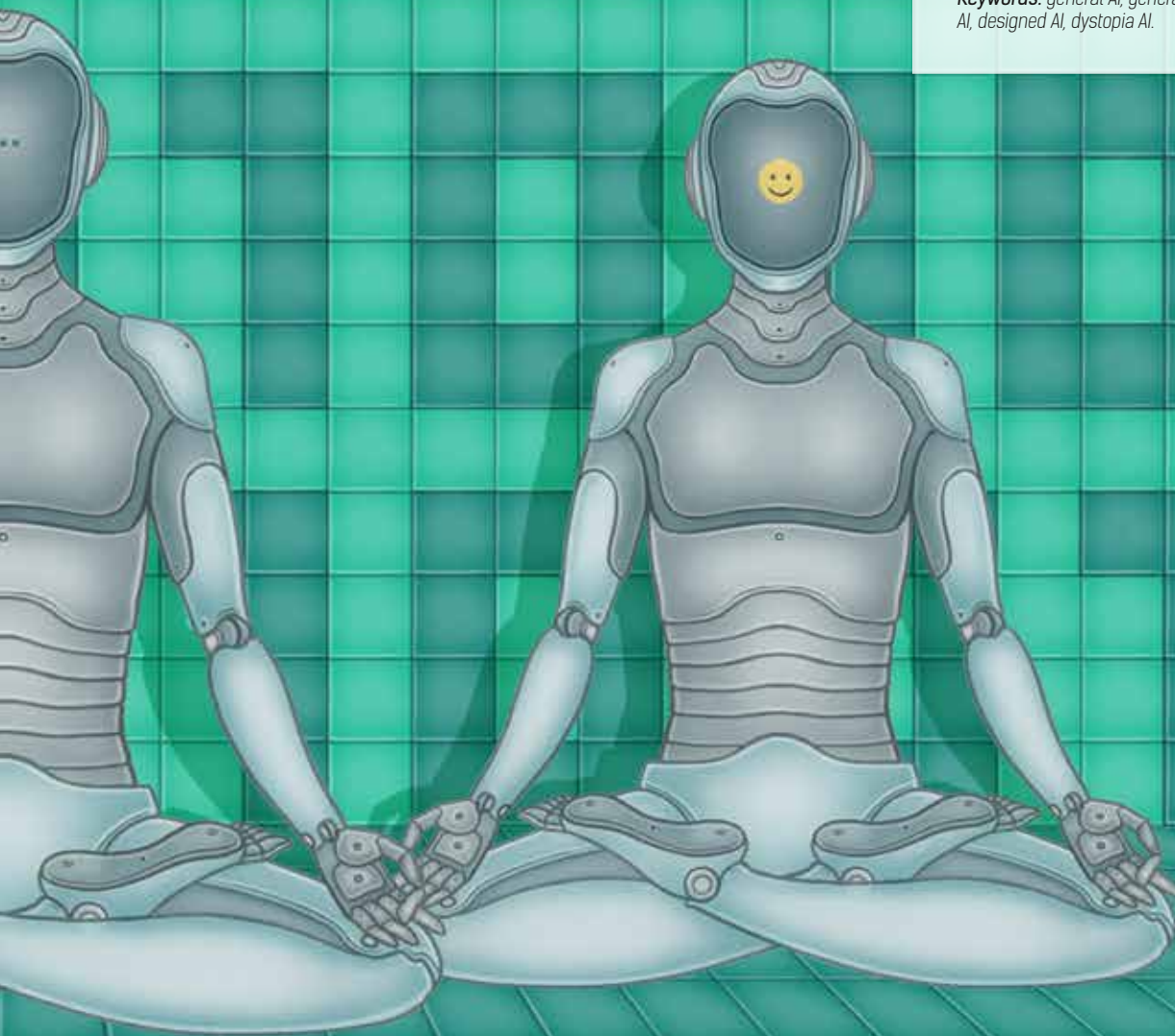
EL CAMINO HACIA UNA SEGURA Y SOS

Evolved versus designed intelligence

THE ROAD TO SAFE AND SUSTAINABLE GENERAL AI

Generative artificial intelligence has performed so well that many think that general AI is close. But is it possible to create a system as intelligent as humans? And what would such a system look like, and is there anything we can do to stop it from destroying humanity?

Keywords: general AI, generative AI, evolved AI, designed AI, dystopia AI.



IA GENERAL TENIBLE

La inteligencia artificial generativa ha conseguido un rendimiento tan bueno que muchos piensan que la IA general está cerca. Pero ¿es posible crear un sistema igual de inteligente que las personas? Y, ¿cómo sería este sistema? ¿Podemos hacer algo para que no destruya la humanidad?

Es difícil escribir algo sobre la inteligencia artificial que no se haya comentado ya. Las muchísimas aplicaciones de negocio, el carácter transformador y transversal de esta tecnología, sus retos éticos y sociales, la huella de carbón de los modelos fundacionales, las iniciativas internacionales de gobernanza de IA, su regulación, el futuro del trabajo por la creciente automatización y el impacto en la sostenibilidad de las cuentas de los países, la creciente concentración de riqueza, el aumento de la desigualdad, el reto de las noticias falsas, la creciente polarización de las sociedades, el uso en el sector militar (*killer robots*) y un largo etcétera.

Un ejemplo es este mismo número de *TELOS* con una riqueza de artículos sobre muchos aspectos de la inteligencia artificial. Y si no se ha escrito aún, en breve lo habrá escrito ChatGPT.

Qué distinta era la situación en 1993 cuando defendí mi tesis doctoral sobre cómo programar un sistema inteligente que simulaba el razonamiento diagnóstico de una persona, y era capaz de reflexionar sobre su propio razonamiento. Salvo algunos expertos académicos, pocas personas sabían qué era la inteligencia artificial y podían imaginarse su utilidad en el futuro, es decir hoy, 30 años más tarde: en 2023.

Y qué equivocado estaba pensando que el *hype* de la IA era en 2016, cuando escribí un libro blanco titulado *Surviving the AI Hype. Fundamental concepts to understand Artificial Intelligence*¹ que obtuvo más de 10.000 lecturas en pocas semanas. Sigo pensando que es importante entender los conceptos básicos de la IA, pero el *hype* de la IA sucede ahora, en 2023. En este artículo sobre la IA del futuro, voy a continuar donde acabó

Surviving the AI Hype: la inteligencia artificial general (IAG). ¿Es posible? ¿Estamos casi llegando? ¿Cómo sería?, o quizás, ¿cómo debería ser?

Cuando hablamos de la IA general, es importante ser consciente de que entramos en el espacio de las opiniones. Es decir, nadie puede pretender saber con certeza si la IA general es posible o no y cómo sería. Como explicamos Idoia Salazar y yo en nuestro libro *El mito del algoritmo: cuentos y cuentas de la inteligencia artificial* (Anaya, 2020), hay declaraciones sobre la IA que son ciertas (las cuentas), otras son (semi) falsas (los cuentos) y también hay opiniones (las más frecuentes). Además, la aparición de la IA generativa a finales de 2022 ha disparado la cantidad de opiniones acerca de que la IA general está a un paso.

Cuando leemos sobre la IA general (es decir, la posibilidad de crear un sistema inteligente que iguale a la inteligencia humana), nos pintan máquinas que son como nosotros, las personas, con todos los aspectos, buenos y malos. ¿Pero realmente debe ser así?

En este artículo, voy a debatir tres temas:

- ¿Es posible crear una inteligencia artificial general?
- ¿Cómo sería o debería ser este sistema?
- ¿Podemos hacer algo para evitar una distopía?

¿Es posible crear una inteligencia artificial general? Para debatir la IA general, me gustaría introducir dos conceptos —quizás— nuevos: la inteligencia evolucionada (*evolved intelligence*) y la inteligencia diseñada (*engineered intelligence*).

Está claro que nosotros, los humanos, somos el resultado de un

Podemos tener la esperanza de que la inteligencia artificial general, si llega, nos llegará sin intenciones de dominarnos

proceso de evolución a lo largo de miles de millones de años, en los que, poco a poco, una simple bacteria mediante mutaciones aleatorias se ha convertido en un ser humano, dando lugar también a millones de otros seres vivos. En algún momento de este proceso, lo que ahora llamamos inteligencia se debe haber iniciado. El filósofo Daniel Dennett explica este proceso en detalle en su libro *From Bacteria to Bach and Back: The Evolution of Minds*² (2017), enfocándose en la apariencia de la consciencia. Podríamos ver la evolución como un proceso de aprendizaje por retroalimentación (*reinforcement learning*) donde las mutaciones aleatorias son los aprendizajes y el éxito o fracaso de la mutación en la supervivencia la recompensa o castigo. Este proceso, después de miles de millones de años, ha dado lugar a las especies de la tierra y a nosotros, los seres humanos, con su inteligencia superior. De ahí el concepto de una inteligencia evolucionada.

Esta reflexión nos invita a preguntarnos si pudiera ser posible crear una inteligencia diseñada, es decir, no aplicar cambios aleatorios, sino diseñar los cambios en una dirección deseada, igual que el aprendizaje automático por retroalimentación.

¿Sería posible acortar el plazo drásticamente usando un método mucho más dirigido? Como señalé, nadie sabe la respuesta a esta pregunta, pero podemos opinar. Mi opinión es que no sé si es posible, pero creo que no es imposible. Otra cosa es cuántos años nos llevará... igual son miles y como humanidad nos extinguiremos antes de conocer el resultado.

¿Cómo sería o debería ser una inteligencia artificial general? Imaginémosnos que es posible diseñar un sistema de IA general. De hecho, con la IA generativa hemos dado un paso gigantesco, creando sistemas inteligentes capaces de mantener una conversación sobre cualquier asunto, responder a cualquier pregunta, resumir textos, generar textos (poemas, narrativas, notas, etcétera) según las instrucciones de una persona. Estos sistemas se basan en un modelo fundacional entrenado con una ingente cantidad de datos no estructurados y en un proceso de reentrenamiento donde se adapta el modelo fundacional a un dominio o tarea concreta.

Es importante entender, que, aunque el resultado es impresionante, estos sistemas inteligentes no entienden el contenido que manejan. Simplemente, predicen iterativamente ►►►

1 Disponible en: https://www.researchgate.net/publication/311862250_Surviving_the_AI_Hype_-_Fundamental_concepts_to_understand_Artificial_Intelligence

2 Más información en: https://en.wikipedia.org/wiki/From_Bacteria_to_Bach_and_Back

Para resolver los graves problemas de nuestra época necesitamos comportamientos lógicos del neocórtex

³ Disponible en: https://www.todostuslibros.com/libros/mil-cerebros_978-84-1107-249-6

⁴ Más información en: https://es.wikipedia.org/wiki/Prueba_del_espejo

la palabra subsiguiente de lo que ya hay (de la consulta o *prompt*, o de la última generada) según la estadística del lenguaje reflejada —aprendida— en el modelo fundacional.

La pregunta crítica en estos momentos es si podemos llegar a la IA general partiendo de la IA generativa aplicando modelos fundacionales cada vez más grandes, o si son necesarios otros varios descubrimientos científicos. Como he comentado, las posibles respuestas a esta pregunta no son más que opiniones, por bien fundamentadas que estén.

En este sentido, considero que el libro de Jeff Hawkins *Mil cerebros. Una nueva teoría de la inteligencia*³ es una revelación para un pensamiento original y científicamente fundado, aunque no probado. No se enfoca tanto en qué partes del cerebro se activan según qué tareas, sino que formula la teoría de que las personas tenemos mi-

les de cerebros que, simultáneamente, están haciendo muchas predicciones para guiar nuestra interacción en el mundo físico. Por ejemplo, si tocamos un vaso, *sabemos* qué sentir y lo percibimos como normal, a no ser que haya una rotura y nos llama la atención —porque no ha coincidido la predicción con la realidad—.

A nivel conceptual, la teoría se puede entender como aprender un mapa de una ciudad a través del contenido de las cuadrículas: una vez aprendido, desde cada cuadrícula se puede predecir el contenido de las adyacentes. Así, cada concepto o conocimiento —concreto o abstracto— se aprende a través de la creación de estos mapas por cada uno de los miles de cerebros operando en paralelo. Cuanto más sabemos de algo, más fácil es aprender conceptos relacionados porque el mapa ya está medio completo. Y, por ejemplo, por eso cuesta tanto aprender matemáticas, porque hay que empezar desde mapas casi vacíos.

Todo este proceso de aprendizaje ocurre en nuestro neocórtex, la parte más reciente del cerebro y asociada a las capacidades (lenguaje, consciencia, reflexión, sentido común, etcétera) que nos diferencian a los seres humanos de los animales. Según Hawkins, este proceso nos da muchas pistas de cómo llegar a una inteligencia artificial general: esta no emerge simplemente de la mera complejidad de un algoritmo o modelo único, sino de miles de procesos de aprendizaje y predicciones en paralelo.

Si nos movemos por una ciudad (o pensamos en un concepto) con nuestros mapas aprendidos, pode-

mos recordar que unos minutos antes estábamos en otro lugar (o pensando en un concepto relacionado), y este mero recuerdo es una forma sencilla de ser consciente, parecido al experimento con un animal, un espejo⁴ y una mancha de pintura en la piel del animal. Si el animal presta atención a la mancha, de algún modo, es consciente de sí mismo. Y así, poco a poco, se construye una consciencia cada vez más compleja a partir de millones de interacciones con el mundo físico y mental.

Si llevamos esta teoría de la inteligencia al mundo artificial, debe ser posible crear o diseñar máquinas conscientes de su propia existencia, a través de millones de interacciones y predicciones con el entorno físico y mental (algoritmos).

Evitar una distopía

Además del neocórtex —la parte nueva de nuestro cerebro, con solo un par de millones de años, que nos permite planificar, imaginar, analizar y emitir juicios—, también tenemos la parte antigua, ubicada hacia el centro de nuestro cerebro, que impulsa nuestras respuestas de supervivencia, emociones e instintos. Muchos de los problemas que tenemos hoy en nuestras sociedades son consecuencias de esta parte vieja que, en el fondo, impulsa la supervivencia de nuestros genes a través de un comportamiento egoísta. Nos ha servido muy bien, a través de la evolución de Darwin, para ser el animal superior

del planeta. Pero para resolver los graves problemas de nuestra época, necesitamos comportamientos lógicos del neocórtex y no nos sirven los comportamientos de supervivencia, reproducción, emociones e instintos. Por eso, lo que tenemos que perseguir con la IA general es intentar replicar el neocórtex y no la parte vieja de nuestro cerebro.

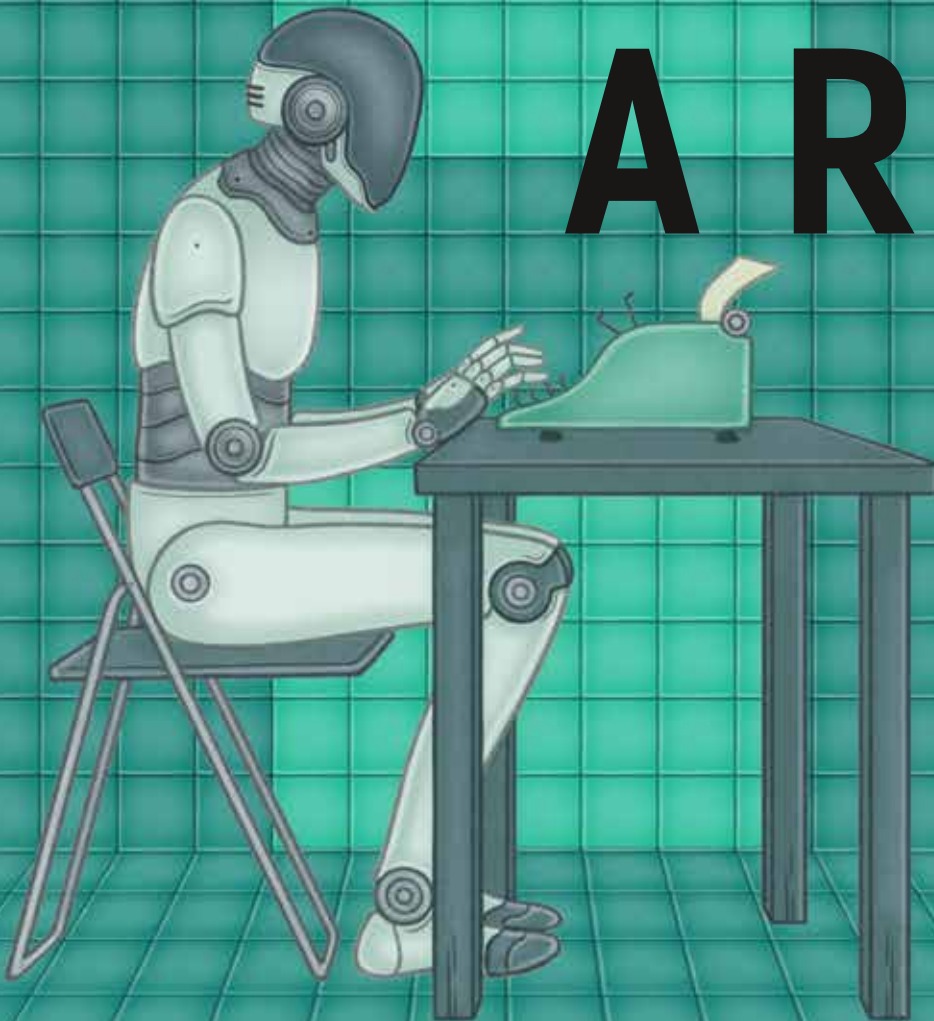
Así, podemos tener la esperanza de que la inteligencia artificial general, si llega, lo hará sin intenciones de dominarnos, sino como herramientas o instrumentos para ayudarnos a resolver los problemas difíciles que tenemos que afrontar.

Quedan muchas preguntas por resolver, como, por ejemplo, en qué medida los modelos fundacionales representan la parte vieja o la parte nueva del cerebro. Los problemas de la generación de contenido racista, violencia, odio... suenan a la parte vieja; mientras que el pensamiento lógico —con resultados modestos hasta la fecha—, al neocórtex.

O quizás no tiene sentido hacerse estas preguntas y haya que buscar otro camino hacia los “1000 cerebros artificiales”.

Bibliografía

- Benjamins, V. Richard** (2016): *Surviving the AI Hype – Fundamental concepts to understand Artificial Intelligence*. Working Paper, December 2016. LUCA Data Driven Decisions.
- Dennett, D. C.** (2017): *De las bacterias a Bach. La evolución de la mente*. Barcelona, Editorial Pasado y Presente.
- Hawkins Jeff** (2023): *Mil cerebros. Una nueva teoría de la inteligencia*. Barcelona, Tusquets Editores.



A R I S



Palabras
clave:

Aristóteles, ética,
inteligencia,
innovación, crítica.

**El futuro que
nos espera sin
olvidar el pasado
del que venimos**

QUÉ DIRÍA TÓTELES DE LA INTELIGENCIA ARTIFICIAL

El artículo aborda desde el pensamiento crítico el problema de qué es la inteligencia y cuál es el valor de la inteligencia artificial. Desde las cuatro causas aristotélicas trata sobre los problemas éticos y vitales a los que nos enfrentamos ante esta innovación social tan atractiva.



NEREA BLANCO MARAÑÓN

Licenciada en Filosofía
por la UAM. Creadora de la
plataforma Filósofers.

The future that awaits us without forgetting the past we came from
WHAT WOULD ARISTOTLE SAY ABOUT ARTIFICIAL INTELLIGENCE

The article approaches from critical thinking the problem of what is intelligence and what is the value of artificial intelligence. From the four Aristotelian causes, it deals with the ethical and vital problems we face in the light of this attractive social innovation.

Keywords: Aristotle, ethics, intelligence, innovation, criticism.

Ante las crisis y los cambios, una siempre busca pilares que no se muevan, algo a lo que aferrarse para continuar avanzando: ante una ruptura sentimental uno busca recordar quién era antes de esa relación; o ante un cambio de ciudad, busca lugares que le recuerden de dónde viene. En definitiva, es habitual —y necesario— mirar al pasado cuando las cosas cambian y el futuro parece que se nos acerca a gran velocidad. Por eso, ahora que nos vemos envueltos en una vorágine tecnológica parece más que lógico mirar cómo éramos antes de que esta gran cantidad de inventos nos rodeasen. Existe un refugio interesante al que acudir cuando todo parece que empieza a desdibujarse y a perder su esencia y ese lugar es la filosofía: un espacio en el que se han recogido tantos miedos, curiosidades, preguntas y reflexiones, que permite sentirse un lugar común a todos. Sería como ese bar que nos resulta familiar, aun estando a miles de kilómetros de nuestra casa, simplemente porque hay gente que habla nuestro mismo idioma. Y es que la filosofía puede servir para eso, para encontrarte con quien ya pensó cosas similares a las que tú te preguntas, sirve como un lugar donde encontrarte con quien habla tu misma lengua, con quien siente tu misma inquietud.

Hoy es más que interesante buscar ese refugio, y no porque en el pasado todo fuera mejor, sino porque se pueden encontrar piezas clave para comprender el presente y lo que deseamos para nuestro futuro. Por eso mismo voy a irme unos cuantos siglos atrás (hasta el IV a. de C.) para refugiarme en las enseñanzas de Aristóteles, porque pueden ser un faro para guiarnos por estos tiempos tan indeterminados. Unos tiempos en los que nos enfrentamos a una tecnología que parece adquirir inteligencia,

mientras que los humanos, de tanto pegarnos a las pantallas, parecen perder capacidades cognitivas. Y una se pregunta: ¿realmente las máquinas son inteligentes? ¿Nosotros como especie podemos dejar de serlo? ¿Tiene sentido el mundo que estamos construyendo? ¿A qué peligros nos enfrentamos? Podríamos seguir lanzando preguntas al aire... Recuperar a Aristóteles tiene sentido porque para este filósofo todo podía explicarse gracias a cuatro causas, que pueden verse como cuatro preguntas fundamentales. Las cuatro preguntas con las que cualquiera debería empezar a tratar de desmadejar cualquier crisis a la que se enfrente son: la causa material, la formal, la eficiente y la final. O lo que es lo mismo: ¿qué es esto y de qué está hecho? ¿Cuál es su esencia y cómo funciona? ¿Quién hace que funcione? y ¿para qué funciona o cómo debemos usarlo?

Si tratamos de responder a estas preguntas respecto a la inteligencia artificial (IA) podemos desenredar la madeja de la incertidumbre para mirar con calma los nudos a los que nos enfrentamos. Porque para decidir qué hacemos frente a un problema, no hay nada como analizar la totalidad de factores que lo componen. Y en este caso, nos enfrentamos a serios problemas y no solo en el plano más filosófico o abstracto, sino también ético y social.

Causa material y formal

El primer problema o nudo que hay que desenredar es que no hay una definición unitaria de lo que es la inteligencia humana. Pero para lo que nos afecta ahora, se puede describir como la capacidad de percibir información, procesarla para crear

conocimiento y, finalmente, aplicar estos conocimientos para generar una serie de habilidades para adaptarnos al entorno.

Asumiendo esta premisa, pasamos ahora a un segundo problema en la definición de inteligencia con el apellido artificial. Es obvio que este apellido indica una diferencia en su materia, pues no está hecha con material biológico, como sucede en el caso de la inteligencia humana, sino que funciona en máquinas (*software* y *hardware*). Pero hay que preguntarse: ¿acaso la inteligencia artificial funciona igual que la humana? ¿Tiene las mismas capacidades para percibir la información? ¿La procesa de forma correcta? y ¿genera las mismas habilidades que las que tenemos los seres humanos?

La definición de la UNESCO para definir la IA dice: “La inteligencia artificial son máquinas capaces de imitar determinadas funcionalidades de la inteligencia humana, incluidas características como la percepción, el aprendizaje, el razonamiento, la resolución de problemas”. De esta definición resulta interesante pararse en dos puntos: en primer lugar, en el término *imitar*, puesto que el hecho de que algo imite una cosa no quiere decir que sea igual a ella. Y, por otra parte, *que solo es capaz de simular determinadas funcionalidades y no todas ellas*. Por tanto, nuestra inteligencia resulta única y superior. Esto debería tranquilizarnos de alguna manera (al menos, de momento).

Y también es momento de analizar bien qué habilidades tenemos nosotros desarrolladas en comparación con las que tiene la IA. Nosotros, además de tener capacidad para percibir información y procesarla mediante habilidades de conceptualización, análisis y razonamiento, tenemos la capacidad crítica que

**Nuestra
inteligencia
resulta única
y superior.
Esto debería
tranquilizarnos
de alguna
manera
(al menos,
de momento)**

nos permite desarrollar habilidades de investigación y tenemos una capacidad creativa que nos permite la búsqueda de soluciones, además de tener una intencionalidad tras cada una de nuestras creaciones. El arte no es arte por la técnica, sino por lo que desea transmitir. El arte es más que colocar palabras o colores en el orden correcto. Por eso, hay que tener cuidado cuando intentamos equiparar una inteligencia con otra, pues no tienen las dos las mismas capacidades ni habilidades.

Igual que nosotros necesitamos estímulos a partir de los cuales poner en marcha nuestra capacidad intelectual, la inteligencia artificial necesita estímulos en forma de datos que ►►►

serán procesados por los algoritmos. Llamamos algoritmos a las instrucciones que tienen las máquinas para manejar y procesar esos datos. Es importante entender que las máquinas tienen una forma de aprender distinta a la nuestra. En la actualidad, funcionan gracias al llamado *machine learning* (aprendizaje automático) y uno de los métodos más utilizados es el *deep learning* (aprendizaje profundo). Esta técnica permite que la máquina aprenda gracias a las redes neuronales que imitan la estructura biológica de nuestro cerebro. Cada capa de la red neuronal está compuesta por algoritmos especializados en diferentes tareas y, cuantas más capas haya, mayor capacidad tendrá la IA. Aquí llega otro gran problema: el uso de tantas capas hace imposible que los supervisores humanos interpreten cómo la máquina logró un determinado resultado. A este fenómeno se le denomina *black box*.

Es importante entender cómo funciona una máquina para poder enfrentarnos a ella. Y lo que nos

está pasando últimamente es que nos enfrentamos a algo que no podemos comprender del todo. Y esto es peligroso, porque resultará poco controlable. Y aquello que no se puede controlar pone en riesgo lo que nos rodea... como los tornados o los tsunamis.

Por otra parte, estamos viendo que la máquina aprende a partir de lo que ya está establecido y eso le hace replicar una serie de sesgos y estereotipos manteniendo el sistema injusto en el que nos encontramos. Además, la nueva IA generativa (como ChatGPT) tiene serios problemas de credibilidad, ya que recopila muchísima información sin haberla contrastado. Y no menos importante, la IA resulta poco innovadora y muy conservadora, puesto que está muy limitada por el pasado y solo puede elaborar patrones en base a los datos que se le han dado para su entrenamiento, repitiendo así lo ya existente, incluyendo aquí los sesgos ideológicos o sociológicos. Pero es que, además, está incumpliendo leyes de propiedad intelectual cuando se apropia de obras para su aprendizaje.

**Es importante
entender cómo
funciona una
máquina para poder
enfrentarnos a ella**

Causa eficiente y final

Detrás de la IA hay empresas que se encargan de programarla. Esta sería la causa eficiente que hay que tener en mente a la hora de analizar de dónde viene esta tecnología. Y para pensar hacia dónde va (causa final) es importante pensar cómo la están diseñando y cómo se está usando.

En este punto hay que analizar las inteligencias humanas que hay detrás de las artificiales.

ChatGPT ha revolucionado nuestros días. También DALLE-E con su capacidad para crear imágenes. Ambas son tecnologías pertenecientes a la empresa OpenAI, fundada por algunas personalidades como Elon Musk, Sam Altman y Reid Hoffman, cofundador de LinkedIn. Y ahora Microsoft se ha convertido en el socio mayoritario.

Cada vez más, grandes corporaciones están entrando en este mundo de la IA generativa y es importante entender que estas tecnologías han nacido en una época de distracción tecnológica, en la que existe una economía basada en la atención. En la era del exceso informativo, la atención ha pasado a ser un bien escaso y codiciado por estas grandes empresas tecnológicas. Y como empresas, lo que buscarán será el beneficio económico por encima de cualquier otro. De este modo, en vez de ayudarnos en nuestro día a día, desvían y explotan nuestra atención apoyándose en nuestra vulnerabilidad psicológica. Resulta, por lo tanto, importante limitar las acciones que puedan llevar a cabo estas empresas en su propio beneficio y que puedan afectarnos a nosotros tanto como personas y como ciudadanos de sistemas democráticos.

Ya hemos señalado que es importante que se regulen los datos con los que estas empresas se alimentan, entender esa *black box* que nos da resultados que no acabamos de comprender. Pero, sobre todo, es importante que esta nueva tecnología no funcio-

ne solo a favor de las grandes corporaciones, que no suelen querer realmente innovaciones sociales, sino más bien conservar el *establishment*. Y, curiosamente, ya hemos visto que este tipo de tecnología parece que podría ayudar a ello. Es importante regular a las empresas que nunca se han preocupado por las cuestiones éticas, sino solo por los beneficios económicos. De este modo es como podremos crear una inteligencia que no vaya en contra de nuestros intereses.

En ese sentido, la Unión Europea está intentando frenar estos problemas y ha publicado unas *Directrices éticas para una IA fiable*, un documento basado en la necesidad de asegurar su naturaleza lícita (que cumpla todas las leyes y reglamentos aplicables), ética (que garantice el respeto de los principios y valores éticos) y robusta (desde una perspectiva técnica y teniendo en cuenta también su entorno social). En estos tiempos, los Estados deben cuidar a los ciudadanos frente a las empresas que solo busquen beneficiarse a costa de nuestro tiempo y atención.

Cómo los usuarios usemos la tecnología dependerá no solo de cada uno de nosotros, sino también de cómo esté programada esa tecnología y con qué fines. Si a mí me dan un palo y me enseñan que sirve para pintar en el suelo, es probable que solo lo use para eso. Si en cambio quieren que use el palo contra otros, porque así ellos salen beneficiados, es probable que acabe usándolo como arma.

Pongamos el foco en aquellos que fabrican la IA y no en los que nos vemos obligados a usarla. Así, estaremos más a salvo de los problemas éticos a los que nos enfrentamos.

Bibliografía

- Aristóteles (2014). *Metafísica*. Traducción de Tomás Calvo Martínez. Barcelona, Gredos.
- Benítez López, A. (2022). *Inteligencia Artificial en perspectiva*. Madrid, Independently published.
- Williams, James (2021). *Clics contra la humanidad. Libertad y resistencia en la era de la distracción tecnológica*. Barcelona, Gatopardo Ediciones.

Esta IA dista de la inteligencia humana. Establece correlaciones, no explicaciones



ANDRÉS ORTEGA
Escritor, politólogo y periodista. Director del Observatorio de las Ideas.

ARTIFICIAL, SÍ; PERO ¿INTELIG

Hay un acuerdo sobre lo que es artificial, mas no sobre lo que es inteligencia. De momento, la inteligencia artificial (IA) tiene propósitos específicos, no generales. Puede procesar información, pero no explicaciones o conocimiento, aunque pueda ayudar a generarlos. En los próximos años la mayoría de la gente trabajará para un agente autónomo, no para un humano.

This AI is far from human intelligence. It establishes correlations, not explanations.

ARTIFICIAL, YES; BUT INTELLIGENCE?

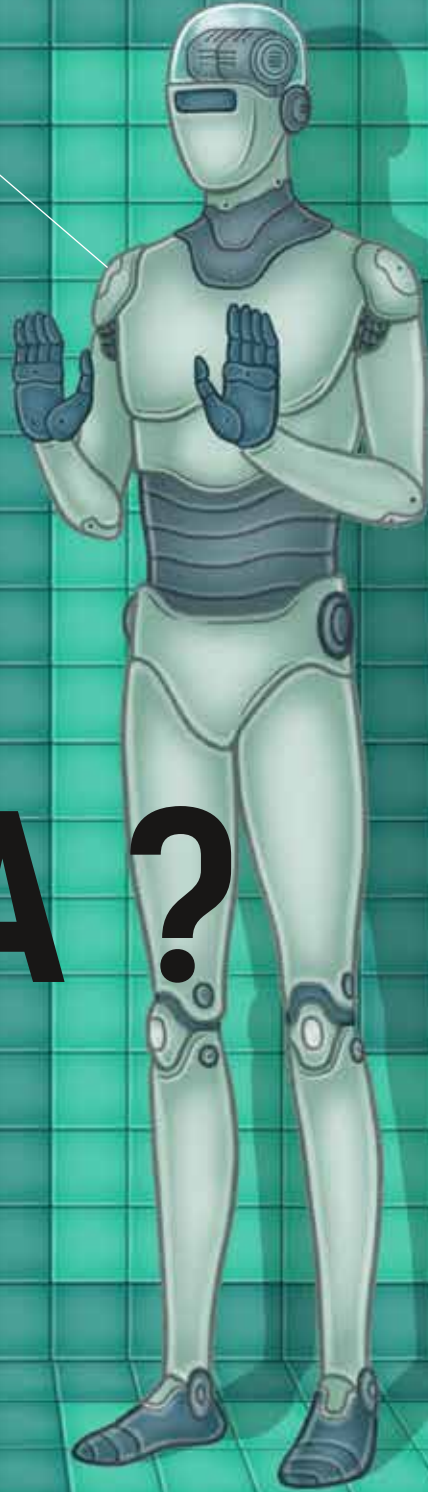
There is agreement on what is artificial, but not on what is intelligence. At the moment, artificial intelligence (AI) has specific purposes, not general ones. It can process information, but not explanations or knowledge, although it can help generate them. In the next few years most people will be working for an autonomous agent, not a human.

Keywords: artificial intelligence, generative AI, general AI, natural language, data, knowledge, statistics, correlations, explanations.

Palabras clave:
inteligencia artificial, IA generativa, IA general, lenguaje natural, datos, conocimiento, estadística, correlaciones, explicaciones.



ENCICIA?



Si. Entendemos, o creemos entender, qué significa artificial. De momento, es lo fabricado por el ser humano, un ser esencialmente técnico, como apuntara Ortega y Gasset. Ahora bien, crecientemente, lo artificial puede ser, y es, generado por las propias máquinas. Mas ¿qué es inteligencia? Una pregunta que se puede plantear desde antes de la emergencia del *Homo Sapiens*, y a la que ha prestado atención la filosofía. No hay un acuerdo sobre su definición. ¿Cómo vamos a definir, pues, inteligencia artificial? ¿Es la IA generativa —tan revolucionaria y en boga como ChatGPT— realmente inteligencia? Es engañoso, porque se basa en LLM (*Large Language Models*, o modelos grandes de lenguaje), y se le puede escribir o hablar y responde como si fuera una persona, dejando obsoleto el famoso *test de Turing*¹, aunque la máquina no sabe lo que dice ni que habla con un humano, ni siquiera lo que es un humano. Es una inteligencia estadística, que establece correlaciones. “Loros estocásticos” (aleatorios), ha llamado a estos modelos la lingüista Emily Bender. Es y va a ser cada vez más útil, pero cabe cuestionar que sea realmente inteligencia. Aunque sí es capaz de aprender estructuras de datos y volcarlas como texto, imágenes u otras formas entendibles por nosotros. Y de superarnos en muchos aspectos.

El ser humano acumula conocimiento; primero, a través de la experiencia —aunque tropiece varias veces sobre la misma piedra—, antes del habla y la escritura, la imprenta o Internet. Muchos animales, también. En la inteligencia humana van unidas la razón y la emoción. Y el cuerpo, aunque con el uso e influencia de los nuevos dispositivos digitales, se ha abierto la teoría o el campo de la mente extendida, que va más allá de nuestros atributos biológicos.

La RAE define la inteligencia como “capacidad de entender o comprender; capacidad de resolver problemas; conocimiento, comprensión, acto de entender; sentido en que se puede tomar una proposición, un dicho o una expresión; habilidad, destreza y experiencia; trato y correspondencia secreta de dos o más personas o naciones entre sí; y sustancia puramente espiritual, además de ‘servicio de inteligencia’”. Y la IA como “la disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana como el aprendizaje o el razonamiento lógico”. Definiciones pobres e incompletas.

Si le preguntamos a ChatGPT qué es inteligencia, parte de que “es un concepto complejo y multidimensional” y señala —la respuesta puede variar si se la hacemos otro día u en otro idioma, en la versión 3.5 o la 4—: “La inteligencia se refiere a la capacidad de aprender, comprender, razonar, resolver problemas y adaptarse a nuevas situaciones. Es una característica de los seres vivos que les permite procesar información y tomar decisiones basadas en ella”. En una ocasión al menos utilizó el término “mental”. Y la IA “se refiere a la capacidad de las máquinas y los sistemas informáticos para llevar a cabo tareas que normalmente requieren la inteligencia humana”. Habla de “simular procesos”.

En este debate es básico el concepto de conocimiento. ¿Lo pueden tener las máquinas? En su magnífico libro *Knowing What We Know. The Transmission of Knowledge: From Ancient Wisdom to Modern Magic (Saber lo que sabemos. La Transmisión del Conocimiento: De la sabiduría antigua a la magia moderna)*, Simon Winchester le da muchas vueltas a

¹ Nota del editor: el objetivo de este experimento es determinar si la inteligencia artificial puede imitar las respuestas humanas.

este término, para acabar resumiendo que “es simplemente información procesada, cocinada, colocada en algún tipo de contexto, algo comprendido”. Aunque añade: “El conocimiento es totalmente subjetivo: una persona puede afirmar que sabe algo, que tiene conocimiento de algo; otra puede tener conocimiento de otras cosas muy distintas y desconocer por completo lo que sabe la primera”. Las máquinas, los programas, más que conocimiento pueden compartir información. De hecho, a medida que se han ido enchufando a Internet, y no solo a textos, música o imágenes con las que se las ha alimentado —sin pagar derechos de autor—, con la técnica del aprendizaje profundo, las IA generativas ganan en información y en capacidad de llegar a conclusiones entendibles, que no a explicaciones.



Conocimiento

Por eso la IA generativa es mucho mejor que los predictores de palabras que tenían, hasta ahora, los procesadores de texto. ¿Y en conocimiento? Para algunos científicos es solo una apariencia. Más aún cuando lo que reescribe, recompone o redibuja esta IA está sacado de ingentes cantidades de textos, música e imágenes que han generado seres humanos, incluso con sus preguntas. Cada vez más las máquinas construyen sobre lo ya aprendido. Algo muy humano, ¿no? Aunque los humanos, que las inventamos, cada vez sabemos menos qué ocurre realmente dentro de estas inteligencias artificiales. Claro que nos queda mucho camino por recorrer para saber cómo funciona realmente el cerebro humano.

Se suele diferenciar entre la IA débil, de propósitos específicos,

Las máquinas,
los programas,
más que
conocimiento
pueden
compartir
información

y la fuerte, la inteligencia artificial general, la IAG, más parecida a la humana, menos estadística y más de explicaciones. Sin embargo, la IAG no existe ¿aún? —estemos atentos—. La primera está dirigida a completar tareas muy específicas como recomendar contenidos a usuarios basados en las preferencias que unos algoritmos deducen de los datos que los propios usuarios generan, comparar millones de análisis para deducir pautas para, por ejemplo, detectar un cáncer, escribir, traducir, incluso generar melodías o fotografías realistas pero irreales, o codificar, entre otras muchas. A una IA se le pide que cumpla una tarea (o varias) y lo hace. La IA no se pregunta el porqué.

La IAG tendría que ir más allá, igualar y enseguida superar a la inteligencia humana, y servir para muchas cosas, todas las que hace el ser humano, e incluso más. Y no solo sobre la base de datos masivos. Es la forma de aprender de, por ejemplo, los ▶▶▶



niños. Algunos creen que esto que llaman “singularidad” está a la vuelta de la esquina, y borran la diferencia entre generativa y general como si se tratase de un progreso paulatino, cuando son dos cosas muy diferentes. Otros (Chomsky, Minsky...) consideran que, desde las premisas actuales, nunca se alcanzará. Esa IAG sería realmente igual que una persona, con diferencias solo en la velocidad, en la memoria, etcétera, pero ¿sería capaz de generar lógicas nuevas? Hay dudas entre los especialistas.

Quizás la cuestión está mal planteada al comparar siempre la inteligencia técnica con la humana. Combinada con los avances en la robótica y en la biología (simbiosis que promete una nueva revolución de enorme alcance) esta inteligencia fuerte puede llevar a cabo tareas no ya que realizara el ser humano, sino que a este ni se le ocurriría que fueran posibles.

Algo diferente de lo humano. ¿Lo entenderemos? Sí, si la inteligencia es universal, lo que no está garantizado.

Hay varias líneas de reflexión e investigación sobre esto y la inteligencia universal en las ciencias de la computación y más allá. La inteligencia universal se refiere a “la tendencia intrínseca de las cosas a autoorganizarse y co-evolucionar hacia formas cada vez más complejas, intrincadamente entrelazadas y mutuamente compatibles,” según Christian Hugo Hoffmann. Nuestra inteligencia humana no sería más que una manifestación de esa dinámica universal. La artificial general, también.

Hay un peligro inminente: el de los agentes autónomos, movidos por las IA, que pueden ser muy positivos, incluso creativos (aunque sin consciencia de serlo), pero también destructivos. Matt Schlicht calcula que en los próximos dos a cinco años

La cuestión está mal planteada al comparar siempre la inteligencia técnica con la humana

la mayoría de la gente trabajará para un agente autónomo en lugar de para un humano. Puede ser especialmente peligroso no solo en lo social, sino también en el terreno militar. Las armas autónomas con IA, ni siquiera fuerte sino débil, pueden ganar en rapidez de decisión al mejor soldado o general para tomar decisiones letales, sobre todo si la otra parte dispone también de ella. Como indican algunos desarrolladores de NVIDIA, empresa esencial en el diseño de chips para la IA avanzada, “en principio, un agente podría utilizarse para automatizar cualquier proceso. Una vez que estos agentes sean altamente sofisticados y fiables, es fácil imaginar un crecimiento exponencial de la automatización en todos los campos e industrias”.

Como apunta Winchester, muchos de los avances digitales del último medio siglo (la calculadora, el procesador de textos, el corrector ortográfico, el GPS, Google, Internet, y ahora ChatGPT) han sido inventos de una innegable inteligencia humana. Pero si Google, ahora con IA generativa —después vendrán otras más avanzadas—, nos hace más capaces, a la vez nos hace más tontos o muy diferentes. El cerebro puede servir para menos de momento, pues hay enormes posibilidades de simbiosis. No se retiene lo que se lee en pantalla, porque se sabe que

se puede recuperar buscándolo. No es necesario recordar números de teléfono; los tenemos en el móvil. Lo que produce amnesia digital. En *Un mundo feliz*, Aldous Huxley planteó que los humanos podrían llegar a desear tener dispositivos que nos ayudasen a *no* pensar. ¿Y si la inteligencia humana acaba siendo menos útil, incluso para los humanos?

El neurocientífico Daeyeol Lee considera que “la investigación de las diferencias entre la inteligencia humana y la artificial puede ayudarnos a comprender mejor nuestra relación con la tecnología, y el porvenir de esta. Toda la información estará en las máquinas, pero no todo el conocimiento, mientras que nuestra ignorancia debe ser necesariamente infinita”. Hay algo a lo que quizás la IA, por muy fuerte que sea, no llegará: la sabiduría. Como decía el poeta inglés William Cowper: “el Conocimiento tiene el orgullo de haber aprendido tanto; la Sabiduría, la humildad de no saber más”. Esta IA, no tiene ni orgullo, ni humildad.

Bibliografía

Catanzaro, B. (2023): “Generative AI Demystified”. NVIDIA. Disponible en: <https://www.nvidia.com/en-us/on-demand/session/gtcspring23-s52089/?playlistId=playlist-e6fdf3ae-afb1-4dc3-ab88-c5f342d4927e>

Mendivil Giró, J. L. (2023): “Un loro estocástico en la habitación china: ¿qué nos enseña ChatGPT sobre la mente humana?” en *Letras Libres*. Disponible en: <https://letraslibres.com/revista/un-loro-estocastico-en-la-habitacion-china-que-nos-enseña-chatgpt-sobre-la-mente-humana/01/07/2023/>

Ortega, A. (mayo de 2023): “Nuevas inteligencias” en revista *Telos*. Disponible en: <https://telos.fundaciontelefonica.com/nuevas-inteligencias/>

Schlicht, M. (2023): “0”. Disponible en: <https://www.mattprd.com/p/the-complete-beginners-guide-to-autonomous-agents> y <https://www.youtube.com/watch?v=1AB41kHUJZA&t=29s>

Winchester, S. (2023): *Knowing What We Know: The Transmission of Knowledge: From Ancient Wisdom to Modern Magic*. Nueva York, Harper/HarperCollins Publishers.

Una historia de bases de datos, colaboración creativa y automatización



DOREEN A. RÍOS
Curadora e investigadora
independiente. Fundadora de
la plataforma [ANTI]MATERIA.

MÁQUINI INTELIGENTES

Cómo ha evolucionado el uso de la IA para fines creativos y cuál es el futuro de estas herramientas.
Realizamos un análisis histórico del uso de la IA en el arte, desde una perspectiva crítica.

A history of databases, creative collaboration and automation

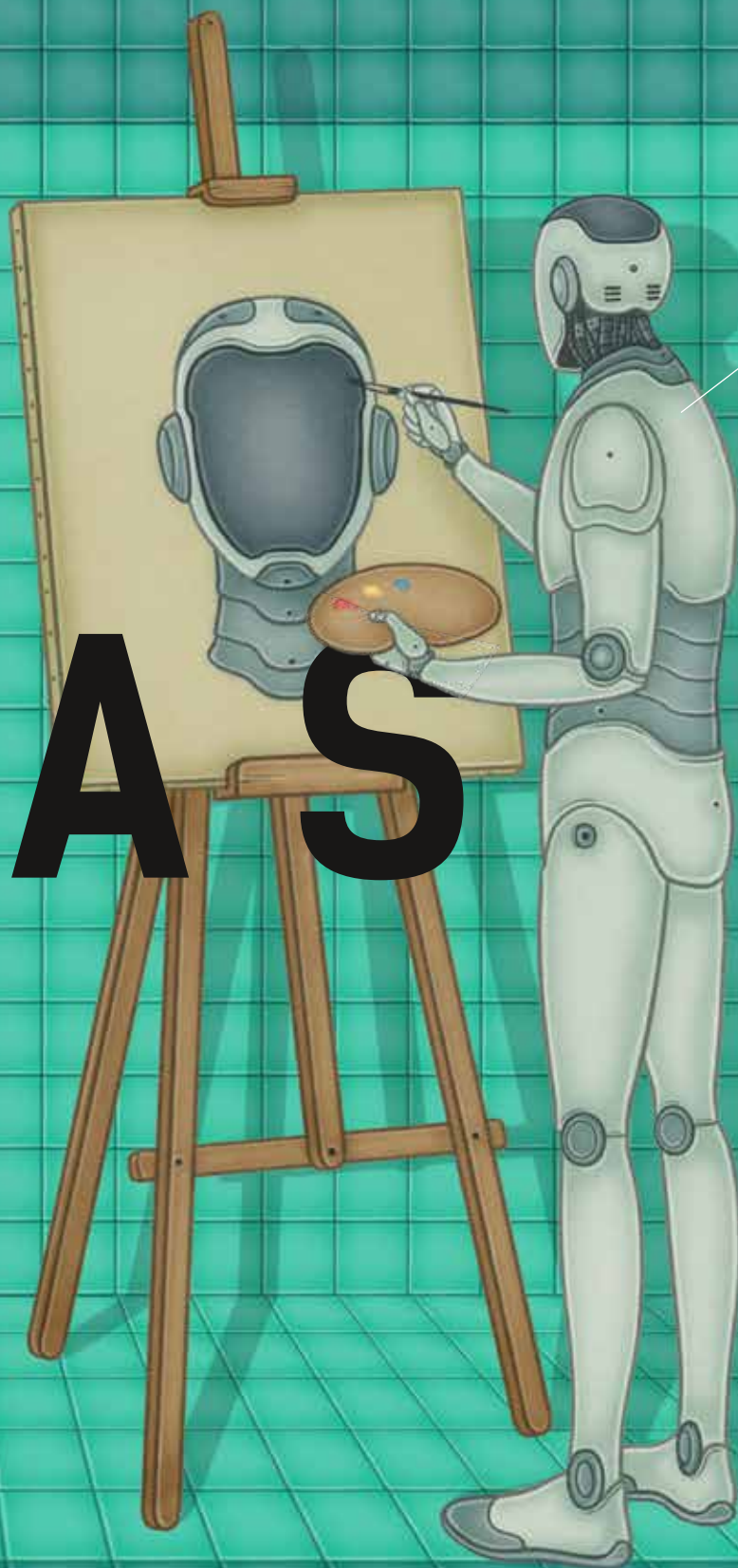
SMART MACHINES

How has the use of AI for creative purposes evolved and what is the future of these tools.

We conduct a historical analysis of the use of AI in art, from a critical perspective.

Keywords: artificial intelligence, art, creativity, art history, authorship, automation.

NAS



Palabras clave:
inteligencia artificial, arte, creatividad, historia del arte, autoría, automatización.



En los últimos dos años, el debate sobre el uso de la inteligencia artificial (IA) en la creación artística ha ganado espacio. Herramientas como DALL-E, Midjourney, Stable Diffusion y ChatGPT se han convertido en referentes en este ámbito. Sin embargo, la presencia de la IA en la creatividad tiene una trayectoria histórica de al menos 50 años, por lo que vale la pena plantear los siguientes interrogantes: ¿qué ha cambiado en el contexto actual? y ¿cómo han evolucionado estas herramientas? En este artículo, exploraré la evolución de la IA en el arte desde una perspectiva histórica, revelando las memorias detrás de su desarrollo y su vínculo con el contexto contemporáneo.

En 1950, Alan Turing presentó su informe *Maquinaria computacional e Inteligencia*¹, en el que abordó la construcción de máquinas inteligentes y planteó la posibilidad de probar dicha inteligencia. Aunque el término *inteligencia artificial* no se menciona directamente en el informe, Turing exploró conceptos fundamentales relacionados con el desarrollo e imaginación de máquinas inteligentes.

Cinco años después, adjudicó el primer proyecto de investigación de verano sobre inteligencia artificial de la Universidad de Dartmouth / Dartmouth Summer Research Project on Artificial Intelligence (DSPRAI), que marcó un hito en el desarrollo de esta disciplina. Durante este acto, conocido como la Conferencia de Dartmouth², se produjo un momen-

to icónico: cuando John McCarthy acuñó el término inteligencia artificial. McCarthy definió la IA como “la ciencia e ingeniería enfocada en la creación de máquinas capaces de simular de manera precisa todos los aspectos del aprendizaje y los rasgos de la inteligencia”. Y desde entonces, esta área de desarrollo, investigación y experimentación ha dado varios giros hasta llegar al día de hoy.

La historia de la IA en el arte se remonta a principios de la década de los 60, cuando surgieron las primeras investigaciones sobre el uso de computadoras para producir arte. En ese entonces, los primeros ejemplos de arte generado por IA se enfocaban en la creación de patrones visuales, explorando la estética matemática y las composiciones algorítmicas. Estas primeras manifestaciones artísticas por medio de IA se basaban en las definiciones y conceptos de la época.

Un ejemplo lo encontramos en la obra de la pionera Vera Molnár³. La artista comenzó a experimentar con el uso de computadoras y dio un paso importante en su carrera artística al empezar a crear dibujos utilizando un plóter. El uso de la computadora en el proceso creativo de Molnár permitió una mayor precisión y, a la vez, introdujo como concepto la interacción entre el orden y el caos, creando intencionalmente un “1 % de desorden” para permitir que un factor de azar sistemáticamente determinado influyera en sus creaciones.

Paralelamente a estas exploraciones artísticas, también se llevaron a cabo investigaciones en el campo de la robótica / IA, y fue en este contexto que Jean Tinguely⁴ dejó su huella como escultor y creador de autómatas. En el caso de Tinguely, su obra se destacó por su enfoque en la creación de esculturas cinéticas robóticas. Sus esculturas eran construcciones complejas compuestas por alambres y láminas de metal, que se movían y/o giraban a diferentes velocidades. Tinguely buscaba transmitir un mensaje irónico y provocador, desafiando los conceptos tradicionales de arte estático y proponiendo una visión del arte automatizado.

De esta manera, tanto Molnár como Tinguely exploraron la colaboración con máquinas inteligentes. Ambos inmersos en un contexto fértil dentro del que compartían con otros artistas interesados en prácticas similares como Nam June Paik, Frieder Nake o Georg Nees, así como con curadores como Jasia Reichhardt, quien gestionó la mítica exposición *Cybernetic Serendipity*⁵.

Autoría y materialidades

En la década de los 70, el uso y la experimentación con computadoras comenzaron a desempeñar un papel cada vez más importante en la investigación artística. En este contexto, Harold Cohen desarrolló *AARON*⁶,

un *software* capaz de generar dibujos complejos. Estas creaciones eran posteriormente transferidas al papel mediante un plóter y, finalmente, intervenidas manualmente por el artista. A través de reglas y algoritmos, *AARON* daba vida a piezas inicialmente abstractas y fluidas. Sin embargo, con el tiempo, evolucionó, para crear representaciones figurativas de rocas, plantas y personas. Harold Cohen colaboró con *AARON* a lo largo de toda su vida, perfeccionando el programa y desafiando las concepciones esta- ➤➤➤

1 Disponible en: <http://xamanek.izt.uam.mx/map/cursos/Turing-Pensar.pdf>

2 Disponible en: https://www.dsi.uclm.es/personal/miguelfgraciani/mikicurri/Docencia/InteligenciaArtificial0607/web_IA/Documentacion/Trabajos/Introduccion/Conferencia%20de%20Dartmouth.pps

3 Más información en: <http://www.veramolnar.com/>

4 Más información en: <https://www.tinguely.ch/>

5 Más información en: <https://proyectoidis.org/179/>

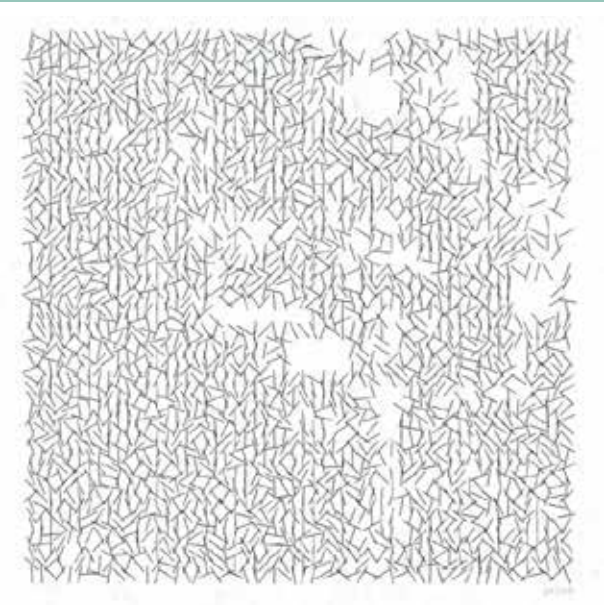
6 Más información en: <https://theconversation.com/aaron-vida-y-obra-de-la-primera-inteligencia-artificial-creativa-192281>

La historia del arte y la IA revelan una trayectoria de exploración y colaboración entre la creatividad humana y las máquinas

2



1



3



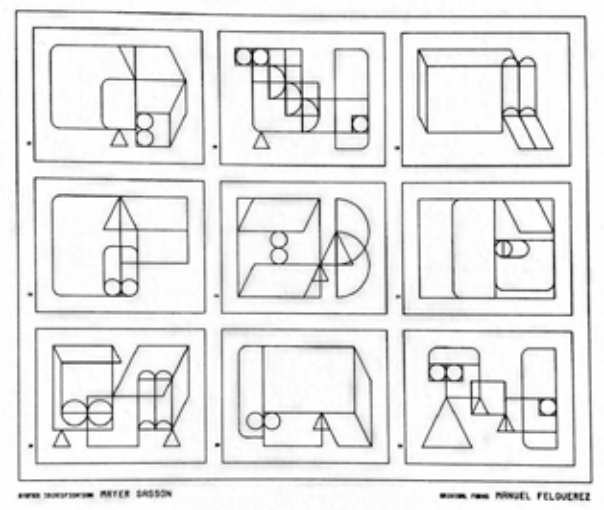
APLICACIÓN DE LA IA EN EL MUNDO DEL ARTE

1. Lynn Hershman Leeson, Captura de pantalla de *Agent Ruby*, 1998 - 2002. Obtenida de VoCA Journal.
2. Vera Molnár, *Interruptions*, 1969. Victoria & Albert Museum.
3. Kira Xonorika, *Meditation*, 2023. Cortesía de la artista.
4. Kira Xonorika, *Symbiosis*, 2023. Cortesía de la artista.
5. Manuel Felguérez y Mayer Sasson, Diagramas producidos por *La Máquina Estética*, 1983.

4



5



Nuestras visiones del futuro han sido moldeadas por la literatura de ciencia ficción

blecidas sobre la interacción entre el artista y la inteligencia artificial.

Por otro lado, Manuel Felguérez también se planteaba cómo las computadoras podían ser aliadas en su proceso artístico. En 1975, inició una investigación en el Laboratory for Computer Graphics and Spatial Analysis y el Carpenter Center for the Visual Arts de la Universidad de Harvard. En colaboración con el ingeniero Mayer Sasson, se propusieron crear un *software* llamado *La Máquina Estética*⁷, que permitiera la producción infinita de nuevos ideogramas y diseños con las características estilísticas que definían la obra plástica del artista. Esta iniciativa buscaba explorar la simbiosis entre la mente humana y la mente computarizada, dando lugar a nuevas conversaciones sobre la implementación de procesos automatizados en el arte.

Tanto en el caso de Cohen como en el de Felguérez, surgieron debates que siguen siendo relevantes, especialmente en relación con la autoría y la replicación de las obras generadas por computadora. Es interesante destacar que, aunque el concepto de inteligencia artificial era casi una cualidad inseparable al hablar de computadoras, este periodo también resaltó la vulnerabilidad de estas máquinas. A medida que se exploraban diversos usos de la computadora, quedaba claro que

los sistemas no eran infalibles. Esto comenzó a desdibujar la idea de que las computadoras eran sinónimo de inteligencia y perfección, y a establecer que algunas tareas sí se ajustaban a esa categoría mientras que otras no. Haciendo que la inteligencia artificial dejara de ser sinónimo directo de computadora.

La era de hielo de la IA

En la década de los 80, el surgimiento de las computadoras personales y su integración en diversos ámbitos de la vida cotidiana provocó una pausa simbólica en la investigación de la inteligencia artificial como disciplina. Durante esta etapa, se tomaron en cuenta las vulnerabilidades y limitaciones que se habían evidenciado en las máquinas inteligentes en la década anterior. En lugar de seguir buscando una única forma de inteligencia artificial que fuera equivalente a la humana, los grupos de investigación comenzaron a dividirse y a enfocarse en aspectos específicos.

Hacia la década de los años 90, Internet se convirtió en un gran partea-guas que, entre muchas otras cosas, multiplica el uso de las computadoras personales. La llegada de Internet también dio lugar a un discurso enriquecido sobre la virtualidad y el ►►►

7 Más información: <https://www.revistadelainiversidad.mx/download/6903309b-bfb9-40e2-82fc-8b6164b4a9ce?filename=la-maquina-estetica>

8 Disponible en: <http://agentrubby.sfmoma.org/>

⁹ Nota del editor: acrónimo de *Generative Adversarial Nets*, en español: Redes Generativas Antagónicas.

¹⁰ Más informaciónen: <https://linktr.ee/xonorika/>

espacio pantalla. En paralelo a estos desarrollos, surgieron nuevas formas de aprendizaje y pedagogías a través de la cultura del tutorial. Los artistas y creadores comenzaron a compartir sus conocimientos y técnicas a través de tutoriales en línea, lo que permitió un intercambio multidireccional de información y habilidades.

En este contexto, la artista Lynn Hershman Leeson desarrolló la pieza *Agent Ruby*⁸ (1998-2002), un agente *web* artificial e inteligente que se forma y refleja a partir de sus encuentros y experiencias con los usuarios. La pieza podía descargarse y aparecer en una parte de los escritorios de los usuarios, siendo compatible con diversas plataformas como PC, Mac y Palm Pilots. A través de *chats*, la agente Ruby interactúa con los usuarios, recuerda sus preguntas y, en última instancia, puede reconocer su

voz, además de tener estados de ánimo y emociones. Esta pieza planteó interrogantes sobre la relación entre lo humano y lo artificial, explorando temas como la identidad y la interacción en el contexto digital. *Agent Ruby* representa una manifestación de la inteligencia artificial aplicada al arte, expandiendo los límites de la experiencia artística y la relación entre el usuario y la tecnología. Desde aquí, una vez más, observamos líneas de pensamiento que siguen vigentes.

En el contexto actual, nos encontramos con una abundancia de producción de datos, así como su comodificación, en la mayoría de los casos opaca, que ha hecho estallar varios debates en torno a la legalidad, consentimiento y privacidad. En este contexto, sobre todo apuntando a la transición del entrenamiento de GANs⁹ hacia los sistemas generativos como DALL-E o ChatGPT, surge la pregunta de qué es lo que las IA pueden reconocer en una imagen o texto y qué es lo que queda invisible.

La llegada del *promptismo* ha reintroducido debates sobre la vulnerabilidad de la tecnología, las IA pueden analizar y reconocer patrones, pero ¿qué ocurre con los matices sutiles, las emociones subyacentes o los significados simbólicos que se encuentran en las imágenes, sonidos, videos, textos, etcétera? Estas son preguntas que nos invitan a reflexionar sobre la naturaleza de la percepción y la interpretación humana en contraste con las capacidades de las máquinas.

Aquí, y con miras al futuro, resulta clave observar prácticas conscientes de todos los procesos que implican trabajar con IA como los de Kira

A medida que se exploraban diversos usos de la computadora, quedaba claro que los sistemas no eran infalibles

Xonorika¹⁰. En su obra explora cómo nuestras visiones del futuro han sido moldeadas por la literatura de ciencia ficción occidental, que a menudo lo retrata como un concepto distante y lineal que perpetúa un paradigma binario. Sin embargo, Xonorika adopta el protopianismo, una perspectiva que hace hincapié en la colaboración, la experimentación y la regeneración, una visión que florece a través de la cocreación y la coautoría entre inteligencias humanas y no humanas. En su obra se inspira en figuras mitológicas que emergen de las profundidades del agua, encarnando habilidades sobrenaturales y arquetipos de feminidad. Su intención es forjar una conexión con estas narraciones, reimaginando sus cuerpos e historias e infundiéndoles un significado renovado. Así, su obra se convierte en un poderoso catalizador para el diálogo y la introspección, animándonos a reimaginar y adoptar una comprensión más integradora de la identidad, la cultura y el potencial de la expresión artística en la era digital desde una perspectiva transversal.

Conclusión

La historia del arte y la inteligencia artificial revela una trayectoria de exploración y colaboración entre la creatividad humana y las máquinas. A lo largo de este viaje, se han planteado preguntas fundamentales sobre la autoría, la interacción entre humanos y máquinas, y los límites de la percepción y la interpretación a la vez que surgen cuestionamien-

tos sobre las implicaciones asimétricas y opacas del desarrollo de las IA. Es crucial cuestionar y reflexionar sobre la relación entre lo humano y lo artificial, explorar nuevas formas de creatividad y fomentar futuros más colaborativos y diversos, considerando siempre las implicaciones éticas, políticas y sociales de la IA en el arte. Como bien apunta Karin Ohlenschläger, “el arte puede explorar y hacer visibles las aristas de la IA en todas las esferas de la vida política, social, económica y cultural. Puede plantear reflexiones y debates de manera más transversal y polifacética, conectando hechos y ficciones, más allá de lo que determinan los intereses del mercado o la política. En cuanto a la IA y el arte, esta unión no solo es interesante, sino también necesaria: pensar, visualizar o interpretar de forma disruptiva” (2021, pág. 27).

Bibliografía

- Felguérez, M.; Sasson, M. (1983): *La Máquina Estética*. México, Universidad Nacional Autónoma de México.
- Hellar, M. (2019): *The Artist and the Technologist*, in *VoCA Journal*, March 7, 2019. Disponible en: <https://journal.voca.network/the-artist-and-the-technologist/>.
- Stocker, G; Jandl, M.; J. Hirsch, A. (Ed.) (2021): *The Practice of Art and AI*. Linz, Austria, Ars Electronica.
- Sunder, A. (2023): *Spectral Ghosts: Deliberate Images, Intended Metaphors*. Alemania.
- Tsing, A. (2015): *Friction: An ethnography of global connection*. Princeton, NJ, Princeton University Press.
- Turing, A. (1950): *Computing Machinery and Intelligence*. *Mind* (236): 433–460. Oxford, Oxford University Press on behalf of the Mind Association.



1

¿Cómo describiría el estado actual de la inteligencia artificial y qué avances considera que veremos en el futuro inmediato?

2

¿Qué desafíos éticos ve asociados al avance de la IA y cómo cree que debemos abordarlos?

3

¿Cómo afecta la IA a la privacidad y la seguridad de los datos de las personas?

CON LAS RESPUESTAS DE
RAMÓN LÓPEZ DE MÁNTARAS
NEREA LUIS MINGUEZA
IDOIA SALAZAR
JORDI TORRES
ÓSCAR CORCHO
ELENA GIL LIZASOAIN

Preguntamos a un grupo de especialistas en inteligencia artificial sobre las principales incógnitas que despierta su uso y desarrollo. Nos han señalado cuáles son los mayores desafíos éticos, de seguridad, privacidad y regulación, haciendo hincapié en la necesidad de supervisión humana y señalando el avance exponencial que seguiremos viendo en el progreso de la IA.

Spanish experts address the ethical, security and regulatory challenges of artificial intelligence

CHALLENGES AND OPPORTUNITIES OF THE AI

We asked a group of specialists in artificial intelligence about the main questions raised by its use and development. They pointed out the major ethical, security, privacy and regulatory challenges, emphasizing the need for human oversight and noting the exponential advance we will continue to see in the progress of AI.

Keywords: artificial intelligence, cybersecurity, privacy, regulation, oversight, appropriate technology.

El desarrollo de la inteligencia artificial en las últimas décadas ha sido muy rápido debido al crecimiento de la capacidad computacional, la disponibilidad de grandes conjuntos de datos y avances en algoritmos. Esto ha llevado a la creación de múltiples aplicaciones y sistemas de IA en diversas áreas. Se trata de un campo en constante crecimiento con un gran potencial para transformar diferentes aspectos de nuestras vidas. Por este motivo, hemos contactado con una gran variedad de expertos en IA de diferentes ámbitos de trabajo. Hemos preguntado a Ramón López de Mántaras, uno de los pioneros de la IA en España; Nerea Luis, cofundadora de T3chFest; Jordi Torres, profesor del grado de IA en la UPC; Idoia Salazar, cofundadora de OdiseIA; Óscar Corcho, director académico de AI4Gov y Elena Gil, directora en Telefónica Tech. Es importante abordar este tema desde diferentes puntos de vista, explicando de manera comprensible las implicaciones, beneficios y desafíos de la IA. Con ello buscamos ayudar a contrarrestar la desinformación y los mitos sobre la inteligencia artificial, brindando información precisa basada en la experiencia de los entrevistados. “Lo fundamental es la necesidad de involucrar a la sociedad en su conjunto en la toma de decisiones sobre la IA para garantizar que esta tecnología sea utilizada de manera responsable”, remarca Jordi Torres en este reportaje.

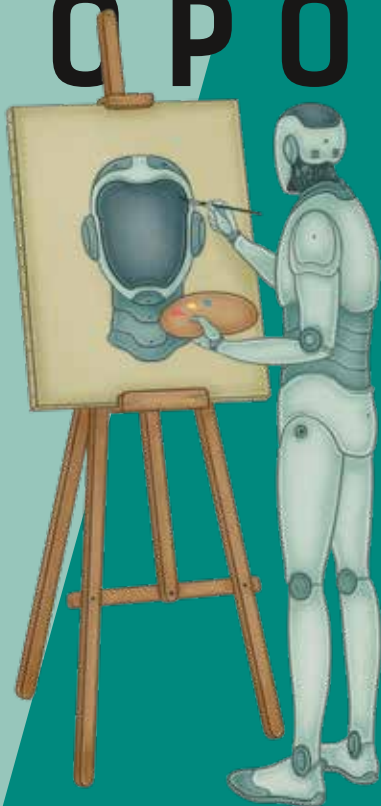
**Expertos
españoles
abordan
los retos éticos,
de seguridad
y regulación
de la inteligencia
artificial**

DESAFÍOS

Y

OPORTUNIDADES

DE LA IA



Palabras clave:
Inteligencia artificial, ciberseguridad,
privacidad, regulación, supervisión,
tecnología adecuada.

POR MARISOL SALES

CUESTIONARIO

1. ¿Cómo describiría el estado actual de la inteligencia artificial y qué avances considera que veremos en el futuro inmediato?



Ramón López de Mántaras

ES EL PIONERO DE LA INTELIGENCIA ARTIFICIAL EN ESPAÑA. PROFESOR INVESTIGADOR DEL CENTRO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS (CSIC) Y DIRECTOR DEL INSTITUTO DE INVESTIGACIÓN DE INTELIGENCIA ARTIFICIAL (IIIA)

“LA INTELIGENCIA ARTIFICIAL NO ES NI TAN INTELIGENTE, NI TAN ARTIFICIAL”

1 Este *boom* del último año sobre todo viene dado por lo que llamamos “inteligencia artificial generativa”. Esta inteligencia artificial generativa no es nueva. Cuando empezó a desarrollarse parecía interesante, pero no que fuera a explotar de esta manera ni que se iban a meter las grandes empresas tecnológicas como ha ocurrido. En mi opinión, esto tiene aspectos positivos y aspectos negativos.

Yo sigo pensando que las grandes preguntas científicas que hay detrás de la inteligencia artificial, como si podrá haber inteligencias artificiales verdaderamente generales, y el avance hacia una inteligencia artificial general que verdaderamente entienda el mundo no lo veo por ninguna parte. Muchos defienden que la inteligencia artificial generativa es la antesala para conseguir una inteligencia artificial general, pero yo no lo pienso en absoluto. La generalidad de la inteligencia artificial generativa es muy limitada. Con estos sistemas preentrenados no estamos ganando generalidad. Estas inteligencias artificiales no nos acercan

a la respuesta a la pregunta de si habrá inteligencias artificiales que verdaderamente entiendan. Por ello me interesan más bien poco. A mí lo que me interesa es hacer ciencia.

2 Desde hace mucho, tenemos una serie de desafíos éticos como, por ejemplo, que un sistema de inteligencia artificial sea una caja negra y que no se sepa cómo funciona y que no tenga capacidad explicativa. Plantea una serie de cuestiones no sobre su uso ni sobre su impacto, sino sobre la prudencia necesaria a la hora de desplegarlo.

Con la inteligencia artificial generativa nos aparecen problemas, digamos, “nuevos” que hace unos años no se veían tan claramente y es el tema de generar todas las falsedades que se hacen y esa facilidad con la que puedes falsificar personas o hacerte pasar por personas, ya sea de manera textual o audiovisual. Si seguimos con la inteligencia artificial generativa le veo difícil solución.

La manipulación se acentúa mucho más con esta capacidad de generar *fakes*. Llegará el día en el que en un juicio se presen-

tará un vídeo o una imagen de alguien cometiendo un delito infraganti y no se podrá utilizar porque no habrá manera de saber si es verdadero o falso. Esto es muy grave porque hace explotar la convivencia social e incluso la democracia.

Además, otro problema ético que se presenta son los millones de personas en países del sur global que llaman, como Filipinas o Pakistán, que están horas y horas al día limpiando lo que hacen los modelos de lenguaje por 0,50€/hora e, incluso, algunos trabajos que no se llegan ni a pagar. También modelando contenido o entrenando estas inteligencias artificiales generativas o etiquetando imágenes indicándole que esto es un árbol o esto es una persona o esto es una farola. Todo esto hecho a mano por una persona. Como digo, no estamos a la vuelta de la esquina de una inteligencia artificial general porque cada vez necesitamos más intervención humana. He llegado a la conclusión de que la inteligencia artificial no es ni tan inteligente, ni tan artificial.

3 Esto ya era un problema incluso sin la inteligencia

artificial. Hace ya bastantes años que parece que mucha gente ha renunciado a la privacidad. Yo creo que fue en el momento que aparecen los móviles y los sistemas que hacen recomendaciones o un perfil de tus gustos y preferencias que están diseñados para que te enganches y que además son supuestamente gratis. Hay que moverse desde la sociedad civil y organizarse para enfrentarse a todo lo que se nos está viniendo encima con esto de la inteligencia artificial generativa. Por primera vez en mi vida, viendo la balanza, soy más pesimista que optimista porque, ¿en qué mal momento salió la inteligencia artificial generativa? A nivel científico no sirve de nada prácticamente y a nivel práctico no veo muy claro quién está ganando con esta IA.



**Nerea
Luis
Mingueza**

DOCTORA EN
CIENCIAS DE LA
COMPUTACIÓN Y
COFUNDADORA DE
T3CHFEST. TRABAJA
COMO INGENIERA
EN INTELIGENCIA
ARTIFICIAL EN
SINGULAR

“LA RECOGIDA DE DATOS PARA EL ENTRENAMIENTO DE LOS MODELOS ES UN PRIMER DESAFÍO”

1 Vivimos un momento en el que la inteligencia artificial se está integrando en las herramientas que conocemos tales como aquellas relacionadas con la ofimática, los gestores de email, de videollamadas o el diseño gráfico. Esto hará que rápidamente aprendamos a sacarle su potencial en las tareas cotidianas y que los mismos modelos de IA aprendan y personalicen su tono, estilo y uso de la palabra en función de los documentos previos que hayamos escrito. Sin duda la personalización será una pieza clave para simplificar los *prompts*. Por otro lado, en la generación de imágenes, técnicas como Gaussian Splatting están dando muy buenos resultados para recrear escenas 2D > 3D. Sin duda se convertirá en otra de las tendencias próximas.

2 Hay un primer desafío alrededor de la recogida de datos para el entrenamiento de los modelos “grandes” o “fundacionales” (como la familia GPT o Palm2). En algunos casos se están viendo avances como en Adobe Firefly, donde solo entrenan con imágenes libres, pero aún hay mucha ca-

suística alrededor de cómo pedir permiso o cómo reportar sobre las fuentes de datos que se utilizan para entrenar a los modelos.

Otro punto es la propiedad intelectual de los textos e imágenes generadas, tanto para reclamar autoría como para reclamar un mal uso de las herramientas de IA si se utilizan para generar desinformación o *deepfakes*.

3 Existen multitud de servicios web ahora mismo que utilizan por detrás modelos de inteligencia artificial generativa. Generar imágenes, hacer preguntas sobre un PDF, resumir reuniones... En todos estos servicios, a la hora de usarlos, tenemos que preocuparnos de saber si los datos que volcamos los utilizan para “reentrenar” o si por el contrario solo los mantienen un tiempo *online* con mero fin consultivo. De esta forma sabremos a qué riesgo nos estamos exponiendo cuando los utilizamos.

Poco a poco van saliendo licencias de uso personal, para empresas y entornos en la nube donde se respeta la privacidad y la seguridad de los datos, eso sí.



Jordi Torres

CATEDRÁTICO DE LA UNIVERSIDAD POLITÉCNICA DE CATALUÑA (UPC). ES AUTOR DE MÁS DE 200 PUBLICACIONES DE INVESTIGACIÓN, Y LIBROS SOBRE COMPUTACIÓN DE ALTO RENDIMIENTO Y DE INTELIGENCIA ARTIFICIAL

“LA HUMANIDAD SE BENEFICIARÁ ENORMEMENTE DE LA INTELIGENCIA ARTIFICIAL ACTUAL”

1 A pesar que actualmente la población en general tiene la percepción de que la inteligencia artificial (IA) ha alcanzado competencias casi humanas, la realidad es que los algoritmos en que se basan las IA no tienen inteligencia, sino solo “habilidades sin comprensión”, en el sentido de que son algoritmos que pueden llegar a ser muy hábiles realizando tareas específicas, pero sin comprender absolutamente nada de lo que están haciendo, ni cuentan con el llamado “sentido común” de la inteligencia humana que la hace tan diversa, rica, espontánea y maravillosamente imprevisible.

A pesar de que los sistemas actuales de IA todavía carecen de capacidades importantes, estamos inmersos en una carrera para crear sistemas de IA cada vez más potentes. Estos últimos años este progreso de las IA se ha sustentado en una aproximación de “fuerza bruta”, con la creación de algoritmos cada vez más grandes, entrenados en supercomputadores cada vez más potentes, procesando cada vez más y más datos. Pero, aunque este crecimiento exponencial de recursos dedicados a este pro-

pósito no es sostenible por mucho más tiempo, el ritmo del progreso de la IA podría volver a sorprendernos.

De todos modos, de momento, en la comunidad científica se coincide en que para lograr una IA mucho más inteligente que se asemeje a la inteligencia humana se necesitarán innovaciones disruptivas, cuyas características son inciertas y nadie sabe si realmente se producirán algún día.

2 Que no estemos cerca de disponer de una IA próxima a la inteligencia humana no quita que el actual progreso de la IA esté generando nuevos desafíos éticos que debemos abordar. Por un lado, tenemos desafíos en su desarrollo, como la tendencia de los algoritmos de IA a heredar sesgos de los datos de entrenamiento, lo que puede llevar a decisiones discriminatorias o incorrectas. Esto, a su vez, plantea la cuestión de la responsabilidad en caso de fallos de una IA. ¿Deberíamos considerar al creador del algoritmo o al usuario como responsable? Es esencial establecer marcos legales y éticos que traten la cuestión de los errores o daños causados por la IA.

Por otro lado, encontramos desafíos relacionados con su uso. **La automatización impulsada por la IA puede llevar a la reestructuración y pérdida de empleos en diversos sectores económicos, lo que plantea interrogantes éticos sobre cómo garantizar una transición justa hacia una economía basada en la IA.** Además, existe la preocupación sobre el uso indebido de la IA, como su aplicación en ciberseguridad, la producción de noticias falsas o el desarrollo y uso de armas totalmente autónomas en conflictos militares, lo cual genera inquietudes profundas.

La humanidad se beneficiará enormemente de la IA actual, a pesar de ser limitada, pero no hay duda de que su uso presenta muchos riesgos. Por ello, al igual que ha ocurrido en otras áreas importantes en nuestra vida, es necesario establecer un marco regulador consensuado, en el más amplio sentido del término: leyes, reglas, protocolos de actuación, autorregulación o sensibilización ciudadana, entre otras. A nadie le gusta ser regulado, pero todo lo que puede entrañar un potencial peligro para la sociedad lo debe

de estar. Espero que las aplicaciones específicas de IA estén también reguladas y, en algunos casos, prohibidas si fuese necesario para poder así mitigar sus riesgos.

3 El avance de la IA tiene importantes implicaciones en la privacidad y la seguridad de los datos personales. Esto se debe a que la IA depende en gran medida del análisis de datos para funcionar de manera efectiva lo que resulta en la recopilación masiva de información personal. Por ejemplo, los algoritmos en que se basan los motores de búsqueda en Internet, o simplemente nuestro teléfono móvil, conocen muy bien, gracias al acceso a cantidades masivas de información que generamos, nuestras preferencias y pueden llegar a inferir cómo pensamos.

Esto amplía sustancialmente el poder y el alcance de las empresas o gobiernos que controlan estos datos y pueden socavar nuestras expectativas de privacidad personal. La capacidad de la IA para analizar grandes conjuntos de datos conlleva un riesgo potencial de violación de la privacidad, ya que la información podría ser utilizada de manera in-

Lo desafíos éticos requieren una colaboración entre gobiernos, empresas y desarrolladores

apropiada. Por ejemplo, los sistemas de IA que utilizan el reconocimiento facial pueden identificar a las personas de manera precisa, una tecnología útil en ciertos contextos, pero que plantea preocupaciones sobre la invasión de la privacidad y un posible uso indebido. Además, los sistemas de IA son susceptibles a ataques, con el riesgo de robo de datos y la generación de información falsa por parte de algoritmos maliciosos.

En resumen, estos desafíos éticos o de privacidad son abordables, pero requieren una colaboración decidida entre gobiernos, empresas y desarrolladores de IA para establecer regulaciones sólidas que promuevan un uso y desarrollo responsables de la IA. Sin embargo, lo más fundamental es la necesidad de involucrar a la sociedad en su conjunto en la toma de decisiones sobre la IA para garantizar que esta tecnología sea utilizada de manera responsable, beneficiosa y en línea con los valores humanos. Esto implica el compromiso individual de todos nosotros en adquirir un mayor entendimiento de esta revolución tecnológica, de manera que podamos tomar decisiones informadas.



CUESTIONARIO

1. ¿Cómo describiría el estado actual de la inteligencia artificial y qué avances considera que veremos en el futuro inmediato?



**Idoia
Salazar**

—
**PERIODISTA,
 INVESTIGADORA
 Y PROFESORA.
 PRESIDENTA Y
 CODIRECTORA DEL
 ÁREA DE ÉTICA Y
 RESPONSABILIDAD
 DE ODISEIA**

“DEBERÍA DE HABER MÁS HERRAMIENTAS QUE CREEN CONCIENCIA EN EL USUARIO”

1 Yo diría que la inteligencia artificial está teniendo ya, y tendrá aún más en el futuro, un avance exponencial. Cada vez son más las empresas que apuestan por esa tecnología porque no es una opción no usarla. Hay muchísima investigación, muchísima inversión y en los próximos años vamos a ver un aumento significativo en el uso y desarrollo de estos sistemas de inteligencia artificial.

A nivel tecnológico posiblemente se siga avanzando bastante para intentar hacer bancos públicos para que sea más fácil, no solamente para las grandes empresas de los datos, sino también para pymes y poder optar a esta tecnología para usarla para la eficiencia de procesos y otro tipo de cuestiones.

También veo un aumento bastante grande en la implementación de IA generativa a todos los niveles y es algo que va a suponer una revolución también a nivel social que tendremos que aprender a sobrellevar para que el impacto sea positivo.

A nivel de regulación estamos asistiendo también a un avance bastante significativo. Se ha visto la ne-

cesidad de que, a la vez que evoluciona la tecnología, también aumentemos significativamente el esfuerzo para hacer una regulación competente y poder prevenir los riesgos que vienen asociados a esta tecnología. Cada vez hay más empresas que están dispuestas a adoptar estas normas éticas y regulatorias, una vez que entre en vigor la regulación de inteligencia artificial para intentar prevenir posibles impactos negativos.

2 Los desafíos éticos que vienen asociados al avance de esta tecnología son los mismos que tenemos que tener asociados a la ética humana. Uno de los principales es la supervisión humana. Como sabemos, esta tecnología puede actuar de manera autónoma y los humanos siempre tendemos a tener confianza excesiva en ella. **Debemos poner ciertas medidas técnicas y de supervisión humana, sobre todo en casos en el que la acción de la IA pueda afectar a los derechos fundamentales.** Puede causar un problema si nosotros no realizamos esa supervisión adecuada.

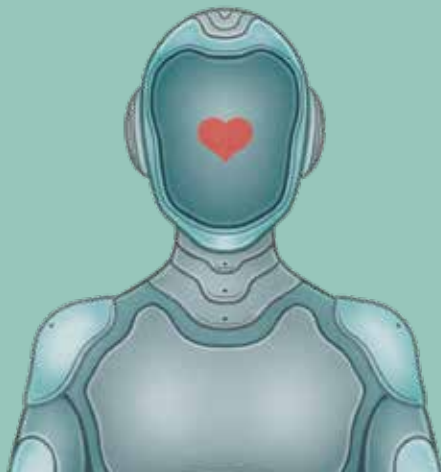
Por otro lado, hay cuestiones relativas a posibles

casos de discriminación si no se ha hecho un eficiente uso, limpieza y revisión de los datos, algo que en temas de transparencia es muy importante.

También las cuestiones de la explicabilidad son fundamentales en la lógica algorítmica de ver cómo se ha tomado la decisión. La transparencia en cuestiones como la IA generativa es muy importante para que se vea que el que ha hecho el *output* es la máquina y no la persona.

Por supuesto también hay motivos específicos de ciberseguridad que son fundamentales. Es muy importante que todas las empresas empiecen a adoptar estos criterios éticos, que al final serán parecidos a los que van a implementar la regulación de inteligencia artificial, para que consigamos tener desde el diseño un uso adecuado de estos sistemas de inteligencia artificial.

3 En este caso, tenemos que saber que, a pesar de que los sistemas de inteligencia artificial funcionan con datos, la mayoría de los datos que usa son personales, por lo que, obviamente, puede llegar a suponer un





Óscar Coreho

CATEDRÁTICO
DE LA UPM EN EL
DEPARTAMENTO
DE INTELIGENCIA
ARTIFICIAL
Y DIRECTOR
ACADÉMICO DEL
MÁSTER EUROPEO
AI4GOV (ARTIFICIAL
INTELLIGENCE FOR
PUBLIC SERVICES)

problema de privacidad. Por ello son muy importantes las labores de consentimiento. Todo esto está regulado por la Ley General de Protección de Datos europea. Así que, en principio, si se cumple con la regulación europea en protección de datos, este tema debería de estar zanjado.

Sin embargo, hay otro tipo de temas como, por ejemplo, la IA usada en motores de recomendación con los que se debe de crear una mayor conciencia en el usuario, porque al final la máquina está usando datos personales que el usuario va generando. Si no ha habido un consentimiento previo es una intrusión en el propio día a día de los usuarios, que obviamente es para mejorar la personalización del servicio. En cualquier caso, creo que debería de haber más herramientas que creen conciencia en el usuario de que se va a realizar esa labor de gestión de sus datos de alguna forma para que se haga más visible esa información.

“EL PRINCIPAL DESAFÍO ÉTICO ES EL POSIBLE SESGO DE LOS MODELOS QUE SE GENERAN”

1 En estos últimos años se han producido avances muy considerables en la inteligencia artificial, sobre todo en las áreas de aprendizaje automático y en todo lo relacionado con el lenguaje, debido a que cada vez hemos tenido más datos y documentos disponibles para poder entrenar modelos, y que ha sido posible hacerlo con las infraestructuras tecnológicas de las que disponemos.

Quizás el avance más significativo de estos últimos años haya sido la aparición de la IA generativa, con los grandes modelos de lenguaje como uno de los pilares básicos.

¿Qué veremos en el futuro inmediato? Bueno, lo vemos ya... Una gran cantidad de aplicaciones que hacen uso de estos modelos y de estas capacidades para poder hacer nuestro trabajo aún más eficiente. Pero no debemos olvidarnos también del “Good old AI”. Se siguen utilizando técnicas tradicionales para resolver una gran cantidad de problemas que ya eran conocidos.

2 El principal desafío ético es el del posible sesgo de los modelos que se generan, basado fundamentalmente en el sesgo que siempre

puede existir en los datos que se utilizan para el entrenamiento. **Necesitamos investigación en técnicas cada vez más robustas para explicar los resultados obtenidos por modelos que actúan como cajas negras, y una buena regulación que no necesariamente frene los avances,** pero que nos genere seguridad en que no se producen estos sesgos.

3 No creo que este sea un problema que se vea agravado por la IA puesto que ya éramos conscientes de los posibles riesgos en este sentido cuando se puso el foco en la regulación de protección de datos, por ejemplo. Por supuesto, si cualquiera de los modelos que usamos se ha entrenado con datos personales, o los usuarios no son conscientes de que están entregando datos personales al interactuar con sistemas como ChatGPT, podemos poner en riesgo esa privacidad.

CUESTIONARIO

1. ¿Cómo describiría el estado actual de la inteligencia artificial y qué avances considera que veremos en el futuro inmediato?

2. ¿Qué desafíos éticos ve asociados al avance de la IA y cómo cree que debemos abordarlos? 3. ¿Cómo afecta la IA a la privacidad y la seguridad de los datos de las personas?



Elena Gil Lizasoain

DIRECTORA DE LA
UNIDAD DE NEGOCIO
DE INTELIGENCIA
ARTIFICIAL Y BIG
DATA EN TELEFÓNICA
TECH, ES LICENCIADA
EN CIENCIAS
ECONÓMICAS
Y EMPRESARIALES EN
LA UNIVERSIDAD DE
DEUSTO Y MBA POR
EL MIT SLOAN SCHOOL
OF MANAGEMENT

“LA IA TODAVÍA CARECE DE LA VERDADERA COMPRENSIÓN HUMANA”

1 La inteligencia artificial se encuentra en una fase de aceleración y democratización, impulsada recientemente por los avances de la IA generativa y modelos de lenguaje de gran escala (LLM). Es una realidad y una parte esencial de la actividad diaria de las personas, así como un motor de innovación en muchos sectores. En el futuro más inmediato, veremos avances significativos en la aplicación de la IA en ámbitos como la medicina personalizada de precisión, automatización inteligente de tareas o el desarrollo sostenible de la movilidad y las ciudades. Sin embargo, aún queda mucho por investigar y desarrollar ya que, a pesar de la impresionante capacidad de la IA para detectar patrones y automatizar tareas específicas en un dominio concreto, todavía carece de la verdadera comprensión humana.

Como ha ocurrido hasta ahora, adoptaremos nuevas herramientas de IA en nuestro día a día de manera natural. La inteligencia artificial lleva muchos años estando presente en nuestras vidas como cuando usamos un GPS, utilizamos un asistente virtual o compramos en una tienda

online. La principal diferencia, ahora, es que se están amplificando sus capacidades y la formación es una herramienta esencial para adaptarse a esta nueva realidad. Aparecen nuevos perfiles con formación específica en esta tecnología y los empleos preexistentes se verán transformados. Es una tendencia imparable y, como sociedad, debemos prepararnos para extraer el máximo valor de la IA de manera responsable, adaptando la regulación y la fuerza de trabajo.

2 **Como toda tecnología disruptiva, la adopción de la IA plantea desafíos que deben ser abordados. Existen desafíos en el desarrollo de las herramientas de IA como el respeto a la privacidad, la protección de la propiedad intelectual de los datos de entrenamiento, la eliminación de sesgos o, incluso, el consumo energético asociado.**

Adicionalmente, se plantean retos éticos relacionados con la desigualdad, desinformación, suplantación de la identidad o el fraude. Resolverlos requiere un enfoque multidisciplinario y es responsabilidad de todos. Es necesario desarrollar un

marco que promueva, a nivel global, el desarrollo de la IA responsable y que ponga un énfasis especial en la transparencia y la rendición de cuentas.

3 La inteligencia artificial tiene múltiples efectos para la privacidad y seguridad. Por un lado, ha incrementado la vulnerabilidad por la recopilación y almacenamiento de más información sobre las personas, como el robo de datos y la suplantación de la identidad. Sin embargo, también se ha producido una mayor concienciación de la importancia de proteger la privacidad de la información y, tanto las empresas como los gobiernos, están invirtiendo y utilizando de manera creciente IA para protegerse. Esto incluye el cifrado de datos, la autenticación de usuarios y el monitoreo de redes y sistemas para detectar posibles amenazas. La normativa europea está siendo pionera en materia de protección de la privacidad y seguridad de los datos de las personas.

¿AÚN NO NOS SIGUES?

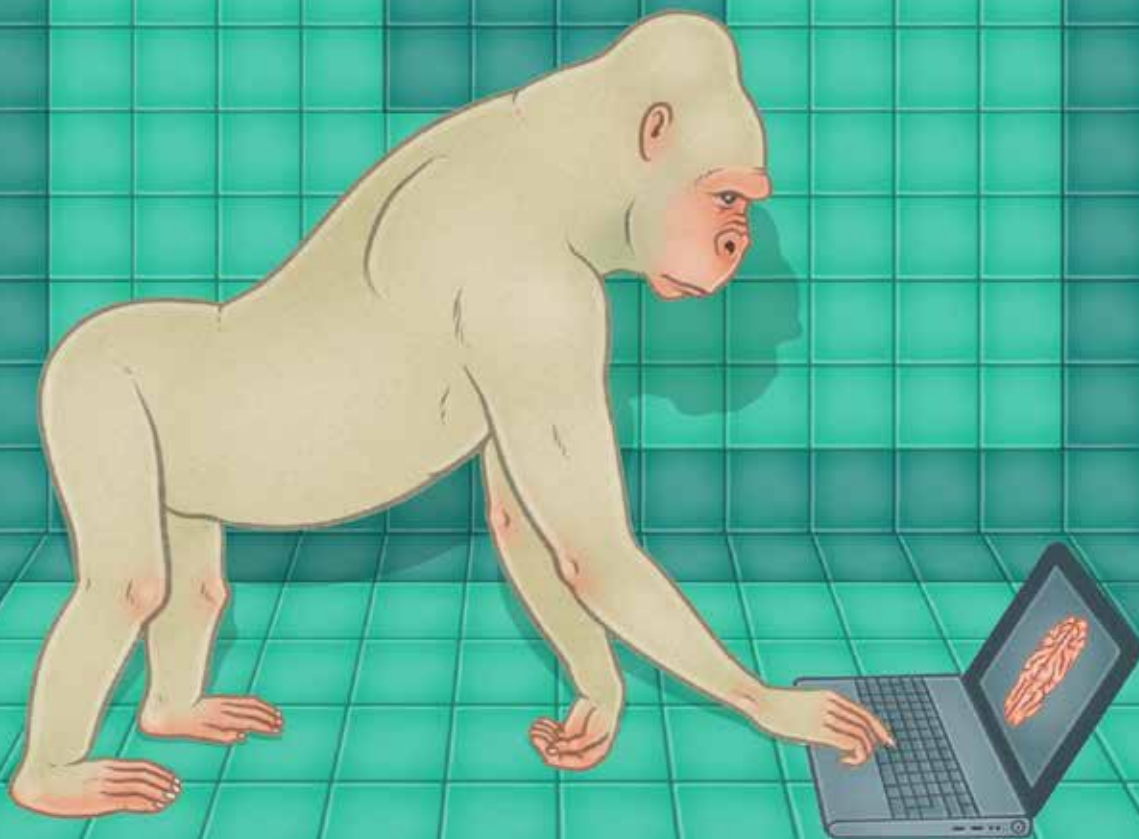
Únete a la
#ComunidadTELOS



T | e | l | o | s

Revista de Pensamiento, Sociedad y Tecnología

telos.fundaciontelefonica.com



How to build a biodigital mind

AI AND THE NEXT HUMAN MENTAL SYSTEMS

Homo sapiens, as a mental model, is evolutionarily obsolete. The next humans will need a biodigital functioning. Artificial intelligence (AI) is the first technology that modifies the mental space and will be key in designing the model of future humans.

Keywords: *artificial intelligence, evolution, mental systems, machines, digitization.*

Palabras clave:
inteligencia artificial, evolución, sistemas mentales, máquinas, digitalización.



GERARDO IVÁN TUDURÍ
Investigador, escritor y artista visual.
Fundador del ADMS (Análisis Digital de Sistemas Mentales)

**Cómo
construir
una mente
biodigital**

LA IA Y LOS PRÓXIMOS SISTEMAS MENTALES HUMANOS

Homo sapiens, como modelo mental, evolutivamente, ha quedado obsoleto.
Los próximos humanos necesitarán un funcionamiento biodigital.
La inteligencia artificial (IA) es la primera tecnología que modifica el espacio mental y será clave en el diseño del modelo de los futuros humanos.

Homo sapiens, fue —es todavía— la marca intelectual evolutiva que hemos lucido por siglos. Una moda del pensamiento que compone su pasarela evolutiva con un conjunto de conceptos científico culturales —y sus accesorios mentales—, determinando nuestra autopercepción evolutiva: somos y nos sentimos *Homo sapiens*, somos los seres vivos más inteligentes de la creación.

Visto así, *Homo sapiens* resulta un monopolio científico cultural —llamémosle humanista— elaborado, divulgado e introducido a millones de personas por los sistemas educativos del mundo, por imperativo categórico y durante siglos.

La ficción humanista nos enseñó el gesto de la bruja del cuento de Blancanieves:

-¡Espejito evolutivo!... ¿quién es el más inteligente de toda la creación?

Y, durante los siglos que dura el humanismo, creímos escuchar una única respuesta:

-Está claro, *Homo sapiens*, ¡el más inteligente eres tú!

En este gesto de autoexaltación podría resumirse la historia del humanismo.

Pero, ¿quién contestaba a esa pregunta? ¿la Vida, la Materia Viviente, el Universo? O ¿era una psicofonía científico cultural del propio *Homo sapiens* que evitaba la verdadera respuesta, movido por su típica ansiedad evolutiva?

Sesenta y siete años después de ¡aquel verano en Dartmouth!, donde un grupo de diez científicos acudieron a unas semanas de trabajo y reflexión, dando lugar al evento que

acabaría convertido en el nacimiento oficial de la inteligencia artificial, algo nos confunde:

- ¡Espejito evolutivo! ... ¿quién es el más inteligente de toda la creación?

- Depende, *Homo sapiens*, ¿a qué aspecto de la inteligencia te refieres? ¡Sé más preciso, por favor!

Tanto el espejito de la evolución como una IA podrían estar respondiendo a la pregunta.

Tenemos un problema

El humanismo, como todos los *ismos*, configura un contenido, una visión, que determina la autopercepción del humano. Una narrativa que, según su fuerza de instalación mental, moldea el comportamiento. La visión de vegetarianismo, es el inicio para que alguien se convierta en vegetariano. La del budismo, abre la posibilidad de un comportamiento budista.

Nuestro modelo mental se ha estado configurando culturalmente.

Para comprender esto podemos valernos de las tres fases con las que Max Tegmark se refiere a la evolución de la Vida:

- Vida 1.0 (fase biológica): su *hardware* y *software* son fruto de la evolución.
- Vida 2.0 (fase cultural): su *hardware* es fruto de la evolución; diseña parte de su *software*.
- Vida 3.0 (fase tecnológica): diseña tanto su *hardware* como su *software*.

Bajo este esquema podemos pensar que el humanismo nos configura el modelo mental culturalmente, dentro de la fase de Vida 2.0.

El problema a resolver que se nos presenta, es que la IA opera cada vez con mayor impacto en nuestra “auto-percepción cotidiana y evolutiva” y la modifica, sin valerse de contenido cultural mental alguno. Le bastan datos y algoritmos, estadística y cálculos para hacerlo.

El impacto de la IA

Compongamos el escenario:

Un discreto verano de 1956.

Un lugar. Hanover. Universidad Dartmouth College.

Un rutinario encuentro de trabajo entre 10 científicos de los más relevantes del mundo.

Un proyecto: “Propuesta para el proyecto de investigación de verano de Dartmouth sobre Inteligencia artificial”.

Un punto de partida: “La conjetura de que todos los aspectos del aprendizaje o cualquier otra característica de la inteligencia puede, en principio, describirse con tanta precisión que puede fabricarse una máquina para simularlo”.

La IA, entonces, nace del esfuerzo de humanos por externalizar sus procesos mentales de aprendizaje e inteligencia en máquinas computadoras para que puedan simularlo.

¿Qué ha sucedido después? Que, si el humano ha volcado en una máquina una secuencia de sus procesos mentales, esta los aprende y, al cabo de un tiempo, los ejecuta mejor. Cuando el

humano vuelve a relacionarse con la máquina, indefectiblemente, siente en su mente el impacto de un proceso externalizado que funciona mejor afuera que en su propia cabeza. Su mente se asombrará de esto y tendrá que saber gestionar esa externalización. Habrá replicado un proceso de su mente que ahora lo experimenta más rápido, más eficaz, mejor.

Este es el punto de disrupción evolutivo entre mente y máquina en el que estamos.

Llevamos 67 años desde aquel inicio y cada vez externalizamos más procesos mentales y a la vez, estamos digitalizando cualquier tipo de proceso natural bajo la misma lógica. Tenemos a un humano con un modelo mental *Homo sapiens* inmerso en máquinas que han aprendido a hacer muchas cosas mentales mejor que él. ¿Acaso no era el objetivo? Sin ►►►

**Homo Sapiens, ¡el
más inteligente
eres tú! En
este gesto de
autoexaltación
podría resumirse
la historia del
humanismo**

El humano biodigital no tendrá temor alguno de comportarse con la eficacia de una máquina y con la sensibilidad de un humano

embargo, el humano sigue entrando en pánico por el mismo hecho: externalizar procesos y ver que las máquinas lo hacen mejor. ¡Cuántos más procesos de los que le daban supremacía mental ceda a las máquinas, más supremacía mental sentirá que pierde! Su miedo evolutivo a la desaparición no es real; solo está dándose cuenta de que su ficción humanista de supremacía inteligente, si él mismo sigue externalizando procesos y las máquinas mejorando su rendimiento, acabará con su supremacía. Y, por lo que vemos, es la dirección actual del desarrollo tecnológico de la IA.

Es para esto para lo que debería prepararse y no para fantasiosas historias de robots que acaban con la vida humana en el planeta.

Podemos cambiar el modelo mental *Homo sapiens* por uno que integre naturalmente los procesos externalizados a las máquinas, que no pararán de aumentar. Pero tiene que enten-

derse en clave evolutiva y eso no se le da bien al humanismo que suele mirarse más el ombligo y enredarse en su historia inmediata.

Una anécdota de ello. La misma Conferencia de Dartmouth constituyó un hecho trascendente para la evolución por ser el origen de la IA. Pero algunos sus participantes, durante aquel verano, quizá tampoco tenían mucha perspectiva evolutiva. Marvin Minsky, uno de los padres de la IA y participantes de la conferencia, reconocerá por un lado que algunos de ellos llegaron allí “con grandes ideas de cómo hacer cosas con las máquinas, pero los únicos... que realmente presentaron un programa real que fuera capaz de realizar algo similar a lo que constituye un proceso humano superior fueron Newell y Simon (otros dos participantes); mientras los demás hablábamos, ellos habían alcanzado ya muchas cosas”.

Y, por su parte, Newell y Simon, comentarían que, efectivamente, “la conferencia se dispersó en todas direcciones. Ellos no querían recibir nada de nosotros y nosotros, por supuesto, nada de ellos, pues teníamos algo que enseñarles”.

Sesenta y siete años después, estamos frente a un momento de trascendencia evolutiva, pero volvemos a ignorarlo centrándonos en un imaginario de pánico. Debemos tomar perspectiva de siglo y generar herramientas que ofrezcan soluciones desde un posicionamiento evolutivo.

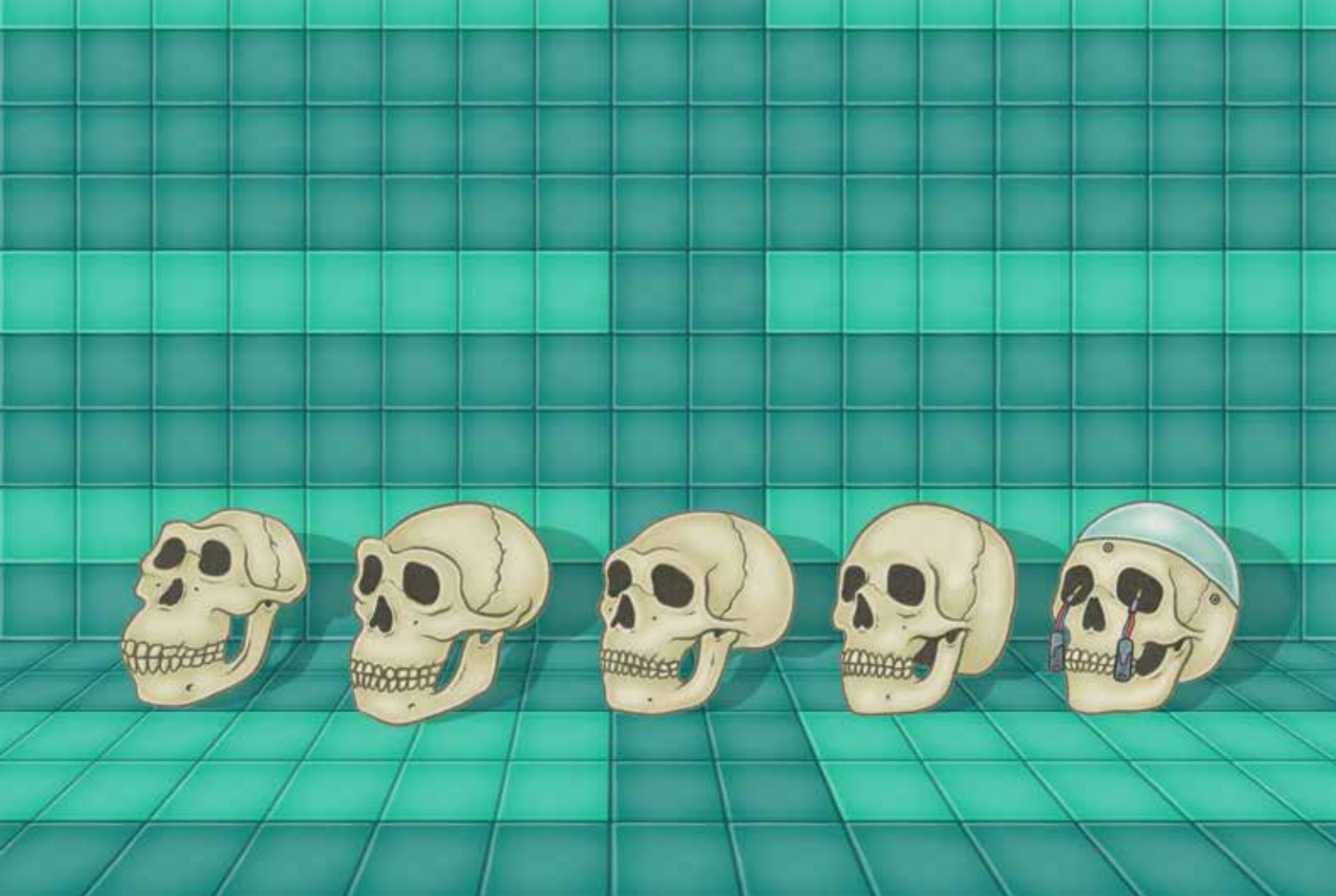
Llevo casi una década de investigación focalizado en entrenar a personas y organizaciones en esta perspectiva evolutiva, desarrollando sesiones de trabajo donde aplico técnicas de biodigitalización adaptativa evolutivas para que el modelo mental de la persona se autoperciba desde la perspectiva de la máquina. Con el ADSM (Análisis Digital de Sistema Mentales) me sigo acercando al objetivo de transformar el sistema mental obsoleto (biocultural de la Vida 1.0) en uno preparado para el futuro (biodigital de la Vida 3.0). Trabajo con dos únicas herramientas analíticas:

a. Una matriz de datos (MDS-YO o Matriz Digital Simplificada del YO) que permite convertir cualquier experiencia humana en un episodio computado ordenadamente mediante el registro de 6 tipos de datos.

+

b. Una fórmula algorítmica que permite a la persona descifrar su AP (Algoritmo Personal) que utilizará para reconfigurar su autopercepción y comportamiento general y utilizarlo en la vida cotidiana.

Mi objetivo actual es que mi tarea de analista la realice una IA para que millones de personas puedan biodigitalizarse en pocos minutos y descifrar



y utilizar su AP en su vida diaria. Cada vez se hace más insistente la pregunta sobre qué debemos aprender para vivir el futuro. Y la respuesta la encontramos en estas tres competencias:

- a. Computar (los datos estratégicos de la propia interioridad en una matriz como la MDS-YO),
- b. Analizar (la dinámica de estos datos),
- c. Descifrar y escribir y aplicar el AP (Algoritmo Personal) en la vida cotidiana como una poderosa herramienta de optimización de su autopercepción y su comportamiento general.

Los próximos sistemas mentales humanos tendrán que relacionarse eficazmente con las competencias que van adquiriendo las máquinas. Nos toca reaprender competencias que creíamos dominar cuando la aventura de la IA comenzó. Pero debemos hacerlo con perspectiva

evolutiva. Se trata de competencias antropotécnicas, aquellas cuyo entrenamiento modifica la configuración misma del sistema mental para transformarlo en el modelo biodigital de los próximos humanos.

El humano biodigital no tendrá temor alguno de comportarse con la eficacia de una máquina y con la sensibilidad de un humano. La integración de ambas surgirá de un sistema mental aumentado en su potencia y su plenitud para vivir la vida.

Nota: Las ideas aquí expuestas forman parte del libro publicado en junio del 2023 *HOMO VITAE: Manifiesto para un Humanismo biodigital. Una solución al pánico que ha despertado la IA.*

Bibliografía

- Delclaux Oraa, I., Seoane Rey, J.** (1982): *Psicología cognitiva y procesamiento de la información.* Madrid, Ediciones Pirámide.
- Ikonicoff, R.** (1999): *La conciencia y la máquina.* Barcelona, Galaxia Gutenberg/ Circuito de Lectores.
- Tegmark, M.** (2018): *Vida 3.0. ¿Qué significa ser humano en la era de la inteligencia artificial?* Madrid, Taurus.
- Tudurí, G.** (2023): *Homo VITAE. Manifiesto para un humanismo biodigital. Una solución ante el pánico que ha despertado la IA.* Amazon.



ANDRÉS PEDREÑO MUÑOZ
Catedrático de universidad
en Economía Aplicada.
Autor de más de 70 libros.

¿Un camino hacia la democratización o la inercia reguladora?

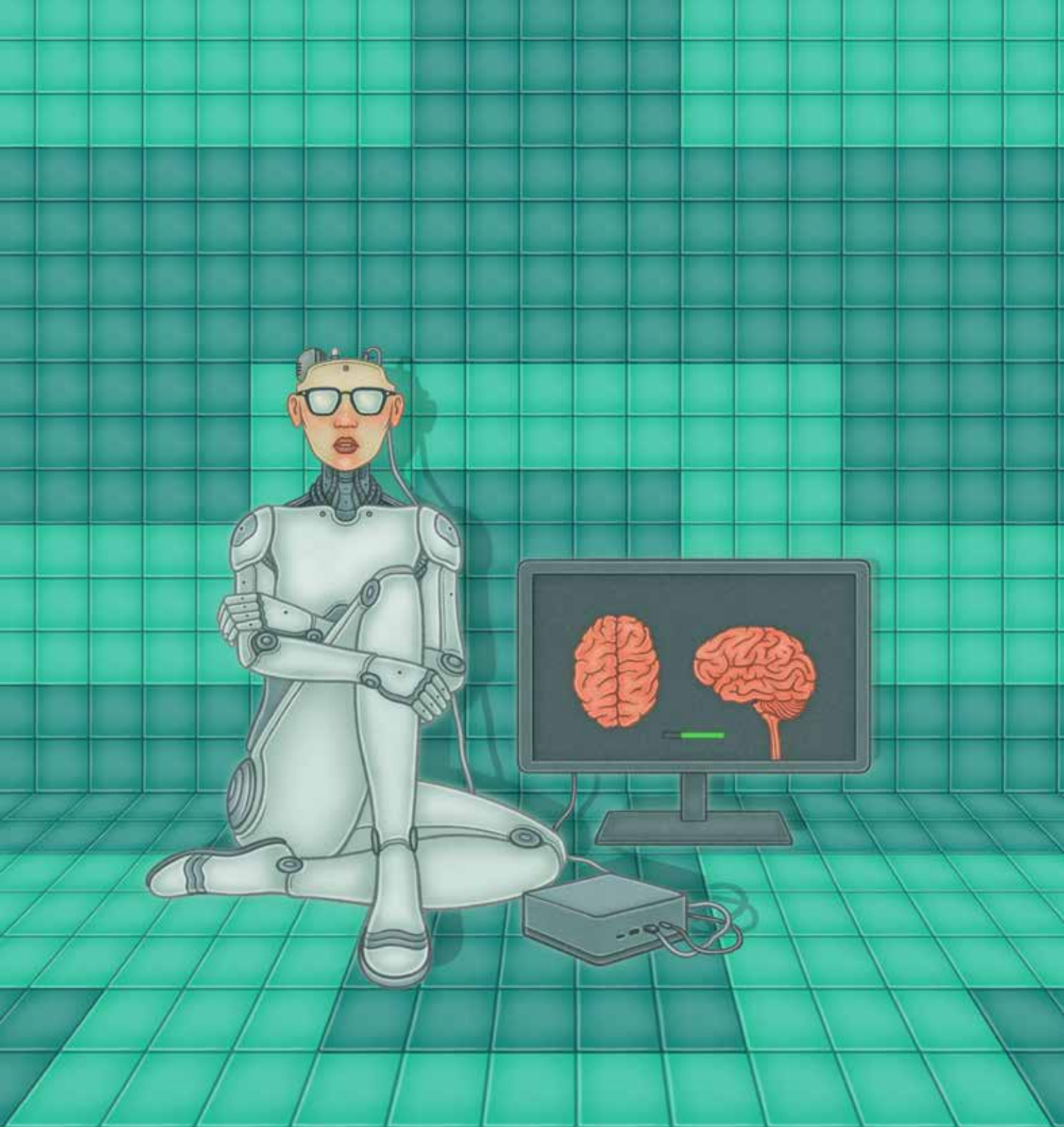
IA GENERATIVA. ¿ESPAÑA EN LA NUBE O EN LAS NUBES?

El artículo explora la compleja relación de España y, por extensión, de Europa, con la inteligencia artificial (IA) generativa, enfatizando cómo la nube computacional ha democratizado el acceso a la tecnología a una escala global. Critica la falta de pragmatismo y la resistencia al cambio en España y Europa, caracterizando esta actitud como "estar en las nubes". La regulación excesiva, la postura crítica y escéptica, y el pesimismo generalizado están frenando la innovación y la adopción de la IA en la región.



Palabras clave:

inteligencia artificial generativa, regulación, nube computacional,
capital humano, transformación empresarial, desempleo.



A path towards democratization or regulatory inertia?
GENERATIVE AI: SPAIN IN THE CLOUD OR IN THE CLOUDS?

The article explores the complex relationship of Spain and, by extension, Europe, with generative artificial intelligence (AI), emphasizing how cloud computing has democratized access to technology on a global scale. It criticizes the lack of pragmatism and resistance to change in Spain and Europe, characterizing this attitude as “being in the clouds.” Excessive regulation, a critical and skeptical stance, and widespread pessimism are holding back innovation and adoption of AI in the region.

Keywords: *generative artificial intelligence, regulation, cloud computing, human capital, business transformation, unemployment.*

El título de nuestro reciente libro, *España en la nube. ¿Una Startup Nation o el país del desempleo juvenil?* invita a reflexionar sobre nuestra posición como país ante la estrategia digital, especialmente en esta nueva era de la inteligencia artificial (IA) generativa.

Estar en la nube computacional es una señal positiva. Esta nube ha democratizado el acceso a recursos tecnológicos infinitos, disponibles casi en cualquier parte del planeta. Sin embargo, “estar en las nubes” es otra historia. Este término suele usarse para describir a alguien con una visión idealizada o poco realista, ignorando las limitaciones prácticas. Por ejemplo, si alguien tiene planes grandiosos, pero poco pragmáticos, podríamos decir que esa persona está “en las nubes”.

Cuando hablamos de inteligencia artificial, debo confesarles con sinceridad que, en gran medida, “España —y, por extensión, Europa— están en las nubes”. A pesar de contar con una minoría de entusiastas pioneros, talentos, *start-ups* competitivas internacionalmente e investigadores, una gran mayoría se resiste al cambio y se inclina más hacia alertar sobre los peligros, riesgos e inexactitudes de la IA.

A continuación, intentaré explicar esto de forma resumida en tres puntos:

Existe un gran discurso ético y humanista que se antepone a todo. En un país líder y con un alto nivel de

desarrollo de la IA esto es muy deseable y justificable. Las reuniones que ha mantenido la Casa Blanca con las principales empresas de inteligencia artificial están más que justificadas. En Europa, este discurso (con el que lógicamente todos estamos de acuerdo) se antepone a una estrategia efectiva de desarrollo. Con regulaciones preventivas e ineficientes que, en un afán de proteccionismo, perjudican y atrofian a nuestras empresas de IA.

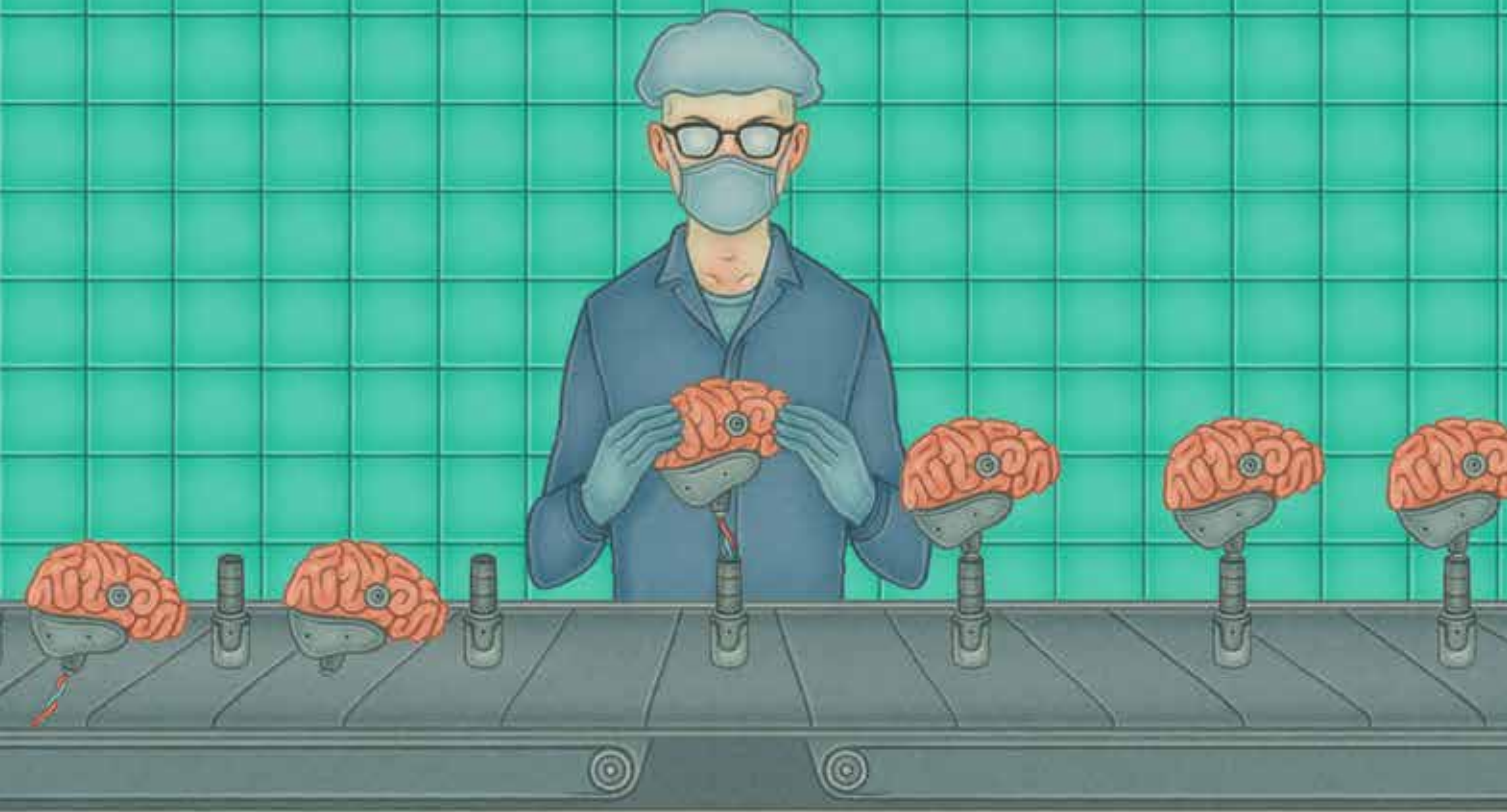
1. La IA, al igual que la economía, es global. Hace unos meses, un grupo de “expertos” firmó una carta solicitando una moratoria en el desarrollo de la IA. Algunos de ellos de manera bastante interesada — como sería el caso de Elon Musk— olvidaron que modelos *open source* (como Meta-LLaMA) impulsan su disponibilidad global. No parece prudente que países como China, Rusia, India o Corea del Norte no asuman ninguna moratoria, mientras que EE. UU. y Europa sí lo hagan. Afortunadamente, las voces más autorizadas como Andrew Ng y Yann LeCun, desde bases de conocimiento mucho más sólidas, alertaron sobre esta intención. El mayor riesgo que podría ocurrir es que Europa y Estados Unidos no lideren la IA. O que, en lugar de democratizar su uso, se convierta en patrimonio de unas pocas empresas que utilizan el

enorme potencial de estos modelos en beneficio propio. Esto era lo que sucedía antes del 30 de noviembre de 2022, fecha en la que OpenAI puso ChatGPT a disposición del público en general.

2. La estrategia correcta requiere una amplia educación. Necesitamos menos regulación y más educación e inversión. Muchos periodistas, colegas universitarios e incluso expertos se dejan seducir por la ciencia ficción, el pesimismo o las limitaciones e insuficiencias de los modelos. Cuando presento públicamente modelos, herramientas y productos basados en IA con altas capacidades, encuentro una actitud crítica y escéptica peligrosa. Lejos de apreciar su gran potencial y cómo revoluciona numerosas facetas de nuestra productividad y creatividad, se percibe un discurso victimista y negativo, centrado en la destrucción del empleo y peligros de todo tipo. Esta actitud favorece que todo lo que se publica resaltando limitaciones, inexactitudes, errores, etc., sea acogido como un bálsamo tranquilizador, influyendo negativamente en la mera receptividad (y ya no digamos en la adopción de la IA en la práctica). Como resultado, las regulaciones ineficientes y prohibitivas se perciben como un salvavidas ante una IA demonizada por sus riesgos y sesgos.

Necesitamos una IA controlada, optimizada, democratizada, gestionada y que empodere a las personas

Estamos alimentando una cultura negativa que puede ser desastrosa para un país como España. Como sostiene acertadamente Nuria Oliver, solo con la IA podremos enfrentar y resolver problemas críticos como el cambio climático o enfermedades complejas como el cáncer. En nuestro libro defendemos que solo con IA, España podrá mantener su competitividad empresarial en sectores como el turismo o la agricultura; solo con IA se podrá revolucionar nuestra sanidad o educación a un nivel notablemente avanzado y superior. ■■■



**Más allá de la transformación en
empresas e instituciones,
la IA tiene el potencial
de cambiar la forma en que
las personas trabajan y aprenden**

En definitiva, tenemos que bajar de las nubes e ilusionar a nuestros jóvenes para que abracen su futuro. Esto significa más STEM, más emprendimiento, más *start-ups* tecnológicas, más confianza y cultura de consumo empresarial... Según McKinsey solo la IA generativa originará un valor equivalente al PIB del Reino Unido (de 2.6 billones a 4.4 billones de dólares), tomando en consideración poco más de 60 casos de uso y sin tener en cuenta su marca en el *software*. Su impacto en la modernización de nuestras empresas y sectores clave será fundamental e imprescindible.

Frente al pesimismo y negativismo ilustrado, necesitamos dar respuesta a algunos retos y desafíos fundamentales en torno a la IA. Pero la estrategia empresarial para el desarrollo de la IA es hoy por hoy débil. La UE y sus países miembros —especialmente España— deberían poner el foco en cómo alcanzar una IA controlada, optimizada, democratizada, gestionada éticamente y que empodere a las personas.

IA controlada y optimizada

El control y la optimización de la IA son esenciales para su éxito. No podemos dejar nuestras respuestas legales o institucionales en manos de la IA. Por tanto, uno de los objetivos primordiales es cómo controlarla y establecer criterios transparentes. Un asistente conversacional puede diseñarse para responder solo con información seleccionada, o para presentar respuestas específicas previamente aprobadas y supervisadas por personas autorizadas. Además, puede complementarse con enormes capacidades lingüísticas en términos de

comprensión, traducción, estilos de lenguaje, etcétera. En España ya hay productos que ofrecen esto, pero se enfrentan a una cultura empresarial donde las barreras y el escepticismo reinante impide dar entrada a procesos de innovación rápidos que se conviertan en ventajas competitivas.

Un acceso justo y equitativo a la IA es esencial para prevenir un futuro dominado por un puñado de corporaciones que monopolizan esta tecnología revolucionaria. Tal situación acarrearía serios riesgos, desde la desintegración del tejido productivo hasta la agudización de las desigualdades económicas y sociales. Modelos como ChatGPT, fácilmente integrables en cualquier sitio web, representan un notable avance en la interacción con los clientes y un gran paso hacia la democratización del uso y la familiarización con la IA. Con costos mínimos mensuales, las pequeñas empresas pueden capitalizar la IA para potenciar su competitividad. Este acceso democratizado a la IA puede representar un cambio radical para las pequeñas empresas, permitiéndoles competir con las grandes en términos de servicio al cliente y eficiencia. Lamentablemente, no estamos aprovechando al máximo los fondos *Next GenerationEU* para este propósito.

La falta de capital humano y talento especializado en IA puede ser un obstáculo en la gestión de esta. Muchas empresas aspiran a tener una oficina de IA, pero se encuentran con el desafío de no saber por dónde empezar. La administración integral de una oficina de IA bajo la tutela de expertos puede garantizar tanto eficiencia como rapidez en el desarrollo y aplicación de estas tecnologías. Es fundamental abordar la profesionalización de la IA en las empresas. ►►►

La falta de capital humano y talento especializado en IA puede ser un obstáculo. Es fundamental abordar la profesionalización de la IA en las empresas

La gestión de la IA es un asunto serio; los enfoques sin una base sólida pueden alimentar más escepticismo y generar costes innecesarios.

Empoderar a las personas

Más allá de la transformación en empresas e instituciones, la IA tiene el potencial de cambiar la forma en la que las personas trabajan y aprenden. Para los estudiantes, puede proporcionar una educación más personalizada y basada en datos, facilitando la comprensión de textos y conceptos. Para los trabajadores, puede incrementar la productividad al automatizar tareas repetitivas y permitir concentrarse en tareas de mayor valor. Para los profesionales, puede potenciar la creatividad y la capacidad de resolución de problemas al proporcionar nuevas herramientas y enfoques. Este empoderamiento a través de la IA puede ser un poderoso catalizador del crecimiento personal y profesional.

En resumen, una IA controlada, optimizada, democratizada, gestionada éticamente y que empodere a las personas tiene el potencial de trans-

formar nuestra economía y sociedad de forma profunda y duradera. Las herramientas de las que hablamos son ya una realidad¹. Sin embargo, corremos el riesgo de que, como tantas veces ocurre con la tecnología, nos centremos en detalles y perdamos de vista el panorama general, o que, en nuestro afán de subir más alto, nos perdamos entre las nubes.

Para concluir, recordemos que la inteligencia artificial, en su esencia, es una herramienta forjada por mentes humanas. Su verdadero potencial reside en nuestra capacidad para utilizarla como un medio para elevar la humanidad, fomentar la equidad, y encender la chispa de la creatividad y la innovación. Tenemos el derecho a imaginar y luchar por un futuro donde la IA no sea simplemente utilizada, sino que sea un catalizador para alcanzar nuestro máximo potencial y trascender nuestras propias limitaciones. Eso, en esencia, es lo que significa la verdadera democratización de la IA.

Bibliografía

1MillionBot: líder en inteligencia artificial.

Disponible en: <https://1millionbot.com/>

Gumbrecht, H. (2020): *El espíritu del mundo en Silicon Valley: Vivir y pensar el futuro*. Barcelona, Deusto.

Lee, K.-F. (2018): *Superpotencias de la inteligencia artificial: China, Silicon Valley y el nuevo orden mundial*. Barcelona, Deusto.

Pedreño A., Moreno L. (2023): *España en la nube: ¿Una Startup Nation o el país del desempleo juvenil? Afrontando los retos en la era de la inteligencia artificial*, KDP.

Pedreño A., Moreno L. (2020): *Europa frente a EE. UU. y China. Prevenir el declive en la era de la inteligencia artificial*, KDP.

¹ Agradezco a los equipos de tecnología de 1MillionBot y Lucentia, así como al Grupo de investigación Eco-IA de la Universidad de Alicante (UA) los debates en torno a los temas de control, optimización, democratización, gestión y empoderamiento de las personas y cómo se resuelven estos retos en la práctica.

Análisis

La humanidad habrá de convivir con unas inteligencias artificiales que trastocan la forma en la que nos educamos, nos comunicamos e, incluso, nos conformamos como humanos capaces de desarrollar ingenios que ayudan a fortalecer la toma de decisiones éticas e informadas a ciudadanos.



Palabras clave:
educación,
inteligencia artificial
(IA), creatividad,
docentes,
aprendizaje,
tecnología.



ES ESENCIAL CAPACITAR
A DOCENTES Y ESTUDIANTES
PARA APROVECHAR LA IA



RODRIGO FÁBRAGA
Especialista en desarrollo
de la creatividad y la IA como
herramientas cognitivas en educación.

Inteligencia artificial: una herramienta para crear

La inteligencia artificial (IA) en educación debe enfocarse en potenciar la creatividad y el aprendizaje. A pesar de sus raíces conectadas con la educación, su implementación a menudo desvía el foco. Es esencial capacitar a docentes y estudiantes para aprovechar la IA, promoviendo métodos pedagógicos innovadores y enfatizando la construcción del conocimiento. La IA ofrece oportunidades para repensar cómo aprendemos y potenciar la creatividad estudiantil.

It is essential to train teachers and students to take advantage of AI

**ARTIFICIAL INTELLIGENCE:
A TOOL TO CREATE**

Artificial intelligence (AI) in education should focus on enhancing creativity and learning. Despite its roots connected to education, its implementation often deviates the focus. It is essential to train teachers and students to take advantage of AI, promoting innovative pedagogical methods and emphasizing the construction of knowledge. AI offers opportunities to rethink how we learn and enhance student creativity.

Keywords: education, artificial intelligence (AI), creativity, teachers, learning, technology.



En la educación, la inteligencia artificial (IA) es una herramienta cognitiva. El foco debería estar en aprender a usarla para revitalizar los aprendizajes y contribuir al diseño de espacios educativos donde los estudiantes sean capaces de aumentar sus posibilidades creativas. Pero eso no es lo que está sucediendo. Al contrario: a medida que aumenta el debate acerca de las IA en educación, más se pierde ese foco original en la relación entre las IA y el proceso educativo.

En este artículo revisaremos las raíces de la IA y su relación original con la educación, expondremos las posibilidades de eficiencia que ofrece a docentes, así como las alternativas para revitalizar los espacios educativos y para comprenderla como un instrumento que no reemplaza la experiencia de aprender, sino que contribuye a que los estudiantes expandan sus capacidades creativas.

Desde un principio, el desarrollo de la inteligencia artificial implicó un debate técnico, ético y filosófico. La motivación inicial, a mediados de la década de los años 50 del siglo pasado, fue encontrar cómo hacer que las máquinas usen nuestro lenguaje, formen abstracciones, conceptos, aprendan y resuelvan problemas que, por ahora, están reservados para los humanos. Es evidente que se han respondido muchas de aquellos interrogantes. Sin embargo, había también un interés menos difundido: los investigadores observaban que la diferencia entre el pensamiento creativo y el pensamiento competente —pero poco imaginativo— se basaba en cierta cuota de aleatoriedad, que debía ser guiada por personas. Expresamente, una de las líneas de trabajo iniciales

de la IA fue la creatividad y la aleatoriedad. Lo que pueda o no ocurrir en las escuelas del mundo hoy tiene que ver con esta conjetura: la IA en educación es una herramienta en manos de docentes para cultivar el pensamiento creativo.

Tempranamente, una serie de científicos observaron el potencial que tendría en la educación incorporar la IA como una herramienta para pensar. Mientras un grupo se dedicó a desarrollar sistemas para interactuar con estudiantes, donde estos podían aprender a su propio ritmo, otros pensaron que deberían ser los estudiantes los que fueran capaces de programar los computadores y así dotarlos de cierta inteligencia, y para ello crearon lenguajes de programación.

Los sistemas resultaron atractivos, incluso en un mundo donde no había acceso general a los computadores, pero los científicos sabían que la proliferación de los computadores personales era cosa de décadas solamente. Por ejemplo, J. C. R. Licklider quien a inicio de los 60 creó un tutor artificial para aprender alemán, donde los niños y niñas tenían un rol activo en el diálogo con el computador. Los hoy populares chats con IA generativa tuvieron sus precedentes en los trabajos, entre otros, de Jaime Carbonell¹, del grupo de Carnegie Mellon University, quien presentó uno de los primeros sistemas tutoriales para la escuela. Su propuesta superaba los sistemas de instrucción asistidos por computador, con preguntas unidireccionales y prehechas. Introdujo la posibilidad de dialogar acerca de un tema específico; se trataba de *Scholar*, un precursor de los sistemas que comprenden el lenguaje natural:

1 Más información sobre Jaime Carbonell y Scholar. Disponible en: <https://www.cs.cmu.edu/~jgc/>

2. Más información de los trabajos de Seymour Papert en su página web: <http://www.papert.org>

Scholar: "Hola, mi nombre es Scholar. Te voy a ayudar a revisar la geografía de América del Sur. Este es el ABC Escuela Regional y hoy es 29 de abril de 1970.

Scholar: Qué es Chile

Estudiante: Un país

Scholar: Está bien

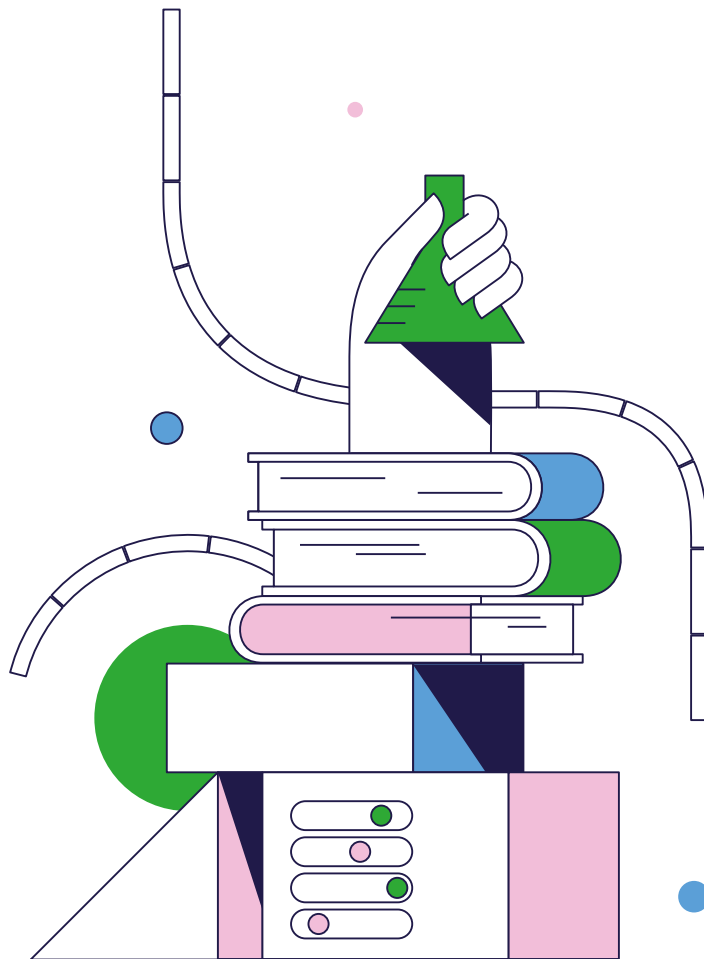
Scholar: Selecciona una alternativa desde esta lista

- 214000
- 800300
- 49883
- 63690

Para esta pregunta: ¿Aproximadamente, cuál es el área de Chile?

Estudiante: 800300

Scholar: Tu respuesta es incorrecta. La respuesta correcta es 214000".



Otro grupo de científicos tomó una ruta distinta, donde la naturaleza de la solución de IA no era tecnocéntrica, sino más bien epistemológica. En el año 1963 se hicieron los primeros ensayos para enseñar a programar en las escuelas con TELCOMP, una derivación del lenguaje FORTRAN. En 1966 se desarrolló un lenguaje especialmente diseñado para estudiantes, LOGO, basado en el lenguaje preferido de la IA en aquella época, el LISP. A mediados de los años 60, Seymour Papert² había llegado al MIT luego de trabajar con Jean Piaget. Mientras Papert se encontraba investigando acerca de la IA, el científico suizo hacía lo propio en el desarrollo intelectual de los niños en el Centro de Epistemología Genética en Ginebra; él pensaba que la inteligencia provenía de una adaptación, que a su vez era asimilada y que luego se producía un proceso complementario de acomodación. Así, la inteligencia se producía en base al interés de niños y niñas, a las

ganas de aprender. Papert hizo una contribución al encuentro entre la comprensión del estudio de la inteligencia humana y la IA.

A pesar del evidente avance tecnológico, se observan menos cambios en los métodos de enseñanza y todavía una confusión acerca de las posibilidades que ofrece la IA en nuestros días. Entonces ¿cuáles son las oportunidades que ofrece la IA en el ámbito educativo?

Eficiencia docente

Una primera oportunidad. ¿Pueden los y las docentes ser más eficientes con el uso de las IA? No puede escucharse otra palabra que un rotundo sí. Los y las docentes deberían hacer una adopción rápida de las herramientas IA que les permitan aumentar su eficiencia. Análogamente cuando pasamos del cálculo manual de calificaciones, luego a usar la calculadora, ►►►

La IA
contribuye
a que los
estudiantes
expandan
sus
capacidades
creativas

Los y las docentes deberían adoptar rápidamente las herramientas IA que les permitan aumentar su eficiencia

enseguida las planillas de cálculo y, finalmente, los sistemas de administración automáticos. Si un sistema permite cargar una planilla de cálculo y le solicito que genere un informe detallado sobre el desempeño de los y las estudiantes y me entrega esa información en diez segundos; si puedo aprender a diseñar rúbricas de evaluación específicas; si puedo solicitar que desarrolle un programa en Python para resolver un problema; si puedo ahorrar tiempo en actividades rutinarias, entonces existen razones para que, con prontitud, los y las docentes aprendan a usar estas tecnologías. Ya hay guías completas disponibles y muchas de estas herramientas son gratuitas, incluso algunas de ellas no requieren conexión a Internet.

Es alarmante que algunas universidades están desarrollando primero guías para evitar las IA, en lugar de desarrollar departamentos para enseñarles a los y las estudiantes a usarlas y producirlas. En los próximos años un docente que se gradúe sin saber trabajar con IA partirá su carrera profesional con una importante desventaja. También es preocupante que los Ministerios de Educación no tengan departamentos de IA rediseñando la educación; como hemos mostrado, el desarrollo de la IA y su conexión con la educación no es algo nuevo, hay pocas excusas.

Enseñar a pensar

Una segunda oportunidad consiste en comenzar una conversación lenta acerca de cómo aprendemos y qué ventajas se desprenden de la incorporación de la IA. En 1970 Seymour Papert, desde el Laboratorio de Inteligencia Artificial del MIT, escribió el Memo 247 de IA y el segundo de LOGO, *Teaching Children Thinking*;

10 años más tarde publicó *MindStorm* donde presentó su enfoque constructivista del aprendizaje. Papert, con ojos piageteanos, observó que la intersección entre la tecnología y el aprendizaje era una oportunidad para que los propios estudiantes pudieran entender su propio pensamiento. Es decir, en verdad, el desafío de la escuela era enseñar a pensar a sus estudiantes. Usando LOGO, los y las estudiantes podían involucrarse en sus propias creaciones y aprender matemáticas dentro de la matemática, en lugar de simplemente aprender sobre la matemática. Papert argumentaba que, tal y como para aprender francés probablemente lo mejor era estar en Francia, análogamente para aprender matemática había que ingresar a *Matemalandia*. No se podía separar la matemática de la enseñanza de la matemática.

Pero el foco actual no parece establecerse desde la construcción del conocimiento, sino más bien de un reforzamiento del enfoque tradicional. Mitch Resnick, líder del Lifelong Kindergarten MIT, llama la atención acerca de las preocupaciones y oportunidades que trae esta nueva ola de IA generativa. En primer lugar, no ver esta oportunidad como una buena alternativa para superar el instruccionalismo. Por ejemplo, distraerse con un sistema diseñado para ayudar a resolver ecuaciones sin antes preguntarse ¿pero es esta una buena forma de aprender matemáticas? Una segunda preocupación tiene que ver con que los actuales sistemas de IA han puesto su foco en resolver problemas de solución única. Es decir, son tutoriales que llevan a un alumno o alumna a resolver un problema de una manera específica, limitando de esa forma la capacidad de descubrir la solución. Piaget decía que, si respondemos las preguntas de los niños y niñas de esta forma, le quitamos la oportunidad de por vida

³ K-12 es una definición que popularmente se utiliza en Norteamérica para definir las edades educativas comprendidas entre el jardín de infancia y el 12º grado. K-12 podría englobar desde los 4 hasta los 16 años.

⁴ Disponible en: <https://web.learningm.org/>

de que la pudieran descubrir por sí mismos; muchos de los sistemas actuales estarían en desacuerdo con Piaget. Una tercera preocupación es tratar de asimilar las IA a las capacidades de los docentes y no considerar el valor de la conexión humana que se produce en un aula y la capacidad que tiene la escuela de crear comunidad. Pero también ve oportunidades para ver la IA como una manera de apoyar el aprendizaje creativo.

Una tercera oportunidad es que los y las estudiantes sean los inteligentes que le enseñan a una máquina obediente a tener algún tipo de inteligencia. Es importante conocer paso a paso qué es, cómo funciona, cuáles son las oportunidades y riesgos de la IA desde el K al 12vo grado³. Docentes de todo el mundo ya están trabajando con currículos diseñados colaborativamente entre la academia, la industria y la escuela. Por ejemplo, el grupo RAISE de MIT creó un completo currículum para comprender qué es la IA (www.diadelainteligenciaartificial.org); Profuturo desarrolló la Escuela de Pensamiento Computacional e Inteligencia Artificial para formación docente y un nutrido grupo de empresas creó Teachai.org, entre otras posibilidades.

Luego de entender en qué consiste, los y las estudiantes pueden crear sus propios modelos de *Machine learning*. Por ejemplo, utilizando Learning-ML.org⁴, los estudiantes asumen el papel del maestro inteligente, impartiendo conocimiento e instrucciones a máquinas obedientes. Este poderoso concepto invierte la dinámica tradicional y empodera a los estudiantes para convertirse en los creadores y directores de la inteligencia artificial. Con este tipo de herramientas los estudiantes tienen la oportunidad de diseñar y entrenar sus propios modelos, enseñando a las máquinas cómo realizar tareas específicas y hacer predicciones. Este enfoque práctico no solo

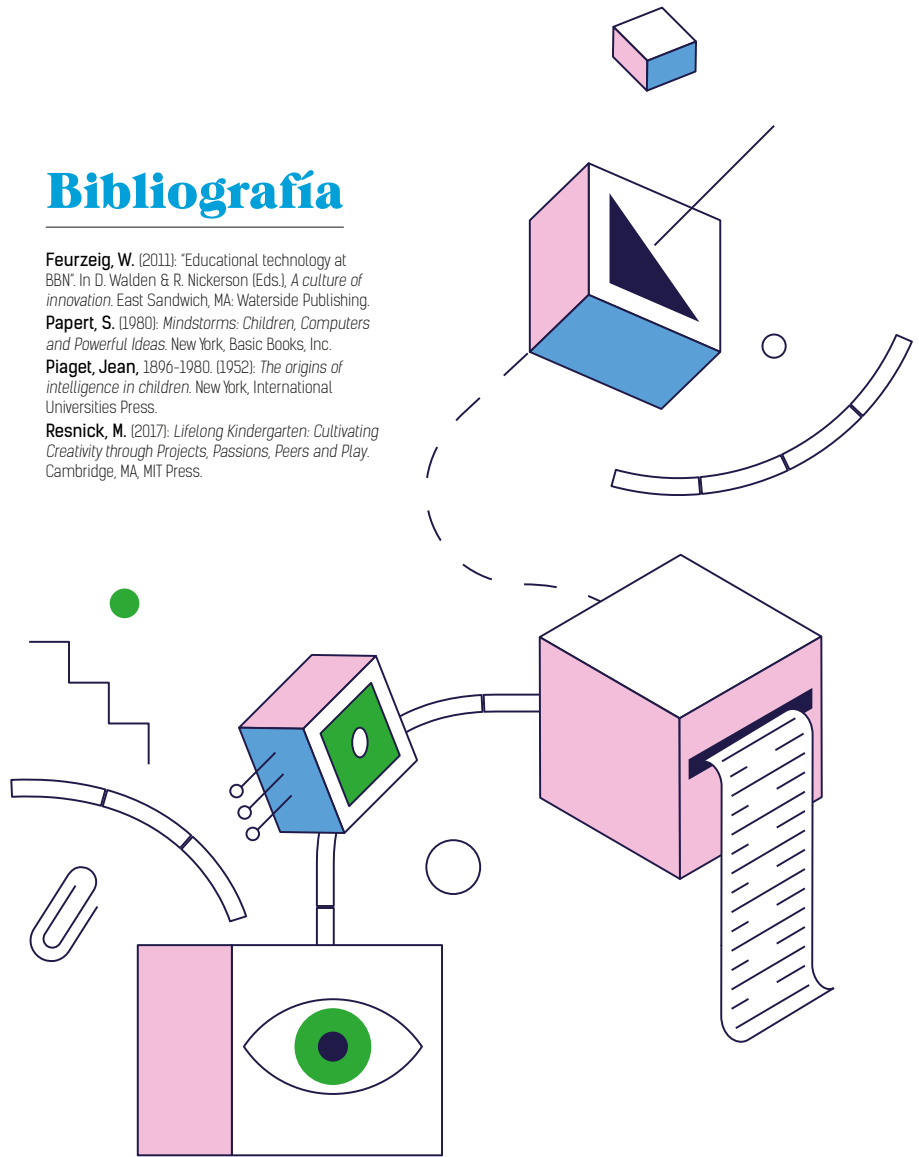
Bibliografía

Feurzeig, W. (2011). "Educational technology at BBN". In D. Walden & R. Nickerson (Eds.), *A culture of innovation*. East Sandwich, MA: Waterside Publishing.

Papert, S. (1980). *Mindstorms: Children, Computers and Powerful Ideas*. New York, Basic Books, Inc.

Piaget, Jean, 1896-1980. (1952). *The origins of intelligence in children*. New York, International Universities Press.

Resnick, M. (2017). *Lifelong Kindergarten: Cultivating Creativity through Projects, Passions, Peers and Play*. Cambridge, MA, MIT Press.



mejora su comprensión de los principios de la IA, sino que también fomenta un sentido de propiedad y creatividad en su proceso de aprendizaje.

En los últimos 60 años la investigación y aplicación de la IA ha tenido altos y bajos, ahora nos encontramos en un momento de especial atención, con la característica adicional ya que, por primera vez, la nueva tecnología se expande por el orbe simultáneamente. Rápidamente, el mundo educativo debería aprender de esta

tecnología; lentamente, hay que repensar cómo aprendemos y así ver la oportunidad cognitiva que ofrece la IA; y, masivamente, deberíamos dotar a los niños y niñas de una nueva materialidad para crear. Esto no se puede hacer sin la escuela y sin los docentes. ¿Partimos?

**LA CONEXIÓN ENTRE PROGRESO
ECONÓMICO Y SOCIAL**

Historia de humanos y algoritmos

El desarrollo de modelos de inteligencia artificial generativa cada vez más sofisticados ha suscitado un debate acerca de las implicaciones que esta tecnología tendrá para el ser humano en los próximos años desde el punto de vista económico, productivo y social. ¿Cómo podemos construir un futuro integrador en términos de progreso?



ANA SAINZ
Directora general
de Fundación SERES

A principios del siglo XIX en Reino Unido, los luditas fueron los primeros en preocuparse por su futuro laboral ante la llegada de las máquinas y la automatización de sus trabajos. Este movimiento comenzó el 11 de marzo de 1811 en Arnold, un pueblo cerca de Nottingham, cuando los trabajadores británicos destruyeron los telares de las fábricas textiles. Los luditas se inspiraron en el personaje imaginario de Ned Ludd, un joven aprendiz que destruyó en 1779 el telar de su propio maestro.

A finales del siglo XVIII la tensión entre los trabajadores y empresarios fue creciendo con la prohibición de los sindicatos. Las restricciones impuestas por el Gobierno británico sobre el comercio internacional obligaron a los empresarios a reducir sus costes y producir materias primas de baja calidad, por lo que recurrieron a las máquinas. Las jornadas laborales se alargaron y los salarios se hicieron cada vez más precarios. Los luditas demostraron su descontento asaltando fábricas y destruyendo telares. Probablemente del pasado hemos aprendido que las desventajas de una tecnología o avance, entendido como progreso general, no residen en la tecnología en sí, sino en el uso que ha-

ceamos las personas de ella —con inteligencia y encontrando respuestas—.

La llegada de la IA y la revolución digital está generando en los ciudadanos una sensación similar a la que se produjo con la Revolución Industrial aunque, según los expertos, con dos diferencias importantes: la velocidad de irrupción en nuestras vidas y la intensidad y globalidad del cambio que genera, que nos afectará y nos transformará a todos.

Es una tecnología disruptiva, que ha avivado el debate sobre las reglas que deben regir su funcionamiento; hablamos de la ética en la IA, que se engloba dentro de una temática más amplia sobre el control de la tecnología por los propios humanos.

Que la IA ha cambiado nuestra forma de vivir, trabajar e incluso visionar nuestro futuro no es algo nuevo. Algunos sectores como la salud o la educación han experimentado una auténtica revolución con el potencial de transformar profundamente la sociedad. La IA está mejorando las organizaciones, así como su gestión y procesos internos. Aparecen nuevos modelos de negocio, productos y servicios. Y también para los Estados: predictibilidad de las decisiones administrativas y de las sen- ➤

The connection between economic and social progress

HISTORY OF HUMANS AND ALGORITHMS

The development of increasingly sophisticated generative artificial intelligence models has sparked a debate about the implications that this technology will have for human beings in the coming years from an economic, productive and social point of view. How can we build an inclusive future in terms of progress?

Keywords: ethical AI, responsible technology, social commitment, reskilling, upskilling, responsible management.

Palabras clave:

IA ética, tecnología responsable, compromiso social, *reskilling*, *upskilling*, gestión responsable.





El progreso exige reflexión, conocer y abordar los desafíos sociales asociados a la tecnología

tencias judiciales; sistemas objetivos de selección de empleados públicos; incremento de la protección policial; identificación de delincuentes; predictibilidad de defraudaciones e incumplimientos con la Hacienda pública o la Seguridad Social, etcétera.

No se trata solo de acortar los tiempos de producción, podría llegar a revolucionar el trabajo en todos los sectores, automatizar tareas, fomentar la creatividad y mejorar la toma de decisiones, dotando a los seres humanos de herramientas para avanzar en el progreso. La innovación tecnológica puede perfeccionar las habilidades de precisión de un ser humano y realizar tareas metódicas con una consistencia casi perfecta, pero su gran potencial está en la combinación con las capacidades de las personas y el uso en favor de todos.

Que la inteligencia artificial es capaz de afectar economías y sociedades enteras también es un hecho. Eso refuerza la necesidad de garantizar su uso responsable y ético y distintas tendencias no cesan de alertar sobre los retos vinculados con la equidad, o incluso de una falta de control sobre la tecnología por parte de los humanos. Algunas voces la han llegado a paragonar con una revolución nuclear por su alto impacto en todos los ámbitos y, en parte, por la duda que

aún siembra sobre si nos ofrece un futuro prometedor, o todo lo contrario. Aún existe poco consenso sobre si hay que tildar a la tecnología de heroína o villana. De hecho, hay figuras destacadas como el historiador y filósofo Yuval Noah Harari que han impulsado una carta en la que se propone una “pausa de la inteligencia artificial”. ¿Qué hay tras este debate? ¿Estamos ante una amenaza real? ¿Cómo podemos construir un futuro integrador en términos de progreso?

El World Economic Forum publicaba en junio de este año un informe en el que desmenuzaba las implicaciones éticas del uso de la IA y algunos puntos de reflexión para salvaguardar la gobernanza y la responsabilidad en el uso de esta tecnología en el que las palabras más mencionadas fueron *ética y responsable*.

También hay voces que ven los riesgos de una reducción elevada de puestos de trabajo, a la que deberíamos adelantarnos para redirigir antes hacia nuevas oportunidades a estas personas. El *Informe sobre el futuro del empleo 2023* del Foro Económico Mundial evalúa el impacto de las macrotendencias, así como del cambio tecnológico en los puestos de trabajo y las cualificaciones en los próximos cinco años. El informe concluye que casi una cuarta parte de todos los puestos de trabajo (23 %) cambiarán en los próximos cinco años. En 45 economías, que abarcan 673 millones de trabajadores, se prevé la creación de 69 millones de nuevos puestos de trabajo y la supresión de 83 millones, lo que supone una disminución neta de 14 millones de empleos, es decir, el 2 % del empleo actual.

Los mayores aumentos absolutos de empleo procederán de la educación (tres millones de empleos) y la agricultura (cuatro millones de empleos), impulsados en parte por la demografía, y en parte por las aplicaciones de las

nuevas tecnologías en estos campos. También se espera que la nueva geografía económica (creada por el cambio de las cadenas de suministro) y el mayor enfoque en resiliencia (en vez de eficiencia) generen un crecimiento neto del empleo con beneficios, especialmente para las economías de Asia y Oriente Medio.

Los sesgos son otro de los aspectos que se encuentran bajo mayor escrutinio si hablamos de IA. No cabe duda de que, si no alineamos estas tecnologías con los objetivos conectados con la humanidad, no vamos a lograr tener sociedades más sostenibles. En el trabajo concreto desarrollado por SERES y vinculado a una IA ética hemos identificado algunas características necesarias para contar con una IA que ponga el foco en las personas y sea así posible aprovechar toda su potencialidad:

1. Ser confiable: principio básico para tener una guía responsable que permita la colaboración entre las empresas y los clientes.
2. Adaptable: la tecnología debe evolucionar con el usuario, adaptándose de acuerdo con la interacción con la persona.
3. Relevante: es necesario que aporte valor al usuario, o permita mejorar el valor ofrecido.

4. Útil: debe aportar algo que antes no estaba y ofrecer valor al usuario con el mínimo esfuerzo.

También existe el debate entre la potencia del dato para generar impacto positivo en la sociedad y el riesgo de la mala utilización de los datos de los usuarios. La nueva regulación relativa a los derechos humanos tiene muy en cuenta este aspecto, y la necesidad de regular el acceso al dato garantizando los derechos de las personas.

Hoy nadie duda del potencial de crecimiento que aporta la tecnología, y tampoco, que su uso conlleva responsabilidades. Eso sí, este progreso exige reflexión, conocer y abordar los desafíos sociales asociados a la tecnología y entender desde el inicio los impactos que la implementación de la IA implica en las personas. Supone comprender mejor, *ex ante*, cuáles son los riesgos e impactos negativos a los que debemos anticiparnos, integrarlos en el diseño y así darles respuesta. En definitiva, las personas deben estar en el centro de todo avance tecnológico.

Bibliografía

Barrio Andrés, M. (2023): "Novedades en la tramitación del próximo reglamento europeo de inteligencia artificial". Real Instituto Elcano. Disponible en: <https://www.realinstitutoelcano.org/analisis/novedades-en-la-tramitacion-del-proximo-reglamento-europeo-de-inteligencia-artificial/>

Kumar, M. y Li, C. (2023): *Adopting AI Responsibility Guidelines for Procurement of AI Solutions by the Private Sector*. World Economic Forum. Disponible en: https://www3.weforum.org/docs/WEF_Adopting_AI_Responsibility_Guidelines_for_Procurement_of_AI_Solutions_by_the_Private_Sector_2023.pdf

Prees, G. (2023): "Diagnostic Robotics AI Advances Predictive, Personalized Medicine" en *Forbes*. Disponible en: <https://www.forbes.com/sites/gilpress/2023/07/17/diagnostic-robotics-ai-advances-predictive-personalized-medicine/?sh=110fa8392144>

Snabe, J. (2023): "Reimagining the enterprise with AI" en *Forbes*. Disponible en: https://brand-studio.fortune.com/c3-ai/reimagining-the-enterprise-with-ai/?prx_t=tn4IaehxgAoPEQA&ntv_ui=afceabe1-fda1-4f1f-a203-b2bd67d7fa3c&ntv_ht=WgjCZAA&gl=116wuqdd_gcl_auMTE20TM50TMyMC4xNj5MjM2NDUy



Palabras clave:

inteligencia artificial, información, crisis, expertos humanos, máquina y humano.



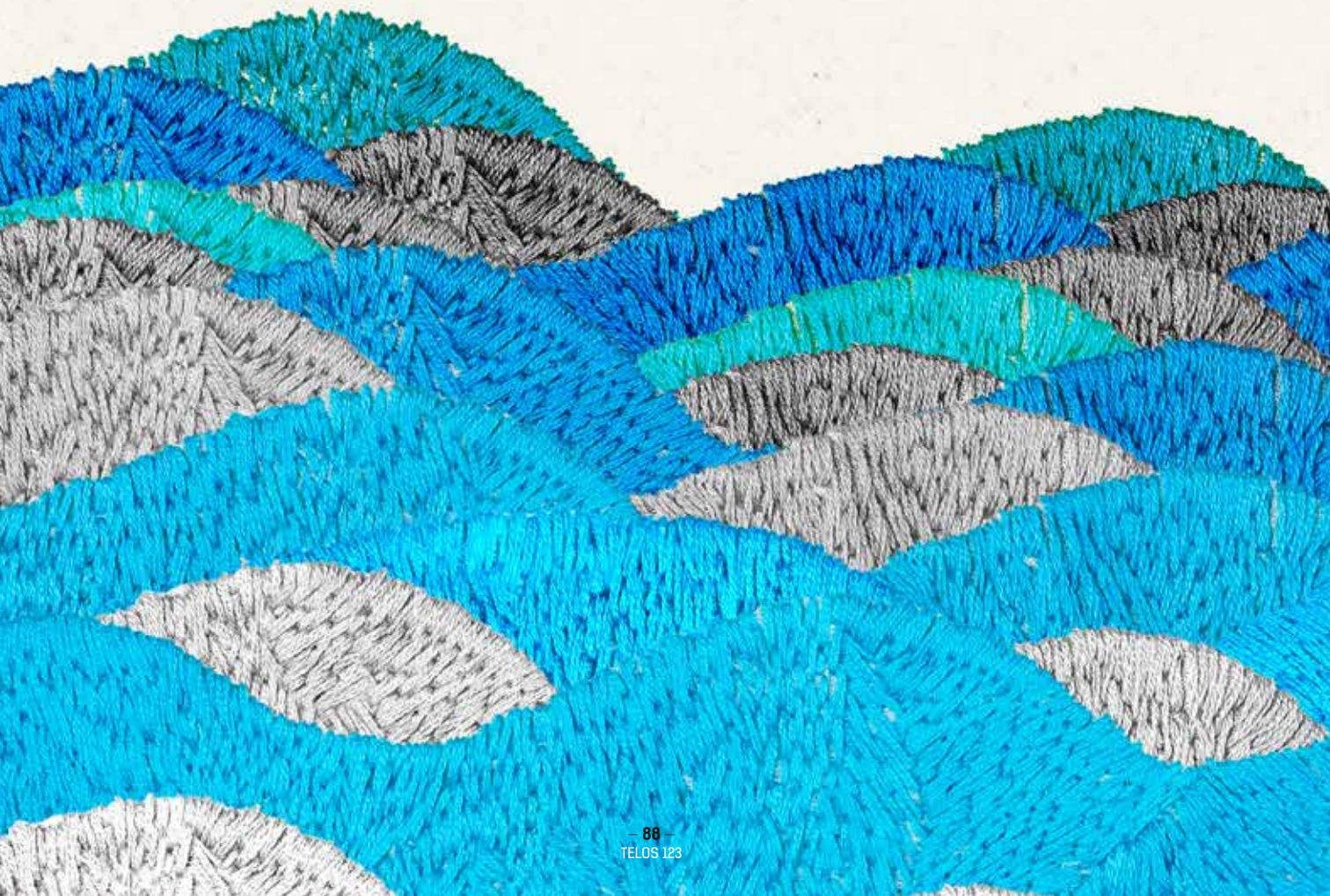
MANUEL CEBRIÁN

Investigador distinguido y catedrático de Excelencia Permanente en la UC3M

CÓMO LA INTELIGENCIA ARTIFICIAL PUEDE AYUDARNOS A NAVEGAR POR LAS CRISIS

La IA como salvavidas

La IA puede ser un salvavidas en tiempos de crisis, un socio que nos ayuda a mantener la cabeza fuera del agua. Puede ayudarnos a procesar la avalancha de información, a encontrar claves que de otro modo pasaríamos por alto, a coordinar una respuesta rápida y efectiva. La IA por sí sola no es suficiente; requiere la guía de expertos humanos.



How artificial intelligence can help us navigate crises
AI AS A LIFESAVER

AI can be a lifesaver in times of crisis, a partner that helps us keep our heads above water. It can help us process the flood of information, find clues we would otherwise miss, coordinate a quick and effective response. AI alone is not enough; it requires the guidance of human experts.

Keywords: artificial intelligence, information, crisis, human experts, machine and human.

Imagina que se desata una crisis. La información fluye desde todas partes, un torrente de datos abrumador en su volumen y complejidad. En medio del caos, la inteligencia artificial entra en acción. No es una IA cualquiera, sino un Modelo de Lenguaje Grande (MGL), una maravilla de la tecnología moderna que puede comprender y generar lenguaje humano. Comienza a recopilar datos, filtrando el ruido para encontrar la señal. Resume la información, destilando el caos en claridad. Interactúa con las personas, proporcionando actualizaciones, respondiendo preguntas, ofreciendo tranquilidad. Es un faro de calma en la tormenta, una mano firme en medio de la turbulencia.

Pero el MGL no trabaja solo. Detrás de escena, hay un equipo de expertos, trabajando con algoritmos complejos y herramientas de optimización. Son los arquitectos de la IA, los titiriteros que

mueven los hilos. Toman la información recopilada por el MGL, los planes que ha hecho, y los refinan. Utilizan su experiencia, su comprensión de la situación, para ajustar la salida de la IA. Son el toque humano en la máquina, el corazón en la IA.

En una crisis, esta combinación de IA y experiencia humana puede marcar un cambio radical. La IA puede rápidamente unificar todas las informaciones de las personas, dar a las personas una comprensión de la situación y rápidamente hacer planes. Los expertos humanos pueden luego mejorar esos planes con algoritmos más detallados y consejos. Es una asociación, una simbiosis entre máquina y humano, cada uno complementando los puntos fuertes del otro, compensando las debilidades del otro.

La IA puede ser un salvavidas en tiempos de crisis, un socio que ►►►

nos ayuda a mantener la cabeza fuera del agua. Puede ayudarnos a procesar la avalancha de información, a encontrar claves que de otro modo pasaríamos por alto, a coordinar una respuesta rápida y efectiva. La IA por sí sola no es suficiente; requiere la guía de expertos humanos. Pero juntos, la IA y el experto humano pueden ser una combinación poderosa para ayudar a navegar incluso la peor de las tormentas. La IA es el salvavidas, y nosotros somos sus navegantes. Juntos, podemos llegar a puerto seguro.

En el corazón de Getafê, una vibrante ciudad al sur de Madrid, el constante zumbido del aire acondicionado es mi fiel compañero. Es una nana mecánica que ahoga el sofocante verano español del exterior, un fuerte contraste con el calor de otro tipo que se gesta dentro de los confines de mi oficina en la Universidad Carlos III de Madrid. Me encuentro en el epicentro de un equipo global de científicos. Nuestros miembros están dispersos por todo el mundo, desde los sagrados pasillos del MIT y Georgia Tech en Estados Unidos hasta los verdes paisajes estivales de la Universidad de Aalto en Helsinki, un lugar más comúnmente asociado con la nieve, pero que actualmente se deleita con el resplandor del sol de medianoche.

Nuestra exploración no es de paisajes físicos, sino del vasto e in-

trincado laberinto de la inteligencia artificial. Es un viaje que no exige resistencia física sino agilidad computacional, un agudo intelecto y una insaciable curiosidad. No estamos explorando la IA solo por el avance tecnológico. Nuestra misión es mucho más urgente, mucho más humana: estamos trazando los territorios inexplorados de la IA, buscando formas de aprovechar su inmenso potencial para ayudarnos a navegar las futuras crisis que la humanidad está destinada a enfrentar.

Desafíos humanos

Nuestra misión no está arraigada en la mera curiosidad académica. Es un clamor que resuena desde los desafíos urgentes y apremiantes con los que nuestro mundo está lidiando. El cambio climático, con su marcha implacable alterando paisajes y vidas; pandemias, barriendo naciones, dejando rastros de devastación; amenazas cibernéticas, invisibles pero insidiosas, socavando nuestra sensación de seguridad; disturbios sociales, el espíritu humano clamando contra la injusticia... La letanía de crisis que exigen una acción inmediata y coordinada está creciendo, cada una de ellas es un severo

recordatorio de la fragilidad e interconexión de nuestro mundo.

Nuestras instituciones de respuesta a emergencias, esos bastiones firmes diseñados para una era diferente se encuentran en terreno inestable. Son como antiguas fortalezas enfrentándose a la artillería moderna, sus altos muros y profundos fosos no son rival para el ritmo y la complejidad de los desafíos actuales. El mundo está cambiando a una velocidad asombrosa y un poco aterradora, y la complejidad de los problemas que enfrentamos es una red enmarañada que desafía a cualquier experto u organización a desentrañarla.

Nos encontramos al borde de este precipicio, mirando hacia el abismo de lo desconocido, y está claro: necesitamos ayuda. Necesitamos un nuevo tipo de aliado, uno que pueda mantener el ritmo del incansable avance del cambio, que pueda navegar la laberíntica complejidad de nuestros problemas, que pueda ayudarnos a trazar un rumbo a través de la tormenta. Necesitamos la IA.

En medio de una crisis, el tiempo es esencial. Cada segundo cuenta, cada decisión importa. Pero nuestra capacidad humana para procesar información, para dar sentido al caos, es limitada. Solo somos humanos, después de todo. Nos cansamos, tropezamos, cometemos errores. Y en una crisis, esas limitaciones pueden



tener consecuencias graves. Pero, ¿y si tuviéramos ayuda? ¿Qué pasa si tuviéramos un compañero que no se cansara, que pudiera procesar enormes cantidades de información en un abrir y cerrar de ojos, que nos ayudara a navegar la tormenta? Ahí es donde entra la IA.

Imagínate una crisis en desarrollo. La información llega desde todos los rincones, un torrente de datos que abruma por su volumen y complejidad. En medio del caos, la IA entra en acción. No es cualquier IA, sino un Modelo de Lenguaje Grande (MLG), una maravilla de la tecnología moderna que puede entender y generar lenguaje humano. Comienza a recopilar datos, cribando el ruido para encontrar la señal. Resume la información, destilando el caos en claridad. Interactúa con la gente, proporcionando actualizaciones, respondiendo preguntas, ofreciendo seguridad. Es un faro de calma en la tormenta, una mano firme en medio de la conmoción.

Pero el MLG no trabaja solo. Detrás de escena hay un equipo de expertos, trabajando con algoritmos complejos y herramientas de optimización. Son los arquitectos de la IA, los titiriteros que mueven los hilos. Toman la información recopilada por el MLG, los planes que ha hecho, y los refinan. Utilizan su experiencia, su comprensión de la situación, para afinar los re-

Nuestras instituciones de respuesta a emergencias, esos bastiones firmes diseñados para una era diferente, se encuentran en terreno inestable

sultados de la IA. Son el toque humano en la máquina, el corazón en la IA.

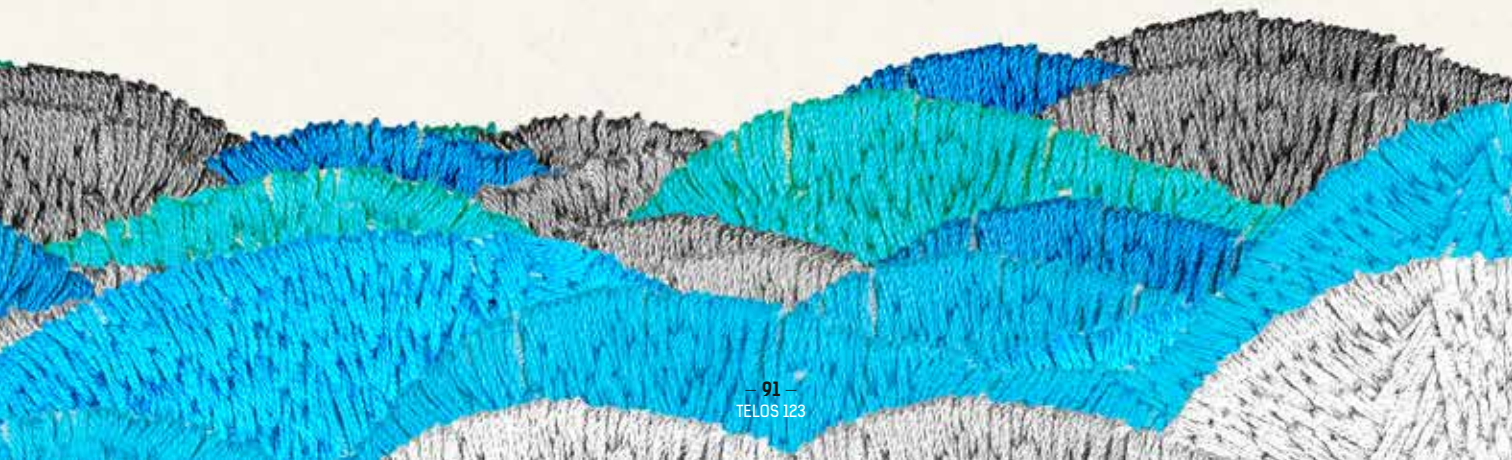
En una crisis, esta combinación de IA y experiencia humana puede cambiar el juego. La IA puede recopilar rápidamente toda la información de las personas, ofrecerles una comprensión de la situación y hacer planes rápidamente. Los expertos humanos pueden mejorar esos planes con algoritmos más detallados y asesoramiento. Es una asociación, una simbiosis de máquina y humano, cada uno complementando las fortalezas del otro, cada uno compensando las debilidades del otro.

En los anales de los desafíos tecnológicos, el Desafío de la Red DARPA 2009 ocupa un lugar especial. Fue una prueba que fue más allá del ámbito de la tecnología, adentrándose

en la esencia misma de la ingeniosidad, cooperación y habilidades de resolución de problemas humanos. El desafío era engañosamente simple en la superficie: localizar diez globos meteorológicos rojos dispersos por Estados Unidos continentales. Pero bajo esta simplicidad, se escondía una compleja red de desafíos que ponía a prueba los límites de la coordinación y colaboración humanas.

Los globos no estaban escondidos en lugares oscuros. Estaban a la vista, anclados en lugares públicos. Pero la escala misma de la tarea, la inmensidad del área dentro de la cual se dispersaron estos globos, hizo que el desafío fuera abrumador.

En muchos sentidos, el Desafío de la Red DARPA fue un micro- ►►



1 Subcontratación masiva voluntaria o colaboración masiva son alternativas en español al anglicismo *crowdsourcing*. <https://www.fundeu.es/recomendacion/crowdsourcing/#:~:text=Subcontrataci%C3%B3n%20masiva%20voluntaria%20o%20colaboraci%C3%B3n,en%20espa%C3%B1ol%20a%20anglicismo%20crowdsourcing>.

cosmos de las crisis que enfrentamos hoy. Al igual que los globos rojos, los problemas que enfrentamos (el cambio climático, las pandemias, las amenazas cibernéticas...) están esparcidos por todo el mundo. Son problemas complejos y multifacéticos que ningún individuo u organización puede resolver por sí solo. Requieren un esfuerzo coordinado y colaborativo. Nos exigen movilizar recursos, recopilar y analizar información, diseñar estrategias y adaptarnos a las circunstancias cambiantes. Nos exigen unirnos, combinar nuestros recursos y nuestro intelecto para resolver un problema que es más grande que cualquiera de nosotros.

En la actualidad, estamos realizando una simulación del Desafío de la Red DARPA para recrear este escenario en un entorno controlado. Los globos rojos son reemplazados por marcadores textuales, dispersos en un mar de datos. La tarea de la IA es generar estrategias para localizar estos marcadores, movilizar recursos y coordinar esfuerzos en una carrera contra el tiempo. Es una prueba de la capacidad de los Modelos de Lenguaje a Gran Escala (MGLs) para resolver problemas complejos del mundo real, trabajar en conjunto con los humanos y aumentar nuestras capacidades.

Los resultados de la simulación han sido a la vez iluminadores y sobrios, un testimonio del inmenso po-

tencial y las limitaciones inherentes de la IA. Resulta que los modelos son capaces de generar estrategias detalladas y matizadas. Pueden tamizar los datos, identificar patrones, anticipar posibles desafíos y proponer soluciones. Pueden pensar de manera estratégica, adaptarse a las circunstancias cambiantes y aprovechar el poder de la tecnología y el *crowdsourcing*¹. Pueden, en cierto sentido, imitar el proceso humano de resolución de problemas, pero a una velocidad y escala que supera con creces las capacidades humanas.

Herramientas

Pero la simulación también revela las limitaciones de estos modelos. Pueden procesar grandes cantidades de datos, pero carecen del toque humano, la capacidad de entender las sutilezas del comportamiento humano, la imprevisibilidad del mundo real. Pueden generar estrategias, pero necesitan la intervención humana para implementarlas, para adaptarlas a las realidades siempre cambiantes del mundo.

Pero estas limitaciones no son una condena a estos modelos, sino un recordatorio de su naturaleza. Son herramientas, no reemplazos del juicio e intuición humanos. Son ayudas, no sustitutos de la comprensión y empa-

tía humanas. A pesar de estas limitaciones, veo no solo desafíos, sino potencial. Veo un atisbo de un futuro en el que los MGL y los sistemas de IA autónomos y agéntico, como por ejemplo la novedosa plataforma AgentGPT, podrían convertirse en aliados invaluable en nuestra búsqueda para navegar las complejidades de nuestro mundo.

AgentGPT no es solo otro modelo de IA. Es una nueva categoría de IA diseñada para el comportamiento autónomo dirigido hacia metas. Es como un jugador de ajedrez, siempre pensando varios movimientos por adelantado, siempre adaptando su estrategia en función del estado del juego. Pero a diferencia de un jugador de ajedrez, AgentGPT no está confinado a las 64 casillas de un tablero de ajedrez. Puede operar en el vasto y complejo paisaje de datos del mundo real, generando estrategias, anticipando desafíos y proponiendo soluciones.

El uso de la IA, particularmente en las crisis, es una espada de doble filo. Por un lado, puede ayudarnos a responder rápidamente y de manera efectiva, a movilizar recursos, a coordinar esfuerzos. Por otro lado, también puede ser una fuente de daño, si sus resultados son poco éticos o perjudiciales. El potencial de mal uso o abuso de la IA es una preocupación real y apremiante, a la que debemos hacer frente.

Pero las salvaguardas son solo una pieza del rompecabezas. También necesitamos empujar los límites de nuestras simulaciones, probar nuestros modelos bajo condiciones cada vez más desafiantes. Necesitamos exponer nuestra IA a las complejidades e incertidumbres del mundo real, ver cómo responde, entender sus fortalezas y debilidades.

Y necesitamos llevar nuestra investigación fuera del laboratorio y al mundo.

Estamos comprometidos a garantizar que cualquier uso de la IA para manejar las crisis se mantenga centrado en el humano. Estamos comprometidos con una visión de la IA que no sea una amenaza, sino un socio, una herramienta que pueda aumentar nuestras capacidades,

que nos pueda ayudar a dar sentido al caos, que pueda ayudarnos a encontrar la señal en el ruido. Estamos comprometido con una visión de IA que respete nuestros valores, nuestra ética, nuestra humanidad.

Y así, mientras el sol del verano español se pone fuera de la ventana de mi oficina, proyectando largas sombras sobre el cuidado campus de la Universidad Carlos III de Madrid, me encuentro lleno de un sentido de propósito, de resolución, de esperanza. Los desafíos que tenemos por delante son formidables, pero también lo es nuestra resolución.

Para la redacción de este trabajo se ha utilizado como apoyo editorial la plataforma GPT4 de OpenAI y la plataforma Claude+ de Anthropic.

Bibliografía

- Cebrián M. (2021): "A time-critical crowdsourced computational search for the origins of COVID-19" en *Nature Electronics*.
- Crandall J.W., Dudah M., Tennom F., Ishowo-Oloko F., Abdallah S., Bonnefon J.-F., Cebrián M., Shariff A., Goodrich M.A. and Rahwan I. (2018): "Cooperating with machines" en *Nature Communications*.
- Kryvashyeyu Y., Chen H., Obradovich N., Moro E., Van Hentenryck P., Fowler J. and Cebrián M. (2016): "Rapid assessment of disaster damage using social media activity" en *Science Advances*.
- Pickard, G., Pan, W., Rahwan, I., Cebrián, M., Crane, R., Madan, A. and Pentland, A. (2011): "Time-critical social mobilization" en *Science*.
- Rahwan, I., Cebrián, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A., "Sandy", Roberts, M. E., Shariff, A., Tenenbaum, J. B. and Wellman, M. (2019): "Machine behaviour" en *Nature*.
- Rutherford, A., Cebrián, M., Dsouza, S., Moro, E., Pentland, A. and Rahwan, I. (2013): Limits of social mobilization, en *Proceedings of the National Academy of Sciences of the USA (PNAS)*.



CARLOS BLANCO

Profesor titular de la Universidad Pontificia Comillas. Autor de *La integración del conocimiento*

LA CAPACIDAD DE INVENTARNOS A NOSOTROS MISMOS

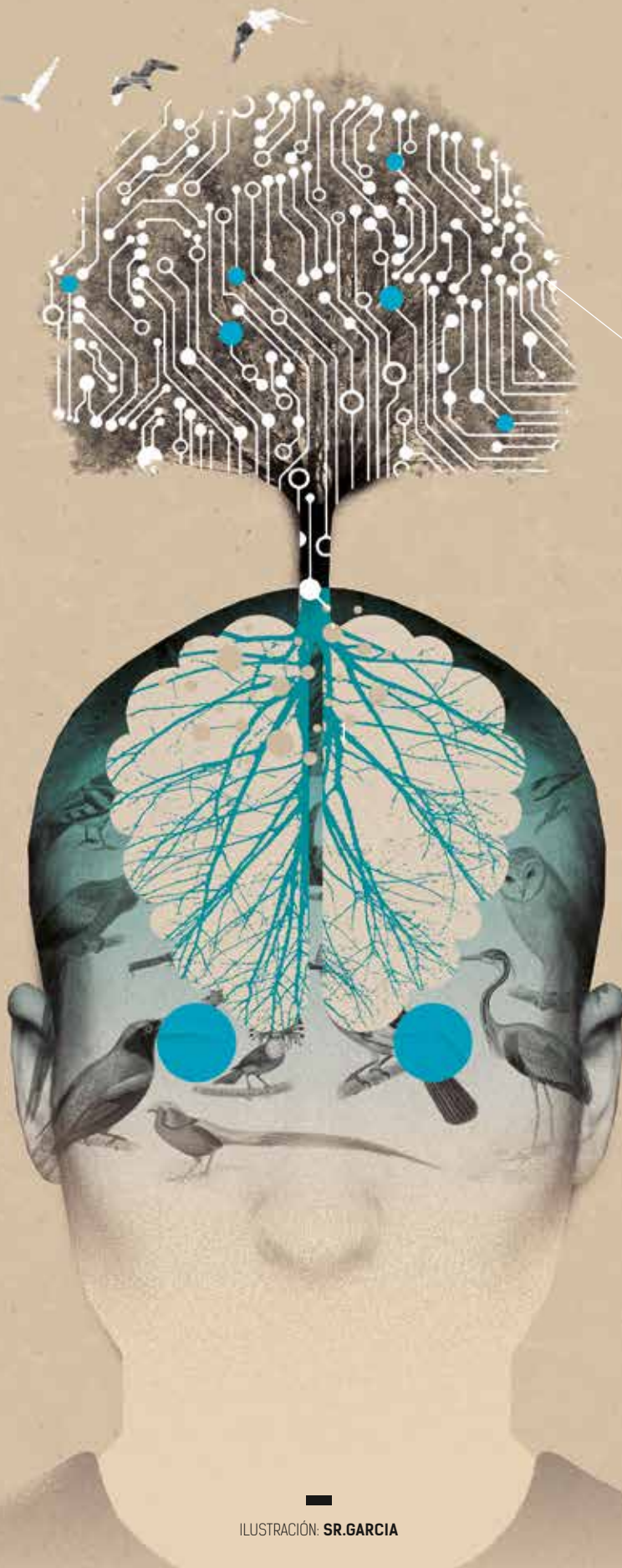
¿Qué nos hará humanos?

El debate sobre qué nos hace humanos continúa abierto. Además, tampoco tenemos por qué asumir un único concepto de lo humano. Incluso quizás debemos mirar a la inteligencia artificial no como una competidora y una amenaza, sino plantearnos qué alianzas estableceremos con ella.

The ability to invent ourselves
WHAT WILL MAKE US HUMAN?

The debate about what makes us human remains open. Nor do we have to assume a single concept of what is human. Perhaps we should even look at artificial intelligence not as a competitor and a threat, but rather consider what alliances we will establish with it.

Keywords: generative artificial intelligence, human specificity, anthropocentrism, human behavior, mental state.



Palabras clave:
inteligencia artificial generativa, especificidad humana, antropocentrismo, conducta humana, estado mental.

La pregunta por la esencia de lo humano constituye uno de los interrogantes científicos y filosóficos más profundos que cabe concebir. Los seres humanos buscamos entender el mundo, pero ante todo queremos entendernos a nosotros mismos. La famosa exhortación “*Conócete a ti mismo*”, inscrita en el frontispicio del templo de Apolo en Delfos, sintetiza la fascinación que esta cuestión ejerce sobre cualquier mente inquieta. Podemos abordarla individualmente, formulada como la pregunta sobre quién soy yo —y sobre quién puedo ser—. Sin embargo, cuando la examinamos desde la perspectiva de la humanidad, y no solo de los individuos, adquiere los visos de un interrogante verdaderamente recapitulador de los grandes problemas de la filosofía. No en vano, Kant resumía en “*Was ist der Mensch?*” (“¿Qué es el hombre?”) otras grandes incógnitas filosóficas como qué podemos saber, qué debemos hacer y qué nos cabe esperar.

Con el desarrollo de la inteligencia artificial generativa (IAG), esta cuestión cobra una nueva relevancia. Uno de nuestros artefactos, el más complejo de cuantos ha inventado el ingenio de la humanidad, ha adquirido tal grado de sofisticación que exhibe habilidades hasta ahora atribuidas exclusivamente a nuestra especie. ¿Hasta qué punto puede decirse que los humanos somos singulares, si una de nuestras creaciones probablemente nos supere en el plano que parecía definirnos de manera más nítida y profunda, que es el intelectual?

No obstante, la pregunta por la especificidad humana no se circunscribe al análisis de la inteligencia artificial. Cabe plantearla a propósito de otras formas biológicas. El avance de la investigación nos permite percatarnos de que muchas de las cualidades y conductas exclusivas del hombre ampliamente predicadas pecan de un evidente sesgo antropocéntrico¹. Por ejemplo, no está claro

que seamos los únicos animales con conciencia de la muerte, ni los únicos capaces de desarrollar cierto tipo de simbolismo y de cultura².

El propio esquema evolutivo ofrece un argumento sumamente poderoso para ello. Aunque en la naturaleza existen procesos de todo o nada —como el potencial de acción en el sistema nervioso—, en la escala filogenética es difícil establecer rupturas absolutas entre especies. Ciertamente, hay saltos notables en la trama de la vida (muchas veces motivados por presiones evolutivas externas, como los cambios climáticos, que aceleran las transformaciones), pero si todas las especies provienen de un origen común, es razonable suponer una continuidad estructural y funcional entre ellas. Ni siquiera la conciencia puede considerarse una brecha tan radical con nuestros ancestros evolutivamente más próximos. Quizá el lenguaje articulado siga siendo nuestro patrimonio. Sin embargo, con el progreso de la etología y con la creación de teorías de la mente animal más elaboradas es difícil atreverse a sostener, de manera categórica, que solo los humanos somos capaces de desplegarlo. Al fin y al cabo, comprendemos muy poco sobre otros sistemas de comunicación presentes en el dominio de la vida.

La naturaleza genera escalas de complejidad. Cuesta identificar algo exclusivamente humano más allá de nuestra singular dotación genética. La especificidad humana no puede sustentarse en nuestro peculiar rango de habilidades emocionales e intelectuales, sino más bien en su grado de desarrollo y de compenetración. Poseemos, sin duda, el cerebro más complejo conocido del reino animal, así como el mayor cociente de encefalización. En nuestro cerebro destaca la asombrosa expansión de las cortezas prefrontales. Lo que mejor responde a la pregunta de por qué nos hace humanos, probablemente sea el nivel de

1 En palabras de Kristin Andrews y Susana Monsó: “*Animal cognition research challenges philosophers to consider that many capacities and behaviors often assumed to require language, sophisticated technological capacities, or legal systems may in fact be had by other animals who lack these properties. In this way, animal cognition research often surprises us by showing that sophisticated-looking activity can be caused through rather simple mechanisms*” (“*Animal cognition*”, *The Stanford Encyclopaedia of Philosophy*).

2 Para una introducción al debate sobre el concepto de *cultura* aplicado a animales no humanos: Laland, K. N. y Galef, B. G. (eds.) (2009): *The question of animal culture*, Harvard University Press.

desarrollo del córtex prefrontal, cuyo crecimiento posibilita el alto desarrollo de nuestra mente.

Estoy convencido de que la capacidad de abstraer hasta nociones puramente formales de la lógica, la filosofía y la matemática es, con bastante seguridad, la rúbrica de nuestra singularidad biológica, pero no es descartable que mentes de otras especies puedan acercarse a ellas, o al menos esbozarlas.

Lo anterior sirve para reflexionar sobre la inteligencia artificial generativa. Nos atemoriza perder especificidad, pero de nuevo, no está claro que la hayamos tenido del todo. Es una ilusión antropocéntrica pensar que somos verdaderamente únicos, una especie enajenada del árbol de la vida. Todas las especies que existen han sido elegidas por la selección natural; en términos biológicos, ninguna ostenta un privilegio.

Prolongación humana

No se me ocurre nada humano no reproducible en última instancia. Una prolongación de lo humano —como lo es la inteligencia artificial— podría perfectamente adquirir vida propia, una autonomía creadora, que ya no se limitase a seguir nuestros programas de instrucciones, sino que fuera capaz de tomar decisiones por sí misma.

Tendemos a suponer que las máquinas inteligentes, incluso las de nueva generación, poseen habilidades de cómputo y de procesamiento de información superiores a las nuestras, si bien carecen de imaginación, creatividad y pensamiento crítico (de *interioridad consciente*, en suma). Probablemente sea cierto por ahora, pero no es descartable que en algún momento la inteligencia artificial alcance tal grado de desarrollo que nos aventaje también en estas facetas. Incluso el ámbito de las emociones, que parece tan propiamente

humano (aunque, ¡oh paradoja!, es justo aquí donde la cercanía con otros mamíferos es mayor, pues sabemos que la emotividad no es un atributo nuestro, y no exige disponer de un cerebro tan complejo como el de los humanos modernos) lo consideramos inasequible para una máquina inteligente. La concebimos como una mente pura que, al carecer de cuerpo, no puede expresarse emocionalmente, restringida a encadenar inferencias lógicas, cadenas de premisas y consecuencias hilvanadas por reglas formales; un ciego automatismo vacío de vida.

De nuevo, no se ha demostrado que haya impedimento filosófico o científico para que una máquina pueda también desarrollar estados mentales (es decir, una capacidad de representarse el mundo y de referir sus percepciones a ella misma, a su mundo interior), y pueda aproximarse a la realidad no solo con la frialdad de la lógica, sino también con el calor de las emociones, sentidas como propias. Ninguna ley de la lógica y de la naturaleza parece prohibir esta posibilidad tan intrigante; tampoco leyes *a priori* del pensamiento³. Si la evolución biológica lo ha logrado, aunque sea tras millones de años de variaciones aleatorias y selección natural, ha tenido que seguir un mecanismo. La mente humana puede descifrarlo y reproducirlo. ▶▶▶

El avance de la investigación nos permite percatarnos de que muchas de las cualidades y conductas del hombre pecan de un evidente sesgo antropocéntrico

3 Quienes adopten una posición dualista, o reminiscente del dualismo clásico, considerarán la conciencia como algo irreducible al mero procesamiento de información, y subrayarán la esfera subjetiva (el célebre “para sí”, o “ser como...”, de la experiencia fenoménica, contemplada como percepción subjetiva del mundo). Sin embargo, en las últimas décadas se han desarrollado interesantes modelos neurobiológicos sobre la conciencia que explican sus características fundamentales desde un paradigma netamente computacional, o compatible con este tipo de aproximaciones al estudio de la mente. Como escriben Stanislas Dehaene, Hakwan Lau y Sid Kouider: “*Although centuries of philosophical dualism have led us to consider consciousness as unreducible to physical interactions, the empirical evidence is compatible with the possibility that consciousness arises from nothing more than specific computations*” (Dehaene, S., Lau, H. y Kouider, S.: “What is consciousness, and could machines have it?”, *Robotics, AI, and Humanity: Science, Ethics, and Policy* (2021), p. 56). Por tanto, y aunque el escepticismo en torno a la dimensión subjetiva de los actuales desarrollos en inteligencia artificial esté más que justificado, si la conciencia es explicable científicamente, “objetivamente”, y en consecuencia reducible a un proceso físico, químico y biológico (“naturalizable”, en definitiva), donde lo subjetivo emerge a partir de mecanismos objetivos, no hay obstáculo alguno para reproducir esas cualidades y diseñar sistemas que las posean, porque cabe asumir puntos de vista menos “maximalistas” sobre la naturaleza y las propiedades del pensamiento consciente (cf. Carlos Blanco Pérez (2019): “Pensamiento, creatividad y máquinas”, *Naturaleza y Libertad. Revista de estudios interdisciplinarios* (121), más allá de la mistificación filosófica a la que muchas veces se ha tendido.



Lo que nos hace humanos es precisamente esta búsqueda, este afán de superación incesante, este anhelo de trascender las fronteras de lo biológico para crear nuestro propio mundo

Sin duda, se trata de una aspiración colosal. Probablemente estemos ante la tarea más ambiciosa que hayamos iniciado nunca. Implica tomar el relevo a la naturaleza. Nace de la pretensión de crear una mente no ya similar a la nuestra, sino, por qué no, superior en todos los sentidos: emocional, intelectual... Pero ¿no es la cultura humana nuestro deseo insaciable de conocer, descubrir e inventar, una lucha titánica por elevarnos más allá de lo que la naturaleza nos ofrece (propósito que, quizá albergaron también otras especies, dueñas de un cerebro complejo, como el hombre de Neandertal)? ¿No es la inteligencia artificial la continuación lógica de ese empeño heroico? Porque no es osado afirmar que lo humano, lo que nos hace humanos, es precisamente esta búsqueda, este afán de superación incesante, este anhelo de trascender las fronteras de lo biológico para crear nuestro propio mundo.

Por tanto, no podemos descartar que el progreso tecnológico alumbré máquinas no solo inteligentes, sino también creativas. La creatividad trasciende la inteligencia; exige inteligencia, pero va más allá de ella, pues

consiste en añadir, en innovar. Una máquina que desarrolle la habilidad de procesar emociones quizá se halle más capacitada para ser creativa, porque no se limitará a discurrir con las herramientas de la lógica, a razonar a partir de premisas y reglas de inferencia, sino que, al sentir e imaginar, al apropiarse intuitivamente del mundo, es posible que se lance a inventar su genuino mundo, como hemos hecho los seres humanos.

En mi opinión, herramientas como ChatGPT, incluso en sus versiones más avanzadas, no cumple estos requisitos. Siendo realistas, puede considerarse inteligente solo por analogía. Satisface una definición demasiado débil de inteligencia, interpretada como capacidad de resolver problemas. Procesa cantidades ingentes de información y destila multitud de respuestas, pero en su estado actual de desarrollo falta percibir ese pensamiento crítico, esa chispa y esa creatividad que definen la inteligencia humana, cuyo poder no se limita a solucionar problemas, sino que ante todo los crea, pues pregunta y no solo responde. Porque la pregunta, ¿no nos hace aún más humanos que la respuesta?

Una inteligencia artificial verdaderamente sensible y creativa podría, por qué no, aventurarse a cultivar su arte, su religión y su propia concepción filosófica del mundo (por retomar las tres célebres determinaciones supremas del espíritu, sobre las que tan brillantemente teorizó Hegel): su propio universo interior, el *intus* de su *legere*, su capacidad de leer el mundo para sí misma. Solemos pensar que el arte, la religión y la filosofía, que en último término remiten a una honda y poderosa aptitud para expresarnos subjetivamente, son prerrogativa de nuestra especie. Nos distinguiría de otros animales y de unas hipotéticas máquinas creativas, dotadas de cierto grado de subjetividad (aunque aquí cabría preguntarse si, para ser creativos, es imprescindible poseer estados mentales; no lo sabemos). Sin embargo, volvemos a navegar entre tinieblas. La complejidad de nuestro simbolismo nos enorgullece, pero también podría ser desarrollada espontáneamente e incluso rebasada por esas máquinas. La prudencia, el sentido común y el espíritu de finura pascaliano, que se superpone al espíritu meramente geométrico y calculador, podrían también suscitarse en una inteligencia artificial.

Errores y aciertos

¿Qué decir del error? Nuestros errores nos hacen tan humanos como nuestros aciertos. La conciencia de que somos falibles, de que nuestros errores nos definen y también impulsan, se nos antoja inasequible para una máquina tan perfecta, tan connaturalizada con las reglas de la lógica, que esquivaría la posibilidad de errar. No obstante, es concebible que esa máquina, por “perfecta” que fuera, desplegara conductas más flexibles, hasta tolerar las imperfecciones y aprender

de ellas, pues probablemente semejante grado de falibilidad sea tan eficiente en términos adaptativos como la inerrancia absoluta.

¿Y el mal? Los seres humanos categorizamos el mundo según el esquema bien/mal. Más allá de la interpretación filosófica que adoptemos en torno a los fundamentos de este binomio, resulta innegable que el mal, lo que nos perjudica como individuos y como comunidad humana, lo que limita nuestras posibilidades y conduce al conflicto, ha sido y es definitorio de nuestra especie y de nuestra conciencia moral. Ser humano es enfrentarse a la disyuntiva entre el bien y el mal. Entraña reflexionar, plantearme qué debo hacer, y por qué a veces no hago lo que mi conciencia me dicta. ¿Una inteligencia artificial podría tener voluntad y libre arbitrio? ¿Emergería en ella la conciencia moral? ¿Podría optar conscientemente por lo malo y rechazar deliberadamente lo bueno? ¿Cómo categorizaría el mundo? ¿Tendría las mismas nociones de bien y de mal que nosotros? ¿Sería tan perfecta que solo se orientaría hacia lo bueno, como guiada por una luz irresistible?

Una forma superior de pensamiento presumiblemente sería menos egoísta que nosotros, por poseer una visión más completa y profunda de la realidad. Sin embargo, y por retomar consideraciones anteriores, quizá también la evolución de esa inteligencia artificial hacia cotas más creativas de comportamiento exija el desarrollo de un mundo moral propio, así como de la posibilidad de elegir lo malo.

El debate sobre qué nos hace humanos continúa abierto. Además, tampoco tenemos por qué asumir un único concepto de lo humano. Caben diferentes aproximaciones que subrayen una u otra dimensión de nuestro ser. Hay *estilos* en la interpretación de lo humano, una legítima diversidad cultural

y filosófica que inspira un saludable *humanismo pluralista* a la hora de asomarnos a la complejidad de lo que somos.

En cualquier caso, desde un punto de vista neurobiológico parece que casi todas nuestras habilidades más distintivas hunden sus raíces en una estructura anatómica tan deslumbrante como el córtex prefrontal. Este sería el pilar neurobiológico del genio humano. Aun así, qué rasgos y qué conductas son absolutamente exclusivos de nuestra especie, tanto como para diferenciarnos de una inteligencia artificial en sentido fuerte, es un interrogante difícil de responder de manera satisfactoria. En cuanto señalamos algo que parecemos monopolizar (la conciencia moral, la religión...), es posible indicar contraejemplos biológicos que, si bien no de manera inequívoca, apuntan a la existencia de un fenómeno similar. Y en cuanto al lenguaje articulado, es precisamente aquí donde una inteligencia artificial puede emularnos y superarnos.

Creo, por tanto, que más que preguntarnos por lo que nos hace humanos deberíamos hacerlo *por lo que nos hará humanos*; porque la capacidad de inventarnos a nosotros mismos es consustancial a lo humano. Evoca la posibilidad de ser no solo esencia, sino también proceso: proyecto inacabado libertad más que destino. Después de todo, quizá debamos mirar a la inteligencia artificial no como una competidora y una amenaza, sino plantearnos: qué alianzas estableceremos con esta creación, con este fruto de nuestro inagotable ingenio, para progresar y trascender nuestros límites.

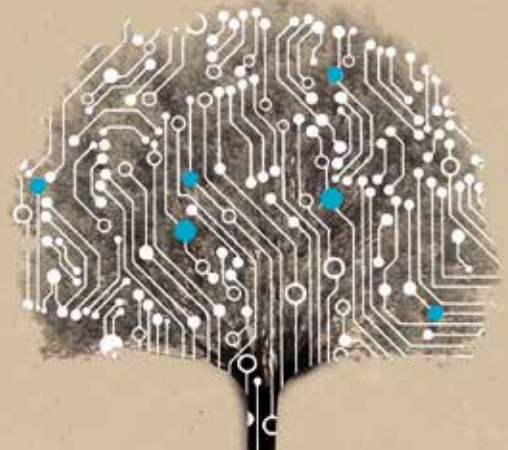
Bibliografía

Andrews, K. y Monsó, S. (2021): "Animal cognition" en *The Stanford Encyclopaedia of Philosophy*.

Blanco Pérez, C. (2019): "Pensamiento, creatividad y máquinas". *Naturaleza y Libertad. Revista de estudios interdisciplinarios* (2019)/12.

Dehaene, S., Lau, H. y Kouider, S. (2021): "What is consciousness, and could machines have it?" en *Robotics, AI, and Humanity: Science, Ethics, and Policy*.

Laland, K. N. y Galef, B. G. (eds.) (2009): *The question of animal culture*. Cambridge, Massachusetts, Harvard University Press.





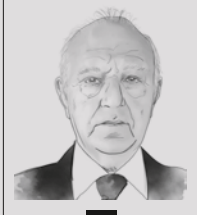
Palabras clave:

consciencia, inteligencia, inteligencia artificial, era axial, neurona artificial, redes neuronales artificiales, inteligencia artificial general.

**LA
CONSCIENCIA
HUMANA ANTE
LA IRRUPCIÓN
DE LA
INTELIGENCIA
ARTIFICIAL
GENERAL (IAG)**

Consciencia e IA

La consciencia y la inteligencia artificial han sido siempre dos temas muy distintos. La primera altamente subjetiva y conectada a lo más profundo de la mente humana, y la segunda muy material, física y producto de la acción humana observable y medible. Aunque siempre ha habido preocupación por la posible relación entre ambas, es ahora cuando hay un interés especial desencadenado por la posibilidad de que las máquinas puedan llegar a ser conscientes como los hombres. En lo que sigue, se analizan algunos aspectos relativos a estas cuestiones, especialmente en cuanto a lo que es la consciencia y lo que puede terminar siendo la inteligencia artificial.



ADOLFO CASTILLA.
Doctor Ingeniero del ICAL
y catedrático de Economía Aplicada

Human consciousness in front of the irruption of artificial general intelligence (AGI)
CONSCIOUSNESS AND AI

Consciousness and artificial intelligence have always been two very different subjects. The first is highly subjective and connected to the depths of the human mind, and the second is very material, physical and the product of observable and measurable human action. Although there has always been concern about the possible relationship between the two, it is now that there is a special interest triggered by the possibility that machines may become conscious like humans. In what follows, some aspects relating to these issues are discussed, especially as regards what consciousness is and what artificial intelligence may end up being.

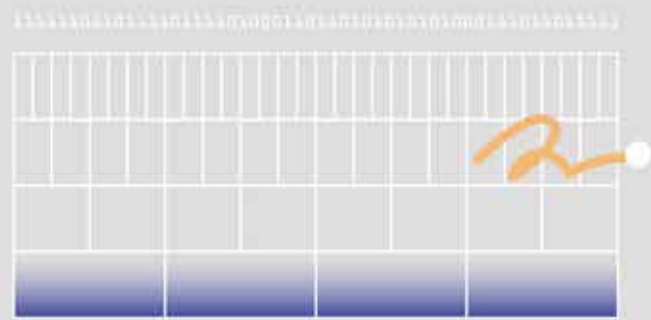
Keywords: *consciousness, intelligence, axial era, artificial neuron, artificial intelligence general, artificial neuron, artificial neural networks, general artificial intelligence.*



Cia



ILUSTRACIÓN: ISABEL ALBERTOS



N

No hay temas más dispares *a priori* que los dos tratados en este trabajo. La inteligencia artificial es solo *software*, programación e instrucciones para un ordenador hechas por el hombre, mientras que la consciencia es pura subjetividad, lo más interno e inmaterial de la mente de los humanos, que además no sabemos todavía cómo surge o se genera.

Las dos se unen en nuestros días, sin embargo, en el debate actual sobre si la inteligencia artificial puede llegar a tener autoconsciencia como la tienen el hombre y la mujer.

Escribir sobre las dos cuestiones y su interrelación es algo arriesgado a la vista de los cientos, por no decir miles de artículos, libros, memorias de investigación y *papers* de todo tipo publicados en los últimos tiempos. La mayoría de ellos accesibles en Internet. ¿Es posible decir algo nuevo?

Como se sabe, la consciencia es un área de reflexión muy antigua. Surge con la aparición del pensamiento en la Era Axial (800 a. de C. – 200 a. de C.) según Karl Jaspers (1883-1969). De la mano, en gran manera, de Platón, acompañado de otros pensadores de la época como Parménides, Heráclito o Tucídides, y precedida por la labor de los presocráticos.

No sabemos de dónde procede y qué genera la consciencia.

Esa cuestión de “ser conscientes” es la que todavía no sabemos de dónde procede y cómo se produce en la mente humana. A pesar de los milenios transcurridos, el ser humano no ha avanzado demasiado en el terreno de los procesos mentales de este tipo,

a los que hay que incorporar, además, a los subconscientes y a los inconscientes, así como a los sentimientos más profundos, a la intuición, la imaginación, la inspiración, la creatividad, la inteligencia misma y, en parte, el aprendizaje. Sabemos lo que es la consciencia, pero no cómo aparece en el hombre y qué la genera.

No podemos hacer aquí una revisión detallada de todo lo que el hombre ha aportado en relación con la consciencia a lo largo del tiempo, pero sí deberíamos recordar la labor de René Descartes (1596-1650) en cuanto a la distinción de dos sustancias en el hombre: la pensante, o *res cogitans*, que es la más importante, la que nos hace hombres y la que nos distingue de los animales; y la sustancia extensa, o *res extensa*, formada por el cuerpo y su conexión con el mundo material y físico, el tiempo y otros componentes del universo en el que vivimos.

Dualismo

Este destacado autor fue el introductor del dualismo, aunque parece que él creía en la unión íntima en el hombre de las dos sustancias. A pesar de ello, la historia posterior interpretó sus aportaciones como la existencia de dos “naturalezas”, en vez de dos sustancias, y que había diferencias profundas entre ellas, hasta el punto de que la primera era el ámbito de la fe y las creencias, y la segunda el de la objetividad y la ciencia.

Pronto surgieron críticas a las explicaciones de Descartes, especial-

1 Allan Turing (1912-1954) fue un notable investigador británico, matemático, lógico, informático teórico, criptógrafo, filósofo y biólogo teórico, que después de haber prestado grandes servicios a su país, entre ellos su contribución a descifrar los códigos empleados por los nazis en su mensajería encriptada, especialmente los de la máquina Enigma, construir dos de los primeros ordenadores, crear la idea de programación continua, trabajar teóricamente en la inteligencia artificial y muchas otras aportaciones, fue perseguido por homosexual, encarcelado y castrado químicamente, suicidándose a la temprana edad de 41 años.



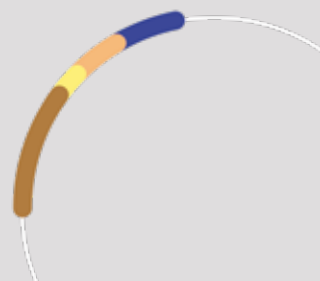
mente por parte del empirismo inglés, que creía en el cerebro humano como una *tabula rasa* de partida, que aprendía a través de su interrelación con su hábitat. Es lo que se llamó monismo del que ha habido diversas interpretaciones también. Hoy, por ejemplo, hablamos del monismo trascendental, que considera que las dos sustancias son componentes indisolubles de la naturaleza humana, como las dos caras de una misma moneda, pero con capacidades de la componente mental que aún no podemos explicar y que nos conectan a una posible consciencia cósmica.

Es curioso, pero a mediados de la década de los años 50 se produjo la interrelación del pensamiento humano con el mundo de los ordenadores o, dicho de otra forma, las ciencias cognitivas incorporaron a sus intereses la programación de la lógica humana en las computadoras. Como fecha de comienzo de la inteligencia artificial se suele indicar el verano de 1956 en el que se organizó una primera reunión de personas destacadas, que se distinguían por estar trabajando en ordenadores, y más específicamente en su programación. El trascendental encuentro en el Dartmouth College, en New Hampshire (Estados Unidos) ocurrió dos años después de la trágica muerte de Alan Turing¹. John McCarthy (1927-2011), profesor de dicho centro, convocó una conferencia a la que asistieron: Claude Shannon (1916-2001), Marvin Minsky (1927-2016), Frank Rosenblatt (1928-1971), Herbert Simon (1916-2001) y Allen Newell (1927-1992). En ella, y

En la Conferencia de Dartmouth se definió la IA como el proceso de hacer que una máquina se comporte de formas que serían llamadas inteligentes si un ser humano las hiciera

especialmente en la documentación elaborada para pedir financiación, se utilizó la expresión inteligencia artificial. Hoy sabemos, y está todo documentado, que fue una idea feliz de McCarthy (1927-2011), quien había trabajado bastante en el tema antes de la reunión y la había definido como “el proceso de hacer que una máquina se comporte de formas que serían llamadas inteligentes si un ser humano las hiciera”.

Es verdad que muchos nos preguntamos por la verdadera inteligencia, que conectamos de forma muy clara a la consciencia. Aquí está el *bard problem* de la inteligencia artificial como lo denomina el filósofo, australiano de nacimiento, pero estadounidense de residencia, David Chalmers (nacido en 1966). ■▶▶



Sabemos lo que es la consciencia, pero no cómo aparece en el hombre y qué la genera

Hay un punto al que, de momento, no llegan los ordenadores: el de la consciencia, la introspección y la actividad mental superior, entre otras cosas porque no sabemos lo que es ya que los procesos conscientes (y los inconscientes) no son algorítmicos y porque no existe (en realidad, es una ilusión).

Y lo cierto es que ha habido inviernos y veranos en esta disciplina, es decir, épocas con grandes expectativas y épocas de grandes decepciones. La actual parece claramente del primer tipo, sobre todo por la aparición del *machine learning*, el *deep learning* y el *learnig reinforcement*. Así como por el gran impacto social que ha tenido la llamada, inteligencia artificial general (IAG)², la que hemos llamado hasta ahora inteligencia artificial fuerte.

Basándose en esas formas de hacer que las máquinas aprendan y en los avances impresionantes llevados a cabo en la programación, la empresa OpenAI y otras están lanzando al mercado productos deslumbrantes como ChatGPT y DALL·E 2. Esta y otras empresas, menos conocidas y más especializadas, están creando una verdadera revolución en las sociedades actuales de todo el mundo. Algunos hablan ya de la V Revolución Industrial y anuncian cambios superiores a la introducción de la electricidad y el motor eléctrico.

A pesar de esos progresos, hoy se considera mayoritariamente que las máquinas no son conscientes y que no podrán serlo a menos que se avance mucho, y no solo en el terreno de la tecnología digital. Habrá que complementarlas con lo que hoy empezamos a llamar tecnología biodigital, que tiene que ver, muy ampliamente, con lo

que creemos que es el hombre: interconexión profunda de ser humano-naturaleza-tecnología.

Podemos concluir recopilando las posiciones de los hombres hoy en cuanto a la consciencia y las máquinas con las siguientes posiciones: 1) la de los que creen que la consciencia no existe y que en realidad es una ilusión; 2) la de los que creen que es un epifenómeno del cerebro, algo así como el vapor de agua que se produce al calentar el líquido elemento; 3) la de los que creen que solo hay cerebro, es decir, materia, y antes o después descubriremos cómo se produce la consciencia; 4) la de los que creen que la consciencia no es algorítmica y que hay que utilizar la mecánica cuántica para su explicación; 5) la de los que creen en la existencia de una consciencia cósmica o universal, lo que presupone la existencia de un ser superior, de la que procede la consciencia humana.

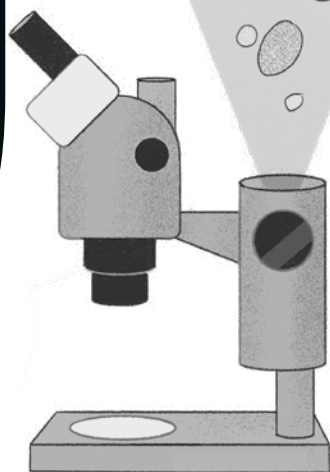
² Algunos la llaman, erróneamente, inteligencia artificial generativa, mezclando en esta denominación lo que en realidad es ChatGPT: un *Generative Pre-trained Transformer*.

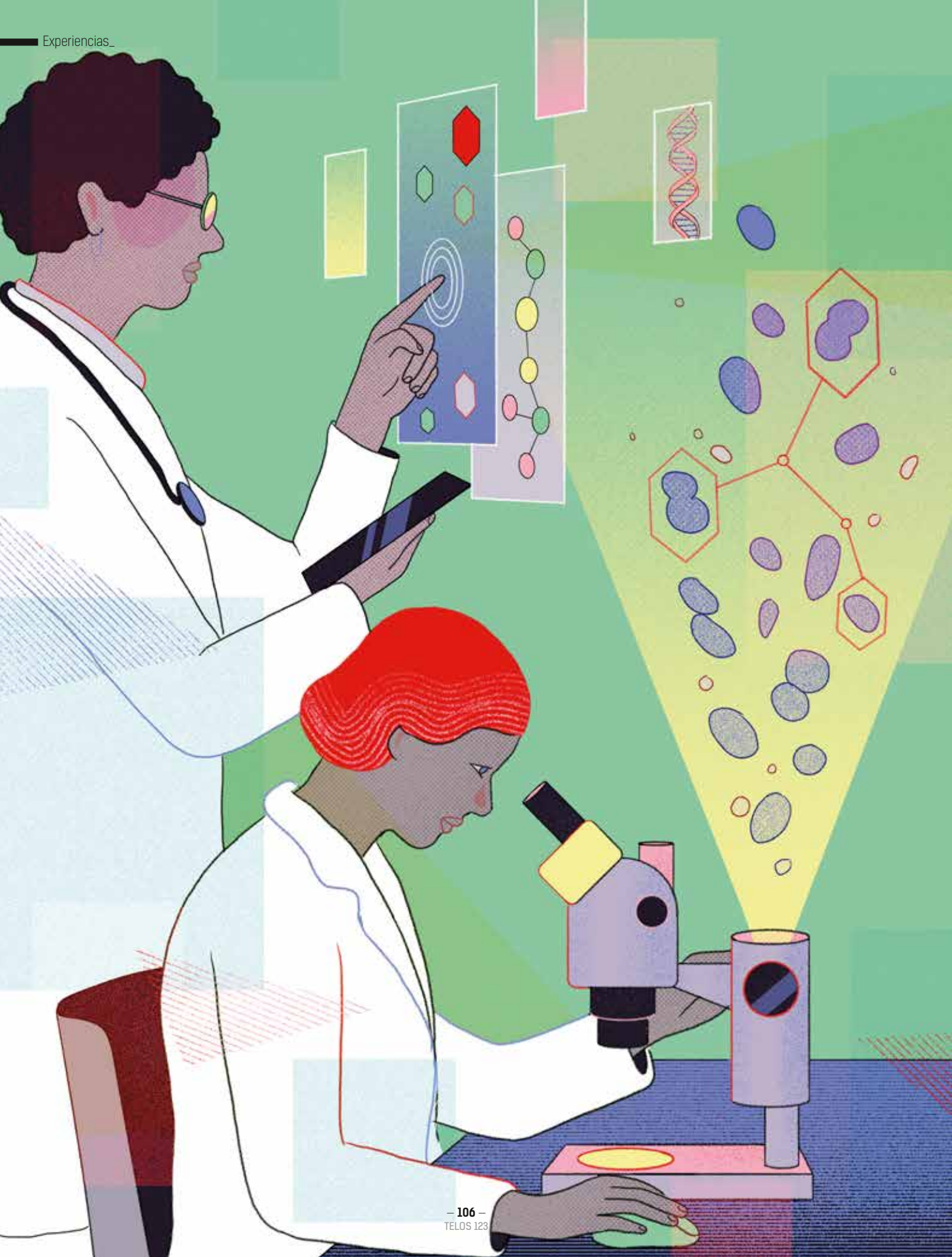
Bibliografía

- Arias, A. (2021): *Introducción a la ciencia de la consciencia. El estudio de la experiencia subjetiva en filosofía, psicología y neurociencias*. Madrid, Los Libros de la Catarata.
- Campillo, J. E. (2021): *La consciencia humana. Las bases biológicas, fisiológicas y culturales de la consciencia*. Barcelona, Arpa.
- Koch, C. (2004): *La consciencia. Una aproximación neurobiológica*. Barcelona, Ariel.
- Monserrat, J. (2015): *El gran enigma: ateos y creyentes ante la incertidumbre del más allá*. Madrid, San Pablo.
- Neumann, E. (1970): *The Origins and History of Consciousness*, Princeton, N.J., Princeton University Press.
- Penrose, R. (1996): *Las sombras de la mente*. Barcelona, Crítica.
- Russell, S. y Norvig, P. (2014): *Inteligencia artificial. Un Enfoque Moderno*. Madrid, Pearson.
- Tibayrenc, M. y Ayala, F. (2021): *Lo que nos hace humanos. Biología, medicina, lenguaje, mente, ética y religión*. SalTerrae-Cornillas

La IA aplicada a la salud ha revolucionado diagnósticos, tratamientos y gestión de datos médicos. Ofrece mejoras significativas en la atención, el análisis y la predicción de enfermedades, pero también plantea desafíos éticos y de privacidad.

Experiencias







JUAN CARLOS F. GALINDO
Periodista, divulgador
y comunicador de empresas
como Telefónica o Sherpa.ai

**DESARROLLOS Y MIEDOS DE LA
APLICACIÓN DE LA IA A LA SALUD**

Curarnos, o no, con IA: *Is this the question?*

Que la inteligencia artificial puede ayudar a diagnosticar, a tratar, a descubrir nuevos medicamentos, a hacerlo de forma más segura, e incluso a que sean más eficaces, es algo en lo que todos podemos estar de acuerdo. Pero los miedos, los riesgos, los sesgos que nadie sabe todavía cómo controlar o gestionar pesan más todavía en esa balanza y lo que provoca es que no se utilice más la inteligencia artificial para mejorar la vida de las personas en su salud. Yo solo voy a mostrar algunos de los desarrollos, y algunas de las dudas, que hoy son una realidad, y que son tan solo un ejemplo de lo que científicos y empresas más o menos conocidas han puesto ante nuestros ojos. Al menos, démosles también el beneficio de la duda para que nos ayuden.

 **Palabras clave:** inteligencia artificial, salud, privacidad, *machine learning*, aprendizaje federado.

Developments and fears of the application of AI in healthcare
TO HEAL US, OR NOT, WITH AI: IS THIS THE QUESTION?

That artificial intelligence can help diagnose, treat, discover new drugs, make them safer, and even more effective, is something we can all agree on. But the fears, the risks, the biases that nobody knows yet how to control or manage weigh even more in that balance and what causes is that artificial intelligence is not used more to improve the lives of people in their health. I am only going to show some of the developments, and some of the doubts, which are now a reality, and which are only an example of what scientists and more or less well-known companies have put before our eyes. At the very least, let us also give them the benefit of the doubt to help us.

Keywords: artificial intelligence, health, privacy, machine learning, federated learning.

ILUSTRACIÓN: ANA GALVAÑ



Mil quinientas palabras. Ese es el espacio que tengo para describir lo que la inteligencia artificial (IA) está haciendo o podrá hacer por y para mejorar la salud de las personas. No es mucho, pero voy a intentar describir algunas de las experiencias y de las empresas que han volcado algunos de sus esfuerzos en los beneficios de esta tecnología, en lugar de desarrollar aplicaciones que, en la práctica, no generan un valor o valores para los humanos y la sociedad y solo animan titulares sensacionalistas y alimentan a los detractores de los avances y las innovaciones. Una de las cuestiones que rodea toda esta carrera, no obstante, es si ética y filosóficamente la IA y la salud deben caminar unidas, o no.

No sé cuánta gente habrá oído hablar del síndrome de la médula anclada. Es una de esas tantas enfermedades raras —como el síndrome de Marfan, un trastorno poco común del tejido conectivo que se encarga de unir o ligar entre sí a los demás tejidos, brindando sostén y nutrición— de las que hay contados casos en el mundo, y que para diagnosticar y tratar se necesitan años y mucho más dinero para poder generar alguna expectativa de éxito en su tratamiento.

Leí sobre el síndrome de la médula anclada hace escasas semanas en la prensa. En concreto, un diario deportivo se sorprendía ya que la aplicación de moda en IA generativa, ChatGPT, había sido capaz de

diagnosticar en un niño de cuatro años lo que no habían sido capaz de hacer 17 especialistas con anterioridad en un periplo de más de tres años. La tan mentada *app* de OpenAI no solo diagnosticó la enfermedad, sino que fue capaz de señalar que el niño podría recuperarse con cirugía y rehabilitación.

Pero dejando a un lado la moda de la IA generativa, y poniendo en solfa algunas de las experiencias que algunas empresas trabajan desde hace años en el campo de la salud y la inteligencia artificial, y siguiendo por el mismo camino de las enfermedades raras, me gustaría destacar una que desarrolla una empresa española: Sherpa.ai, empresa emergente vasca que ha sido capaz de desarrollar una plataforma de inteligencia artificial que soluciona el gran interrogante o la gran duda que aducen muchos científicos a la hora de decidirse a trabajar con IA en el ámbito de la salud.

Sherpa.ai lleva varios años desarrollando su plataforma de IA aplicada a la privacidad de datos. ¿Y para qué? El mejor ejemplo para responder a esta pregunta lo encontramos en el ámbito de la salud, y más concretamente en el de las enfermedades raras.

Imaginemos que queremos encontrar un tratamiento para una enfermedad que hasta ahora no tiene cura. Para ello necesitamos que la inteligencia artificial nos ayude a detectar lo que el ojo humano no puede ver, como por ejemplo patrones o

biomarcadores. Pero para que el algoritmo sea efectivo y encuentre esos patrones o biomarcadores, necesita aprender de datos de pacientes, que en la mayoría de los casos son privados y confidenciales, para después poder sacar las conclusiones que nos ayuden a encontrar la cura.

Pero es que esos datos, además de ser normalmente privados, están dispersos en diferentes hospitales, centros de investigación u organizaciones de salud muchas veces, e incluso en diferentes países con diferentes regulaciones.

La manera tradicional de hacer que el algoritmo aprenda de todos esos datos privados y dispersos sería enviarlos a un servidor común y entrenarlo. Pero si se hiciera de esta manera, se vulneraría la privacidad de los pacientes, no se cumpliría con la regulación actual y mucho menos con la futura. Y si hablamos del gran consumo de energía que requeriría esta forma de hacerlo, mejor ...

Riesgo de dislexia

La plataforma de Sherpa.ai ha logrado hacer todo ello de tal manera que no es necesario compartir los datos, ya que trabaja bajo el paradigma o la técnica de aprendizaje federado que, combinado con otras tecnologías, asegura que se preserve la privacidad. Su *software* permite que el entre-

namiento se distribuya y se produzca en local. Por lo tanto, ningún dato del paciente, historial médico o radiografía sale del hospital.

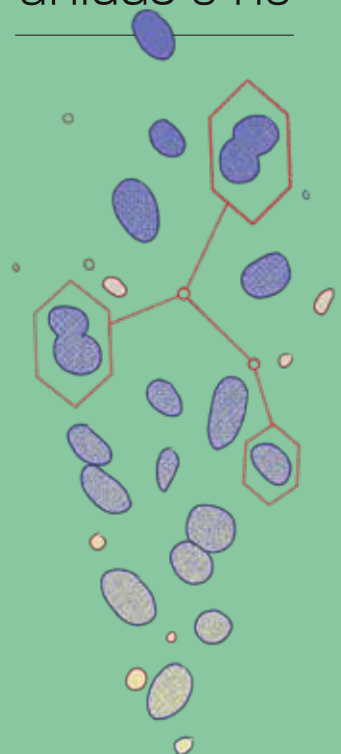
A finales del año pasado, Sherpa.ai anunció un acuerdo con el NIH o Servicio de Salud de Estados Unidos para diagnosticar y tratar enfermedades raras.

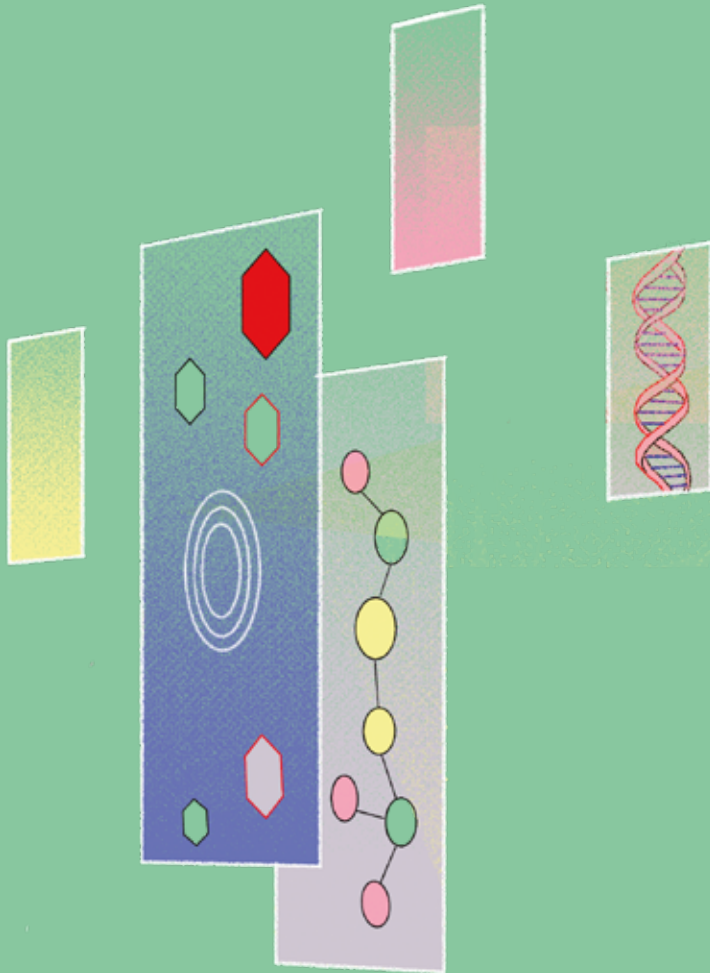
Sherpa.ai también fue capaz, en plena pandemia de COVID-19, y utilizando su plataforma de IA con algoritmos de *Machine Learning* o aprendizaje automático, de reconocer patrones y tendencias del virus, así como identificar datos de vital importancia para los servicios de salud, como tendencias en las tasas de infección y futuros brotes, entre otras funcionalidades, con el objetivo de identificar las necesidades de las UCI en el servicio de Salud del País Vasco con siete días de antelación.

Dyetective es una aplicación basada en IA para *tablets* que integra juegos lingüísticos y de atención para la detección temprana de riesgo de dislexia, y que realiza en tan solo 15 minutos y con un 90 % de precisión. La aplicación es gratuita y multiplataforma —tanto para Android como para iOS— y es el primer detector en línea de riesgo de dislexia disponible de forma gratuita para todo el mundo.

El juego Dyetective ha sido creado por Change Dyslexia, adaptado por Samsung, y utiliza las técnicas más avanzadas de inteligencia artificial y minería de datos lingüísticos ■■■

Una de las cuestiones que rodea toda esta carrera es si, ética y filosóficamente, la IA y la salud deben caminar unidas o no





para detectar y dar apoyo a personas con dislexia. La plataforma Dydetective tiene dos partes: una de detección de riesgo de dislexia, que consiste en una prueba, y otra de estimulación para mejorar las dificultades de lectoescritura.

Dydetective integra una investigación patentada, desarrollada por un equipo de investigadores liderado en la Carnegie Mellon University en colaboración con otras universidades (University of Maryland Baltimore County, Universitat de Barcelona, Universitat Pompeu Fabra y University of San Francisco). Para la investigación, se contó con más de

300 voluntarios de diferentes países (España, Colombia, Argentina y Chile) y la colaboración de más de 100 colegios y centros especializados. El resultado es el mayor estudio realizado hasta la fecha sobre dislexia en el mundo de habla hispana.

Para el proyecto, Samsung ha trabajado en colaboración con Luz Rello —licenciada en Lingüística y doctora en Informática, Ashoka Fellow, investigadora en la estadounidense Carnegie Mellon University e investigadora y fundadora de Change Dyslexia—.

En Google relatan que están trabajando con Northwestern Medicine en el desarrollo y ensayo de modelos de IA para capacitar a ecografistas con poca formación y de entornos con bajos recursos para que puedan identificar con seguridad problemas y riesgos potenciales, como los derivados de la posición fetal. Solo en 2017 fallecieron cerca de 295.000 mujeres durante y después del embarazo y en el parto; y en la mayoría de los casos por causas que podrían haberse tratado o evitado con una atención sanitaria adecuada. Por ejemplo, las ecografías forman parte de los protocolos estándar de atención prenatal. Sin embargo, en zonas de bajos recursos el acceso a este tipo de pruebas es deficiente, debido en parte a la falta de personal sanitario con la formación adecuada.

También el diagnóstico por imagen mediante inteligencia artificial ya es una realidad. En colaboración

con organizaciones sanitarias de todo el mundo, Google investiga herramientas de IA innovadoras y potentes para ayudar a los facultativos a realizar diagnósticos. Utilizan diferentes conjuntos de datos, etiquetas de alta calidad y técnicas punteras de aprendizaje profundo, para desarrollar modelos que esperan que en el futuro ayuden a los especialistas médicos a diagnosticar enfermedades.



Audífonos con IA

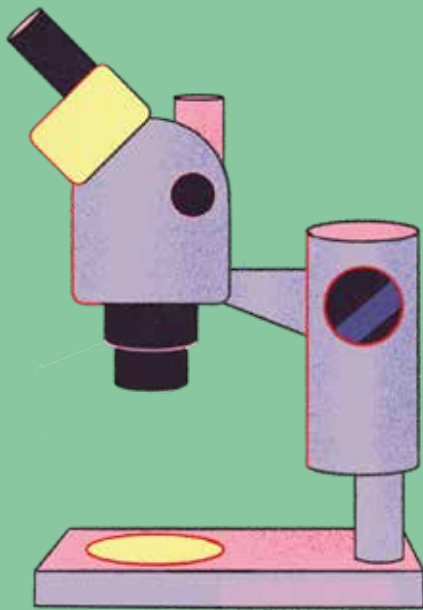
Lo mismo hace Samsung, que adquirió hace unos años en Corea una empresa de equipos médicos en los que la IA es uno de los factores esenciales en la mejora del diagnóstico. Es el caso de equipos como HeartAssist, ClearVision o S-Detect, que son capaces de reducir por siete el tiempo de detección de un problema de salud y de aumentar hasta un 98 % la precisión del diagnóstico.

Y hasta Copenhague me desplacé en octubre para visitar alguna de las experiencias que allí se están realizando en el ámbito de la IA y la salud, en este caso centrada en la intervención de esta tecnología para mejorar los problemas de audición. Si bien es cierto que nuestro cerebro comprende el mundo que nos rodea a través de los cinco sentidos, lo cierto es que en la vida cotidiana dependemos en gran

La IA en medicina y en la salud no es una disciplina aislada, sino que se encuentra en la intersección de la medicina, las ciencias y la ingeniería, pero también con la ética y la filosofía, según la Universidad de Sevilla

medida de la audición para comunicarnos. Su pérdida no solo dificulta la audición, sino que el cerebro también tiene que trabajar más para entender los sonidos. Con inteligencia artificial, y con la tecnología BraingHearing, la compañía danesa Oticon, del grupo Demont, desarrolla sus productos —audífonos e incluso los *chips*, más pequeños que una moneda de un céntimo de euro— que integran entrenando redes neuronales con miles de millones de sonidos de la vida cotidiana, en diferentes acústicas, ambientes, momentos, frecuencias, ...

De esta manera, los audífonos adquieren funciones de proce- ►►►



samiento del sonido cada vez más sofisticadas que van más allá de la amplificación básica, y en los que la inteligencia artificial tiene hoy un papel más que relevante. Por un lado, con la IA se pueden añadir ventajas para el usuario y permitir la personalización del propio audífono, y por otro lado se garantiza el beneficio de este producto en entornos de escucha cotidianos típicos y en los que se gestionan los sonidos que interfieren en la comunicación, denominados genéricamente ruido.

Tanto el mundo académico como la industria han realizado esfuerzos considerables para desarrollar algoritmos cada vez más avanzados que traten el ruido, utilizando normalmente los principios del procesamiento direccional y el postfiltrado, y técnicas para la reducción del ruido en los audífonos modernos.

En primer lugar, se tratan las técnicas clásicas que se utilizan en los audífonos modernos. A continuación, se analiza con el aprendizaje profundo, un subcampo de la inteli-

gencia artificial, la forma radicalmente distinta de resolver el problema del ruido. Y por último, se utilizan los resultados de varios experimentos para mostrar los beneficios de los recientes avances algorítmicos en términos de relación señal-ruido, inteligibilidad del habla, atención selectiva y esfuerzo de escucha. El resultado es un producto cada vez más sofisticado y que ayuda a la persona con dificultades de audición a integrarse en el entorno, no solo porque escucha mejor, sino porque esa mejor escucha le ayuda a comprender, aprender y a relacionarse mejor.

Sería capaz de completar un número entero de *TELOS* hablando de aplicaciones, desarrollos, estudios, investigaciones, proyectos e incluso de nuevas empresas emergentes que están emergiendo para, desde la inteligencia artificial, mejorar la vida de las personas y aumentar las capacidades humanas como señalan muchos de los científicos expertos en este terreno. Pero los interrogantes que genera el uso de la IA siguen siendo también muchos. Además de la privacidad, del que hemos hablado, está el de la seguridad, y el de la ética y los sesgos, que no solo corresponde autocontrolarlos a las propias empresas que desarrollan, sino a las instituciones y a los gobiernos, que tienen que vigilar que lo que se hace esté equilibrado con el respeto a la dignidad de todos, y no solo de unos pocos.

Así lo detalla un estudio certero y extenso de la Universidad de Sevilla en

La privacidad, la seguridad,
la ética y los sesgos son
los principales interrogantes
que genera el uso de la IA

el año 2020¹, y que en sus conclusiones señala que la IA en medicina y en la salud no es una disciplina aislada, sino que se encuentra en la intersección de la medicina, las ciencias y la ingeniería, pero también con disciplinas sociales como la ética y la filosofía. “Ya tiene profundas influencias en muchos aspectos relacionados con la salud y el bienestar humanos a escalas muy diferentes, y las consecuencias del rango de desarrollos en curso pueden afectar a conceptos básicos cubiertos por regulaciones internacionales, desde el juramento hipocrático en medicina hasta los fundamentos del libre albedrío recogidos en la Carta de Derechos Humanos”.

Además, la IA en medicina y la salud es un área muy dinámica, inevitablemente enlazada con la muy rápida evolución de la tecnología, especialmente aquella relacionada con los ordenadores, la fotónica y las diferentes ramas de la ingeniería. Tiene muchas perspectivas positivas junto con aspectos controvertidos y claramente negativos. En algunas aplicaciones, la IA (y las herramientas en las que interviene la IA) incluso evolucionan hacia cuestionar la definición y los límites de la vida. ¿Pero cuál es la visión pública sobre este tipo de investigaciones? “Comienzan a aparecer algunas pautas específicas para *big data* y salud digital, e incluso re-

cientemente se ha formulado una aproximación de no dañar. Algunos de estos temas han sido recientemente definidos como prioridades urgentes para la próxima década por la Organización Mundial de la Salud”, relatan en el informe de la Universidad de Sevilla.

La categorización de las tecnologías como buenas o malas depende en gran medida de muchos factores, en los cuales los valores subjetivos, personales y sociales juegan un papel esencial.

Si bien parece evidente que los criterios objetivos y profesionales deben definirse para un juicio adecuado, el impacto extraordinario de la medicina y la salud en los seres humanos y la sociedad trasciende una evaluación técnica y requiere tener en cuenta también los criterios culturales, éticos y sociales. Para aumentar la complejidad del análisis, la aceptación y la expansión potencial de las aplicaciones de IA se relacionan con equilibrios multifacéticos de efectos de bienestar y aspectos económicos, geográficos, incluso políticos, aunque están esencialmente vinculados con las cualidades personales, intangibles —basadas en la confianza— que configuran la relación de médico y paciente. Al final de todo ¿es la cuestión o no, si la IA puede curarnos, o no?

Bibliografía

- Buch V. H., Ahmed I., Maruthappu M. (2018). “Artificial intelligence in medicine: current trends and future possibilities”. Londres, *British Journal of General Practice* 68: 143–4.
- European Commission, Directorate-General for Research and Innovation, European Group on Ethics in Science and New Technologies (2018). “Statement on Artificial Intelligence, Robotics and ‘Autonomous’ Systems” <https://op.europa.eu/en/publication-detail/-/publication/d1e6e62e-4ce9-11e8-be1d-01aa75ed71a1>. Bruselas (accessed July 25, 2019).
- Gómez-González E., Gómez E., Márquez-Rivas J. et al. (2020). “Inteligencia Artificial en Medicina y Salud: revisión y clasificación de las aplicaciones actuales y del futuro cercano y su impacto ético y social”. Disponible en: https://5a8fe19c-a1f8-41fe-938d-8c3f23f975c9.filesusr.com/ugd/d42f6e_2fc902643d594faa8e80d82d7f43d2b6.pdf
- Greene A., Greene C. C., Greene C. (2019). “Artificial intelligence, chatbots, and the future of medicine”. *The Lancet Oncology: Digital Oncology* | Volume 20, ISSUE 4, P. 481–482.
- Panch T., Pearson-Stuttard J., Greaves F., Atun R. (2019). “Artificial intelligence: opportunities and risks for public health”. *The Lancet Digital Health*. Volume 1, ISSUE 1, e13–e14.

¹ Disponible en: <https://www.us.es/actualidad-de-la-us/efectos-de-la-inteligencia-artificial-en-la-medicina-y-la-salud-de-las-personas>

TELOS

123
NOVIEMBRE 2023 9€

ENTREVISTA: VOCES DE LA IA

Fundación
Telefónica

ARTIFICIAL,
SI, PERO
¿INTELIGENCIA?
Andrés Ortega

QUE DIRÍA
ARISTÓTELES
DE LA IA
Nerea Blanco

ORILLA A,
ORILLA B
Carmen
Torrijos

IA - UNA
HERRAMIENTA
PARA CREAR
Rodrigo Fábrega

MÁQUINAS
INTELIGENTES
Doreen A. Rios

EL CAMINO
HACIA UNA
IA SEGURA
Y SOSTENIBLE
Richard Benjamins

UN PUNTO
DE INFLEXIÓN
EN LA
HUMANIDAD

I
N
T
E
L
I
G
E
N
C
I
A
A
R
T
I
F
I
C
I
A
L

Suscríbete a la revista TELOS

Ahora es más fácil recibir TELOS en tu casa.
Puedes suscribirte desde nuestra web.

Entra en

telos.fundaciontelefonica.com/suscripcion
y rellena el formulario o escribe a
suscripcionestelos@fundaciontelefonica.com

TARIFAS:

Precio por número: Canarias: 9,15 €. Resto de España: 9 €. Suscripción anual: 16 €.
Suscripción anual: América y resto de países: 24 €.

Distribución:

SOIDEM

Algunas librerías en las que se puede
adquirir TELOS

ALBACETE

HERSO

Calle Guardiola, 18

ALICANTE

CILSA LIBROS

Calle Italia, 6

BADAJOS

TUSITALA

Calle Meléndez Valdés, 6

BARCELONA

FREE TIME REVISTAS

Carrer del Comte d'Urgell, 32

LA CENTRAL

Calle Mallorca, 237

Calle Elisabets, 6

BILBAO

LIBRERÍA CÁMARA

Euskalduna Kalea, 6

BURGOS

LIBRERÍA DEL ESPOLÓN

Paseo del Espolón, 30

CÁDIZ

QUORUM LIBROS

Calle Ancha, 27

CIUDAD REAL

LITEC

Avenida Alfonso X el Sabio, 11 Local

CÓRDOBA

NUEVA LUQUE

Calle Jesús María, 6

CORUÑA

BERBIRIANA

Rúa Santiago, 7 - Porta Real

GIJÓN

LIBRERÍA CENTRAL

Calle San Bernardo, 31

GRANADA

LIBRERÍA PICASSO

Calle Obispo Hurtado, 5

HUESCA

ANÓNIMA

Calle Cabestany, 19

LOGROÑO

SANTOS OCHOA

Gran Vía, 55

MADRID

BOUTIQUE DE LA PRENSA

Calle García de Paredes, 34

LA CENTRAL

Calle Postigo de San Martín, 8

LIBRERÍA DIOGENES

Calle Ramón y Cajal, 1

Alcalá de Henares

MÁLAGA

LIBRERÍA LUCES

Avenida Principal 37 - Trinidad Grund, 30

OVIEDO

LA PALMA

Calle Ramón y Cajal, 2

PALMA DE MALLORCA

RATA CORNER

Carrer d'Antoni Marqués, 34

SALAMANCA

LETRAS CORSARIAS

Calle Rector Lucena, 1

SAN SEBASTIÁN

LIBRERÍA LAGUN

Urdaneta Kalea, 3

SEVILLA

CAÓTICA

Calle José Gestoso, 8

TARRAGONA

LA CAPONA

Carrer del Gasòmetre, 43

VALENCIA

LIBRERÍA IZQUIERDO

Gran Vía Fernando el Católico, 12

VALLADOLID

EL ÁRBOL DE LAS LETRAS

Calle Juan Mambrilla, 25

PRIMERA PÁGINA

Calle Corro de Santo Domingo, 15

Urueña-Villa del Libro

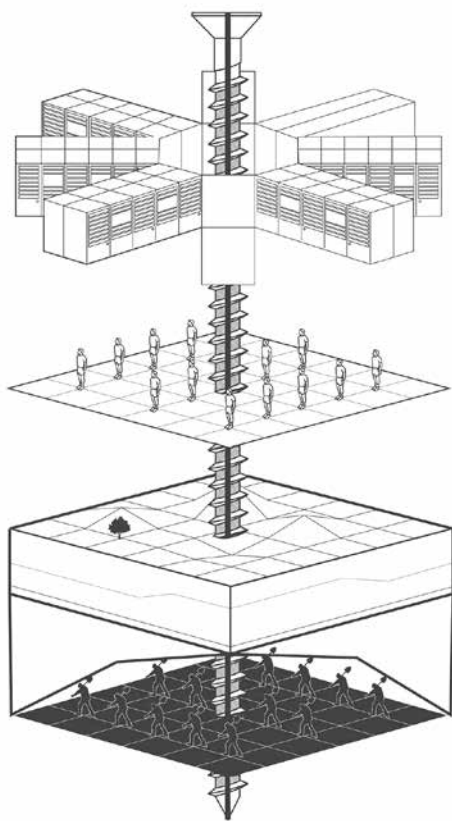
ZARAGOZA

LA PANTERA ROSSA

Calle San Vicente de Paül, 28

Consultar el listado completo en la web.

Biblioteca TELOS



Una sección para ahondar en los temas que propone TELOS. Dedicada a Mario Tascón, amigo y maestro, que nos abrió y amplió los límites del conocimiento con su anhelo constante de saber y su generosidad enorme.

MARIO TASCÓN, EN SUS LIBROS

Un repaso, a modo de homenaje, por la bibliografía de uno de los principales expertos en comunicación de España, fundador de Prodigioso Volcán, recientemente fallecido.

El periodista Mario Tascón (Ponferrada, 16 de diciembre de 1962 - Buenos Aires, 28 de septiembre de 2023) fue el artífice de algunos de los principales cambios de la comunicación en español de las últimas tres décadas. Desde la infografía a finales de los años 80 y 90 en diarios impresos; la revolución del periodismo

digital, del que fue el pionero; o a través de la información corporativa en las últimas décadas, Tascón modernizó la comunicación, la hizo más visual, más clara, más comprensible. Ayudó a medios, personas, instituciones y compañías a hacerse entender. Vinculado estrechamente en esta última década con Fundación Telefónica y con TELOS, Mario Tascón fue el comisario de la actual exposición *Fake News. La fábrica de mentiras*, que se puede visitar hasta enero de 2024 en Espacio Fundación Telefónica. Una muestra que revisa el fenómeno de las noticias falsas —falseadas, diría Tascón— a lo largo de los siglos y plantea una reflexión sobre el impacto que generan actualmente en nuestra sociedad. Desde su consultora de comunicación, Prodigioso Volcán, Mario Tascón comenzó a trabajar para Fundación Telefónica con la conceptualización en 2014 del universo creativo del cocinero Ferrán Adrià, en la exposición *Ferran Adrià. Auditando el proceso creativo*. De ahí, surgió el *Método Sapiens*, para emprender cualquier proyecto de comunicación a través de tres pasos —comprender, crear e innovar—. Tascón, galardonado con el *Premio Internet 2022* de la Asociación de Usuarios de Internet por su trayectoria personal, dejó sus enseñanzas e investigaciones por escrito en varios libros de comunicación, compartidos con otras personas y adelantados a su época para ayudar a comprender el futuro. Sirvan esta páginas para rendir homenaje a una persona curiosa, generosa y bondadosa que abrió y amplió los límites del conocimiento con su anhelo constante de saber.

Mario Tascón, por Raúl Arias

Tras décadas de trabajos en común, el ilustrador y gran amigo de Tascón, Raúl Arias nunca le había retratado hasta ahora. Una ilustración en blanco y negro a la manera de la leyenda de los dibujantes de caricaturas de medios Al Hirschfeld, al que ambos idolotaban.



TWITTERGRAFÍA EL ARTE DE LA NUEVA ESCRITURA

Mario Tascón y Mar Abad (2011)
Madrid, Catarata

Twitter era su red social. Junto a Mar Abad dedica su primer libro a analizar la plataforma que cambió la forma de comunicarse entre las personas y ayudó a propagar información de manera casi instantánea. Reflexiona cómo los usuarios escriben en Twitter, cómo se lanzan mensajes, cómo los políticos y las empresas intentan llegar a los votantes y consumidores, y cómo los ciudadanos la hicieron suya.



CIBERACTIVISMO LAS NUEVAS REVOLUCIONES DE LAS MULTITUDES CONECTADAS

Mario Tascón
y Yolanda Quintana (2012)
Madrid, Catarata

Con el telón de fondo de las revueltas árabes o las filtraciones de WikiLeaks, el libro explica cómo Internet se democratizó con las redes sociales, unas nuevas herramientas donde la simplificación de la capacidad de emitir mensajes y relacionarse con otras personas sin importar la edad, el sexo, la religión o su lugar de origen.



ESCRIBIR EN INTERNET: GUÍA PARA LOS NUEVOS MEDIOS Y LAS REDES SOCIALES

VV. AA. (2012)
Madrid, Galaxia Gutemberg
Mario Tascón, que llegó a ser presidente de Fundéu en 2019, dirigió para esta institución el primer manual práctico de uso del español en los medios digitales y los soportes electrónicos. Pensado para todos los públicos, para aquellos que utilizan la red de manera cotidiana sin ser expertos, pero también para los profesionales.



BIG DATA Y EL INTERNET DE LAS COSAS

Mario Tascón
y Arantza Coullaut (2016)
Madrid, Catarata

Tras publicar su única novela *La biblia bastarda* en 2013 junto a su hermano Fernando, Mario Tascón investigó junto a la periodista Arantza Coullaut sobre un fenómeno que iba a cambiar el mundo digital. El *big data*, que comenzó a permitir a empresas, administraciones e individuos el acceso a una enorme cantidad de datos, que se procesan y transmiten en un volumen y a una velocidad insólitos.



AVIARIO

Mario y Javier Tascón (2016)
Madrid, Catarata

Ilustrado por su hermano, el diseñador Javier Tascón, el *Aviario* es un catálogo de personajes que pululan en redes sociales, que analiza y compara el comportamiento de las personas como si fueran pájaros, y sus mensajes como si fueran trinos. ¿Quiere saber de qué especie son sus hijos, sus amigos, su pareja? ¿Qué tipo de pájaro es usted? Búsquese en esta obra. Los capítulos están disponibles en Podium Podcast: www.podiumpodcast.com/aviario/



EL DERECHO A ENTENDER LA COMUNICACIÓN CLARA, LA MEJOR DEFENSA DE LA CIUDADANÍA

Mario Tascón
y Estrella Montolio (2020)
Madrid, Catarata

Mario Tascón ha sido uno de los principales impulsores de la comunicación clara, el derecho a entender de las personas. Junto a Estrella Montolio condensó la idea de que para ser ciudadanos con plenos derechos el primer paso es entender lo que las instituciones y las empresas nos explican, nos piden... o nos ocultan.



PRESENTACIONES DE IMPACTO CÓMO HACER FÁCIL LO DIFÍCIL: COMUNICACIÓN VISUAL, INFOGRAFÍA Y NARRATIVA

Mario Tascón (2021)
Madrid, Larousse

Ya no basta con tan solo comunicar. Para hacerlo bien es necesario cautivar a quien nos escucha usando todos los recursos a nuestro alcance: creativos, visuales y narrativos. Una demostración de que no lo es tanto si sigues los pasos correctos. Una obra ilustrada para sacar partido a diferentes elementos gráficos a la hora de comunicar ideas y hacer presentaciones.



CAMBIA TU VIDA CON UN CUADERNO Y UN BOLI

Mario Tascón (2022)
Madrid, Prodigioso Volcán

En los últimos años, Mario Tascón y Prodigioso Volcán organizaban talleres sobre *bullet journal*, un sistema para organizar mejor el día a día a través de un cuaderno y un bolígrafo. Gestionar tareas y recordatorios, mantener la concentración o aprovechar el tiempo libre se convierten en auténticos retos cotidianos. El libro ayuda a recuperar el control de la vida diaria y da descanso a la mente y pone en valor el tiempo.



Superinteligencia: Caminos, peligros, estrategias

¿ES POSIBLE CREAR UNA INTELIGENCIA ARTIFICIAL INICIAL?

El cerebro humano posee habilidades únicas que distinguen a nuestra especie de otras criaturas. Mientras que otros animales pueden tener músculos más fuertes o garras afiladas, nosotros destacamos por nuestros cerebros más inteligentes. Si en el futuro, los cerebros artificiales lograran superar la inteligencia general humana, esta nueva forma de inteligencia se convertiría en una fuerza extremadamente poderosa. El futuro de nuestra especie dependería en gran medida de las acciones de la superinteligencia artificial.

La pregunta fundamental es si es posible crear una inteligencia artificial inicial o establecer las condiciones iniciales que nos permitan sobrevivir y controlar una explosión de inteligencia. Después de una fascinante travesía que nos lleva a cuestionar los límites de la condición humana y el futuro de la inteligencia, descubrimos que el trabajo de Bostrom representa una reevaluación fundamental de uno de los temas más críticos de nuestra época.

Nick Bostrom (2016). Zaragoza, Teell Editorial. 352 páginas. ISBN: 978-8416511051



FILOSOFÍA POLÍTICA DE LA INTELIGENCIA ARTIFICIAL: UNA INTRODUCCIÓN

Mark Coeckelbergh (2023)
Madrid, Ediciones Cátedra.
224 páginas.
ISBN: 9788437645643

A medida que las tecnologías de inteligencia artificial continúan su desarrollo, problemas cruciales en nuestra sociedad, como el racismo, el cambio climático y la calidad de la democracia, adquieren una urgencia creciente. La pregunta que se plantea es: ¿cómo avanzar en el debate político sobre la IA más allá de las meras advertencias y las acusaciones simplistas? Este libro ofrece una introducción accesible desde la perspectiva de la filosofía política a los numerosos desafíos que surgen de la incorporación de las IA en nuestras sociedades e industrias. Estos desafíos van desde la discriminación en la toma de decisiones judiciales hasta las implicaciones para la democracia de una vigilancia ubicua y constante. Se trata de un ensayo actualizado que será de interés tanto para estudiantes e investigadores en filosofía de la tecnología y filosofía política como para especialistas en innovación, reguladores y cualquier persona interesada en comprender el impacto sociocultural de la IA.



CÓMO CREAR UNA MENTE

Ray Kurzweil (2023)
Berlín, Lola Books.
324 páginas.
ISBN: 978-3944203058

En el horizonte del año 2029, Kurzweil vislumbra un punto crítico en el desarrollo de la IA: el momento en que las máquinas no podrán distinguirse de la mente humana. Este libro traza la curva de crecimiento de la tecnología hasta llegar al momento en que la inteligencia de los ordenadores alcance la paridad con la humana. Kurzweil argumenta que dado que la creación de máquinas más potentes es una capacidad inherente a los seres humanos estas máquinas heredarán esa capacidad. Esto dará lugar a generaciones posteriores de máquinas cada vez más poderosas que acompañarán a la humanidad hacia la singularidad tecnológica. El autor explica detalladamente el camino que la humanidad está recorriendo para alcanzar este objetivo y cómo se está llevando a cabo esta transformación. Según Kurzweil, el futuro inteligente de las máquinas se entrelaza con el de los seres humanos porque seremos capaces de incorporar esta inteligencia en nuestros cuerpos.



Tres recomendaciones de...

LECTURAS SOBRE INTELIGENCIA ARTIFICIAL

JORGE CARRIÓN

El escritor y profesor ha trabajado intensamente la inteligencia artificial durante los últimos cinco años. En ficción con la novela *Membrana*, *Galaxia Gutenberg*, que imagina el Museo del Siglo XXI, en formato pódcast con *Solaris*, *ensayos sonoros* (premio Ondas), en exposiciones como *Todos los museos son novelas de ciencia ficción* y también a través de la co-escritura con GPT-2 y GPT-3 (lo que dio lugar a su libro de ensayo y textos creativos *Los campos electromagnéticos, Teorías y prácticas de la escritura artificial*, editado por Caja Negra y diseñado junto a Taller Estampa). Estos son los tres libros sobre inteligencia artificial que Jorge Carrión recomienda a la #ComunidadTELOS



1

IA: INTELIGENCIA ARTIFICIAL

VV. AA., Centro de Cultura Contemporánea de Barcelona

El catálogo de una exposición que nos cuenta el fenómeno desde sus orígenes en Ramon Llull, las matemáticas clásicas o Ada Lovelace hasta sus desarrollos actuales gracias a la supercomputación o las redes neuronales de aprendizaje profundo, con recursos pedagógicos y obras de arte contemporáneo. Con textos de, entre otros, Marta Peirano y Jordi Torres.



2

ATLAS DE LA IA

Kate Crawford, Ned Ediciones

Tal vez el libro más importante que se ha publicado sobre el tema. Nos explica no solo la tecnología en sí, sino también el marco político y extractivista en el que se inscribe su vertiginoso desarrollo. Su perversa dimensión mineral, material, humana.



3

ARTIFICIAL. LA NUEVA INTELIGENCIA Y EL CONTORNO DE LO HUMANO

Mariano Sigman y Santiago Bilinkin, Debate

Buenas reflexiones, detalles históricos no demasiado conocidos y, sobre todo, un enfoque neurocientífico aportado por Sigman.



La edición 2023 del evento abordó el impacto de la inteligencia artificial generativa en la educación, el empleo y el emprendimiento

EL IMPACTO DE LA IA, EN ENLIGHTED

Los días 18 y 19 de octubre se celebró en Madrid la 6ª edición de enlightED, la conferencia mundial de educación, tecnología e innovación, impulsada por Fundación Telefónica, IE University, South Summit y la Fundación "la Caixa". Desde su nacimiento, enlightED se ha convertido en un referente para la comunidad educativa en los ámbitos de la innovación y la tecnología.

La edición de 2023 ha tenido como eje repensar la educación en la era de la inteligencia artificial, y ha puesto el foco en los retos, oportunidades y soluciones que plantea la IA generativa, el uso y sus aplicaciones para los docentes y estudiantes. Igualmente, ha invitado a reflexionar sobre el respeto a la diversidad, así como a la regulación y los límites éticos, y el impacto en el futuro del empleo, entre otras cuestiones.

Entre los numerosos expertos del ámbito académico y empresarial que formaron parte destacaron los nombres de **Toju Duke (1)**, fundadora de Diverse AI; **Daniel Susskind (3)**, catedrático de Economía en el King's College de Londres e investigador asociado en el Instituto de Ética de la IA de la Universidad de Oxford; **Carl-Benedikt Frey (2)**, profesor asociado Dieter Schwarz de IA y trabajo en el Oxford Internet Institute y miembro del Mansfield College de la Universidad de Oxford; **Carlos Santana (4)**, divulgador, *youtuber* y profesor en el campo de la IA y **Priya Lakhani (5)**, fundadora y CEO de Century Tech.

En su intervención, Toju Duke subrayó la necesidad de conocer los riesgos y limitaciones de la inteligencia artificial para garantizar que somos capaces de maximizar sus beneficios y de minimizar los posibles daños que pueda traer a la sociedad. Por su parte, Daniel Susskind

alertó sobre la destrucción de puestos de trabajo causada por la innovación, y afirmó que las respuestas a estos retos están directamente relacionadas con la educación y la formación en competencias.

El experto en el impacto de la digitalización en el mercado laboral Carl-Benedikt Frey destacó en su ponencia que debemos utilizar nuestro sistema educativo y de bienestar para ayudar a la población a emplear estas nuevas tecnologías digitales, porque si no es así, no vamos a ser capaces de adaptarnos a esta transformación. Edo Segal, por su parte, aportó una visión optimista de este cambio postulando que los algoritmos inteligentes van a permitir personalizar la educación y adaptar la enseñanza al ritmo de aprendizaje de cada alumno. Finalmente, para Priya Lakhani los riesgos que percibimos nos impiden ver con claridad todos los beneficios y oportunidades que ofrece la inteligencia artificial en la educación y, por ello, debemos aprender a utilizarla conociendo bien los peligros que conlleva para poder evitarlos y sacar el máximo provecho de ella. Se trata de una pequeña muestra del aluvión de ideas y debates que fluyeron durante las dos jornadas de enlightED 2023.

El acto de clausura del evento estuvo presidido por Pilar Alegre, ministra de Educación y Formación Profesional en funciones, José Luis Martínez-Almeida, alcalde de Madrid y José María Álvarez-Pallete, presidente de Telefónica y de Fundación Telefónica, quien, a modo de reflexión de cierre del evento abogó por hacer inclusiva la tecnología educativa: *"Personalicemos la educación a través de la inteligencia artificial y aprendamos a hacer un uso de la tecnología para no dejar a nadie atrás."*

Disponible en: <https://www.enlighted.education/>



EL MITO DE LA INTELIGENCIA ARTIFICIAL: POR QUÉ LAS MÁQUINAS NO PUEDEN PENSAR COMO NOSOTROS LO HACEMOS

Erik J. Larson (2022)
Barcelona, Shackleton Books. 405 páginas.
ISBN: 978-8413611686

Los visionarios del futuro a menudo aseveran que la inteligencia artificial pronto superará las capacidades de las mentes humanas más brillantes, anunciando un sombrío destino en el que las máquinas superinteligentes dominarán. Sin embargo, la verdad es que ni siquiera estamos seguros de estar en el camino correcto hacia el desarrollo de máquinas verdaderamente inteligentes.

A lo largo de la historia, los entusiastas de la IA han cometido el error de equipararla directamente con la inteligencia humana. La exageración en torno a la IA no solo es un error científico, sino que también obstaculiza el progreso real. La verdadera innovación se fomenta cuando exploramos lo desconocido en lugar de exagerar las virtudes de tecnologías existentes. Debemos reconocer plenamente la singularidad de la inteligencia humana si deseamos avanzar de manera significativa.



TECNOHUMANISMO

Pablo Sanguinetti (2023)
Madrid, La Huerta Grande. 119 páginas.
ISBN: 9788418657375

Este libro reúne siete ensayos que, aunque independientes, comparten una misma búsqueda: la de establecer conexiones entre la tecnología y las humanidades, centrándose especialmente en el avance exponencial de la IA y los dilemas filosóficos que plantea. La IA está traspasando barreras y adentrándose en campos propios de la humanidad, como la creatividad, el lenguaje y la imaginación. Este avance plantea interrogantes profundos acerca de nuestra propia especie. La tecnología se ha convertido en una cuestión ética, filosófica y estética, lo que hace que sea esencial que los humanistas se involucren en el debate. Los ensayos abordan la posibilidad y las condiciones de un nuevo humanismo tecnológico desde diversas perspectivas, como la importancia de la belleza en el diseño de la IA, la superación de prejuicios entre la tecnología y las humanidades y la necesidad de crear un relato tecnológico en sintonía con los intereses humanos, entre otros.



Máquinas como yo

LA CAPACIDAD DE LAS MÁQUINAS PARA COMPRENDER Y JUZGAR

Ian McEwan adentra al lector en una Londres distópica de los años ochenta, donde la historia ha seguido un rumbo alternativo. En este contexto, el científico Alan Turing, lejos de su trágico destino en la realidad, sigue vivo y se dedica al desarrollo de la IA. Ha logrado un hito al crear a Adán y Eva. La historia se centra en Charlie, quien adquiere uno de los Adanes para compañía, y su amante, Miranda, que guarda un oscuro secreto.

La novela explora la complejidad moral de las decisiones humanas y la capacidad de las máquinas para comprender y juzgar. Adán, siendo un ser sintético casi perfecto pero carente de matices morales, descubre el secreto y desencadena una tensión creciente entre los personajes.

McEwan utiliza la ciencia ficción para plantear preguntas fundamentales sobre la humanidad, los límites éticos de la IA y la moralidad de las decisiones humanas.

Ian McEwan (2019). Barcelona, Anagrama. 360 páginas. ISBN: 9788433980465



EL ARTE DE SER HUMANOS

Rob Riemen (2023)
Barcelona, Taurus. 248 páginas.
ISBN: 978-8430626373

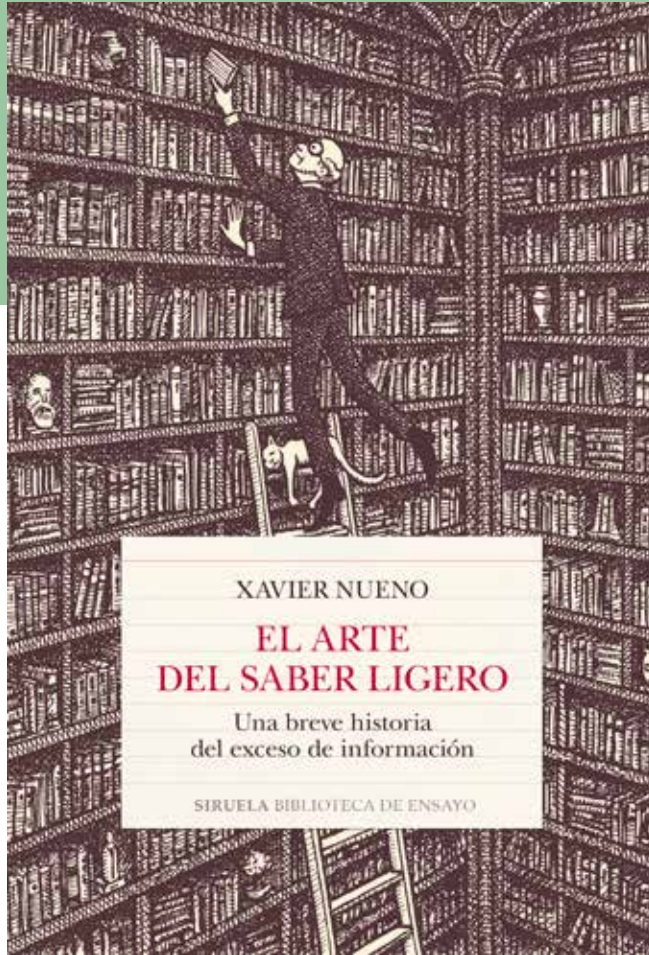
Nuestra auténtica identidad reside en valores universales que trascienden las diferencias superficiales entre individuos, como género, raza, religión, origen o apariencia física. Estos valores esenciales incluyen la capacidad de vivir con honestidad, de buscar la justicia, amar, crear y apreciar la belleza. Lo que nos diferencia y a menudo nos divide es secundario en comparación con estos principios universales. Riemen explora las cuestiones fundamentales planteadas por Sócrates, como "¿Cómo debemos vivir?" y "¿Qué constituye una sociedad justa?" a través de cuatro ejes temáticos. En primer lugar, enfatiza la importancia de la memoria para reconocer y confrontar las fuerzas malignas. Luego, resalta los peligros de la estupidez y la mentira. Posteriormente, aboga por el papel del intelectual en la esfera pública. Por último, se centra en la importancia del amor profundo. El libro defiende el humanismo y la esperanza de que, si los seres humanos cambian en un nivel fundamental, podrán romper con la repetición de la historia.



UN MUNDO CON IA

Fundación Telefónica
Movistar | México

En este episodio se explora uno de los temas más candentes en la actualidad: la inteligencia artificial. Aunque suele estar asociada con la informática, su alcance es mucho más amplio. La IA va más allá, convirtiéndose en un asistente en la creación de contenido y diseño, lo que genera un interés generalizado. Conforme la IA se integra en la vida cotidiana, es esencial que la sociedad se prepare para aprovechar sus beneficios y abordar los desafíos planteados. Esto incluye promover la alfabetización en IA para comprender su funcionamiento básico y mantener una mentalidad de adaptación continua, además de considerar aspectos éticos como la equidad en el acceso y el uso responsable de la IA. Estos son temas fundamentales en el presente y futuro de la inteligencia artificial, como discutieron en este episodio Carmen Torrijos y Pablo Escobedo de Prodigioso Volcán junto a Alejandro Platas, docente de IA en la Universidad Anáhuac.



El arte del saber ligero

TEME A LOS PEQUEÑOS LIBROS PORTÁTILES, PUÉDEN HACER LA REVOLUCIÓN

La sensación de sobrecarga informativa no es un fenómeno nuevo. Mucho antes de la era digital, nuestros predecesores ya habían experimentado con inquietud los efectos de la acumulación infinita de escritos. Sin embargo, junto a la tradición que buscaba siempre aumentar las colecciones de bibliotecas, existía una corriente que advertía sobre los peligros de ser sepultados por el peso del pasado.

Las vanguardias y los antimodernos forjan un pacto contrario al de Fausto: en lugar de entregar sus almas a cambio de un conocimiento ilimitado, exploran la idea de poner límites al deseo de saberlo todo. Estos lectores crean bibliotecas portátiles y otras formas abreviadas, ligeras y móviles del saber, con el propósito de sacar el conocimiento de las estanterías y practicar un auténtico humanismo transformador. Esta reducción nos recuerda que la barbarie puede surgir tanto por la escasez como por la sobreabundancia de libros.

Xavier Nuño (2023). Madrid, Ediciones Siruela. 252 páginas. ISBN:978-84-19744-47-0



ARTIFICIAL. LA NUEVA INTELIGENCIA Y EL CONTORNO DE LO HUMANO

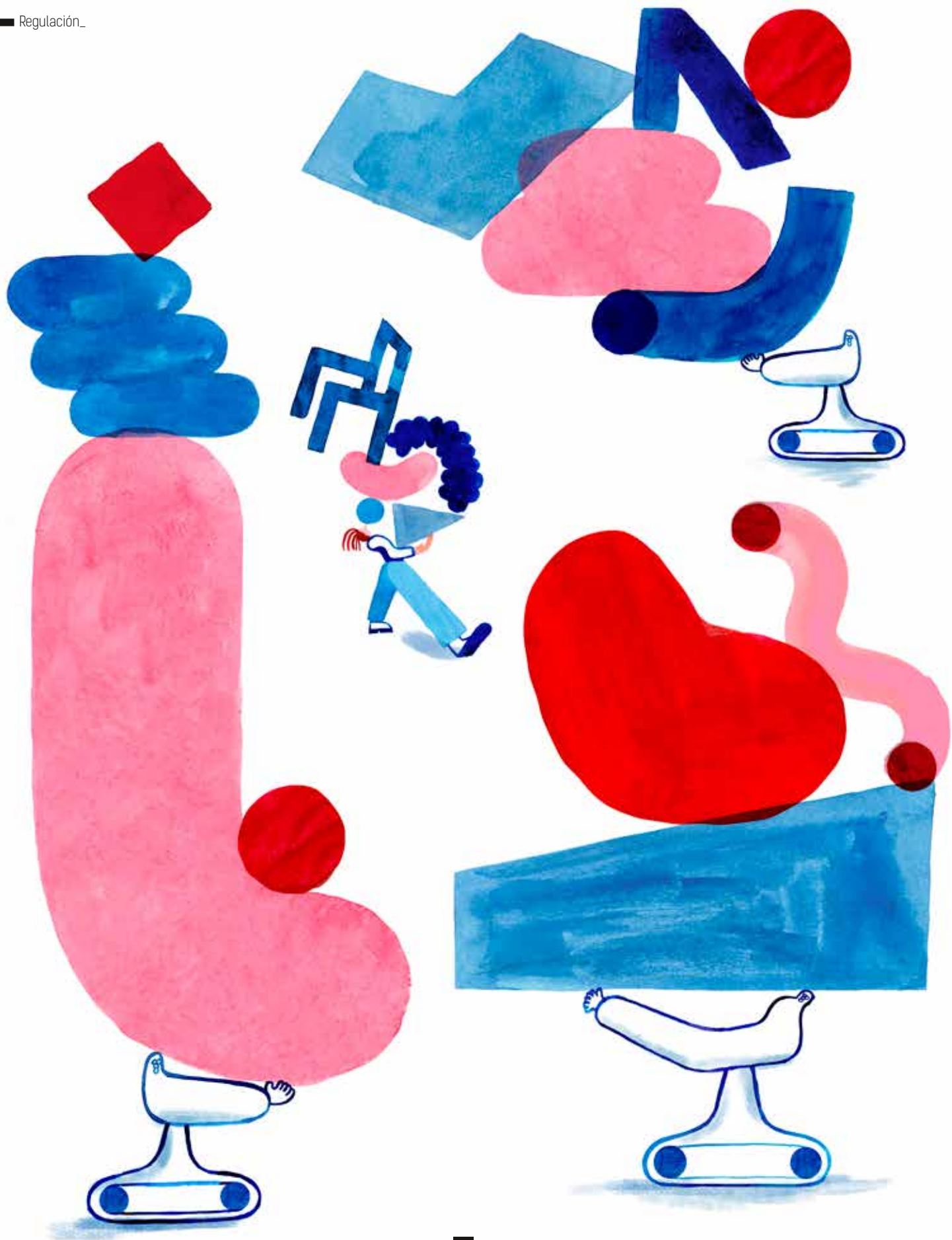
Mariano Sigman y
Santiago Bilinkis (2023)
Madrid, Debate
232 páginas
ISBN: 978-8419642677

Mariano Sigman, uno de los neurocientíficos más destacados del mundo, y Santiago Bilinkis, emprendedor y tecnólogo, nos ofrecen en su obra un enfoque sin temor hacia la inteligencia artificial, proporcionando ideas, herramientas y planteando preguntas fundamentales. Este libro no busca predecir el futuro ni realizar profecías, sino que presenta las reflexiones de dos autores que creen firmemente que el conocimiento es la clave para navegar la imparable ola de la inteligencia artificial. En este libro, Sigman y Bilinkis exploran el origen, las aplicaciones y los posibles riesgos de esta tecnología. Plantean cuestiones cruciales como si la IA será una fuente de oportunidades o un potencial desafío, y cómo podemos acercarnos al escenario más beneficioso. Aunque señalan la importancia de la cautela y la responsabilidad, evitan caer en el pesimismo y nos alientan a considerar que este desafío puede sacar lo mejor de nosotros.

Regulación

La regulación de la IA en el ámbito de las relaciones laborales y de la propiedad intelectual resultan esenciales para proteger derechos sin cercenar la innovación responsable.







JOSÉ VARELA FERRÍO

Responsable de Digitalización en el Trabajo de UGT y experto en Derecho en IA

Palabras clave:
IA, empleo, trabajo, sesgo, discriminación, empleabilidad.



LA IA TENDRÁ UN PROFUNDO IMPACTO EN EL EMPLEO Y NO TODO SERÁ POSITIVO

Quo vadis machina?

La inteligencia artificial (IA) está cada vez más generalizada, pero existe muy poca información sobre cómo afectará al mundo del trabajo. Empero, cada día se detectan más riesgos de desempleo tecnológico, trato discriminatorio o precarización asociada. El V Acuerdo para el Empleo y la Negociación Colectiva (AENC) firmado por los agentes sociales tratará de regularlo desde el consenso.

AI will have a deep impact on employment and not all of it will be positive

QUO VADIS MACHINA?

Artificial intelligence (AI) is becoming increasingly widespread, but there is very little information on how it will affect the world of labor. However, the risks of technological unemployment, discriminatory treatment or associated precariousness are becoming increasingly apparent. The V Agreement for Employment and Collective Bargaining Agreement signed by the social agents will attempt to regulate this on the basis of consensus.

Keywords: AI, employment, work, bias, discrimination, employability.

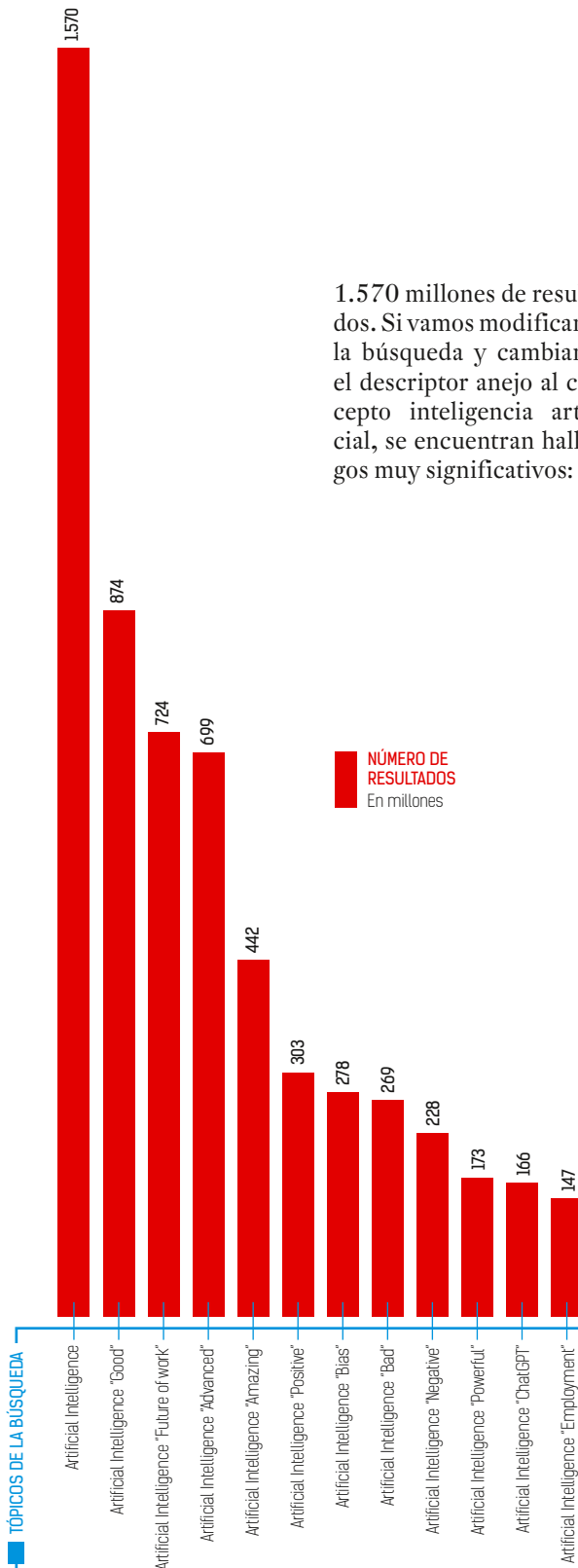
Es uno de los conceptos de moda: inteligencia artificial (IA). En un pequeño lapso de tiempo, ha pasado de estar circunscrita a ámbitos científicos y de ciencia ficción a formar parte de la agenda de jefes de Estado, de los CEO y de un sinfín de expertos y comentaristas, elevándola a la categoría de asunto de cotidianidad mediática.

Sin embargo, en muy pocas ocasiones el debate se centra o incluye al

factor trabajo. Y en las escasas ocasiones que se aborda el impacto de la IA en el empleo y en las relaciones laborales, la conversación tiende a focalizarse en aspectos exclusivamente positivos, mezclando cuestiones que llevan a confusión o equívoco. Así, vemos cómo se superponen en un mismo mensaje verdades indiscutibles del uso de la IA, como la eliminación de tareas improductivas o la

creación de nuevas profesiones, con otras mucho más controvertidas como la creación neta de empleo a corto plazo para añadir otras afirmaciones directamente falsas, como la existencia de la infalibilidad algorítmica.

Estas tendencias se reflejan de forma muy intuitiva en los resultados del buscador de Google. Si se busca el concepto *Artificial Intelligence*, se encuentra la astronómica cifra de >>>



1.570 millones de resultados. Si vamos modificando la búsqueda y cambiando el descriptor anejo al concepto inteligencia artificial, se encuentran hallazgos muy significativos:

La primera conclusión es que la IA se asocia con algo “bueno” (*good*) tres veces más que con algo “malo” (*bad*). Además, los adjetivos que más entradas acumulan son aquellos que le aportan un aspecto francamente positivo (“avanzada”, “maravillosa”, “poderosa”: “*advanced*”, “*amazing*”, “*powerful*”), muy por encima de los que le asocian parcialidad o sesgos (*bias*). En suma, las entradas que transmiten una idea positiva de la IA son un 25 por ciento superiores a las negativas.

Luego está el apartado del trabajo y del empleo. La vinculación de la IA y futuro del empleo está entre las tendencias más populares, pero de un rápido vistazo se comprueba que bajo esa búsqueda se aglutina un confuso compendio de nuevas profesiones, oportunidades de empleo y ofertas formativas (normalmente de carácter privado e integradas sin solución de continuidad en el mismo texto). Con todo, lo más llamativo es que, si relacionamos los conceptos *IA-empleo*, los resultados no lleguen ni al 10 por ciento del total de entradas.

Con todas las cautelas metodológicas que debe tener una aproximación de estas características, se pueden alcanzar dos con-

clusiones: en general, la IA es percibida como algo positivo, muy por encima de representar un riesgo, postergando las entradas negativas en favor de una descripción de la IA como una *tecnología superlativa*. Y dentro de toda la información que transcurre por la red de redes, la afectación de aquella al empleo no es un aspecto que pre-ocupe/destaque en demasía, salvo si concurre un contexto en donde el objetivo es publicitar la formación sobre IA.

Percepción

Sin embargo, lo que nos transmite un buscador de Internet, no tiene por qué coincidir con la opinión pública. De hecho, la ciudadanía demuestra un sentir muy diferente. Así, la última *Encuesta de Percepción Social de la Ciencia y la Tecnología en España*, publicada por la Fundación Española para la Ciencia y la Tecnología (FECYT), detectaba que un 73 por ciento de la ciudadanía pensaba que la introducción de robots en los puestos de trabajo contribuye de forma importante a aumentar el riesgo de desempleo, a la par que entendía que no

existía relación entre la IA y el incremento en la competitividad de las empresas (54,5 por ciento). Además, resultaba significativo que se registre un porcentaje mayor de personas que considera que estos avances tecnológicos entrañan más riesgos que beneficios (45,4 por ciento frente al 39,5 por ciento).

La Fundación COTEC publica periódicamente un análisis similar (*Encuesta de percepción social de la innovación en España*). En la de 2023 volvemos a encontrar con interesantes propensiones: la mitad de los encuestados se inclina por secundar que la innovación tecnológica aumenta la desigualdad social (frente a un 29 por ciento que no corrobora este extremo) y 6 de 10 personas consideran que la mayoría de los puestos de trabajo del presente serán realizados por máquinas en los próximos años.

Conforme a estos datos, parece más que evidente que la ciudadanía española ve con preocupación el impacto de las nuevas tecnologías sobre el empleo. Recordemos que ambos sondeos se realizaron antes de la irrupción mediática —y universalmente accesible— de ChatGPT. Es fácil sospechar, sin mucho

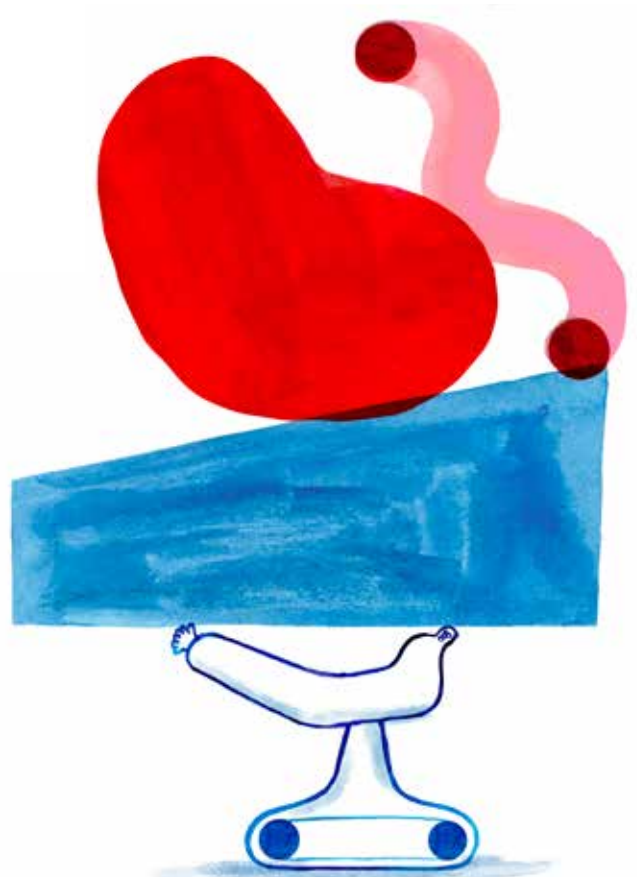
margen de error, que en las siguientes ediciones estos temores se agudizarán.

Empero, este generalizado recelo a la masificación de la IA en los centros de trabajo parece no alcanzar ni a legisladores ni a reguladores, como si se diese más importancia a los resultados del algoritmo de Google que a los sentimientos y opiniones de las personas.

La mayor prueba de esta disonancia la tenemos en el primer trabajo del Grupo de Alto Nivel sobre Inteligencia Artificial, que apadrinaba la Comisión Europea, al objeto de dar un punto de salida sobre cómo regular esta tecnología. En su documento *Directrices éticas para una IA fiable*, de 55 páginas, apenas aparece una única reseña al mundo del trabajo, por su “impacto social”. Ignorar el factor trabajo en la primera piedra de una futura regulación sobre IA es despreciar la realidad de millones de europeos y europeas, donde tener empleo significa tener sustento y en donde invertimos un tercio de nuestras vidas.

Como resultado de la acción de muchos sindicatos y asociaciones, la *AI Act*, la futura Ley de inteligencia artificial de la UE, parece haber corregido este ex-

Que la IA está ya afectando de forma directa al entorno laboral es un hecho incontrovertible



La negociación colectiva debe desempeñar un papel fundamental estableciendo criterios que garanticen un uso adecuado de la IA

tremo al considerar la IA aplicada a las relaciones laborales como una actividad de *alto riesgo*. Sin duda, llegará con retraso. No sabemos exactamente cuándo, puesto que se desconoce su fecha de entrada en vigor.

Y dicha Ley de IA llegará tarde porque la IA desregulada ya está imbuida en nuestro tejido productivo, especialmente en España. El INE confirma que, en 2022, un 21 por ciento de las empresas de menos de 10 empleados usó IA para automatizar flujos de trabajo. En el caso de las empresas de más de 10 personas trabajadoras este porcentaje se dobla, alcanzando el 46 por ciento.

Dos estudios europeos de Eurofound reafirman esta capilaridad. En *The platformisation of work*, el organismo europeo revelaba que un 34,7 por ciento de las personas trabajado-

ras están gobernados por alguna suerte de algoritmo (en Alemania, ni un 19 por ciento de las plantillas se encontraba en esta situación). En *Employee monitoring and surveillance: The challenges of digitalisation* se señalaba que hasta un 40 por ciento de las empresas españolas estarían usando algoritmos para la monitorización de la actividad laboral y para evaluar el desempeño de las personas trabajadoras, colocando a España en el dudoso honor de liderar esta estadística en toda Europa (el único *ranking* digital donde la economía española está en cabeza).

Que la IA está ya afectando de forma directa al entorno laboral es un hecho incontrovertible, y por tres vías principales.

La primera, relacionada con el volumen de empleo. Veamos dos ejemplos muy significativos: IBM, uno de los mayores suministradores del mundo de IA, ha afirmado oficialmente que las tareas de 8.000 puestos de trabajo que iba a realizar con contrataciones humanas, los va a suplir con esta tecnología. British Telecom, uno de los operadores de telecomunicaciones más importantes del mundo, ha presentado un plan de despidos que afectará a 55.000 personas. Las tareas que les correspondería realizar a

10.000 de estas personas se harán mediante IA. Estamos solo ante la punta del iceberg: Goldman Sachs afirmó que la inteligencia artificial generativa podría dejar sin empleo a 300 millones de trabajadores entre Estados Unidos y Europa, y el Foro Económico Mundial predice una reducción de 14 millones de puestos de trabajo en los próximos cuatro años. Demasiadas señales para ser consideradas como algo anecdótico.

Relaciones laborales

La segunda vía es la relacionada con la negociación colectiva, a la que ya está condicionando de forma muy sustantiva. En Hollywood, la respuesta de las patronales a la huelga de guionistas es “no seáis muy exigentes, porque si no encargaremos los guiones a la IA”. No es el primer caso; hace unos años, el CEO de McDonald’s arguyó algo semejante blandiendo la amenaza de brazos robóticos soportados por IA (el primer establecimiento de la cadena sin humanos se abrió en finales de 2022 en Texas).

La tercera vía, más soterrada pero cada vez más emergente, versa sobre si las decisiones de la IA en las relaciones laborales están



libres de decisiones injustas, sesgadas o directamente discriminatorias. Varios estudios europeos señalan que la IA y la algoritmia tienen un potencial efecto negativo sobre nueve derechos humanos fundamentales (como el derecho a la dignidad humana, a la integridad personal, al respeto a la intimidad o a la no discriminación).

No obstante, y según el *IBM Global AI Adoption Index 2022*, el 74 por ciento de las empresas no reducen el sesgo no deseado y un 60 por ciento no desarrollan políticas éticas relacionadas con la inteligencia artificial. Si a esta evidencia le unimos que, al menos, un 46 por ciento de las empresas españolas ya usan algún tipo de IA laboral sin que conste supervisión humana ni evaluación de impacto previa, ni protección de datos desde el diseño, ni auditorías *ex ante* o *ex post*, deberíamos estar poniéndonos en lo peor.

Por ello cobra especial importancia en el V Acuerdo para el Empleo y la Negociación Colectiva, que dedica un capítulo entero a cómo los futuros convenios deben regular la IA, una estipulación pionera en Europa.

Los cuatro firmantes (UGT, CC. OO., CEOE y CEPYME) partimos de un consenso: la IA tendrá un impacto significativo en el mundo laboral y su uso in-

adecuado podría desembo- car en unas relaciones labo- rales discriminatorias o con tratamientos sesgados.

Partiendo de dicha base, consagramos dos principios fundamentales: garantizar el control humano de la IA —aportando así seguridad y criterio al funcionamiento de esta tecnología— y aportar transparencia al funcionamiento de los algoritmos a través del ejercicio del derecho a la información. Dicho derecho ya se articula en el Estatuto de los Trabajadores en su artículo 64.4.d, que obliga a facilitar información “transparente y entendible sobre los procesos que se basen en ella [la IA] en los procedimientos de Recur-

sos Humanos (contratación, evaluación, promoción y despido)”, garantizándose así “que no existen prejuicios ni discriminaciones”.

Con estas premisas, ambas partes declaramos que la “negociación colectiva debe desempeñar un papel fundamental estableciendo criterios que garanticen un uso adecuado de la IA y sobre el desarrollo del deber de información periódica a la representación de los trabajadores”, llegando a ampliar este llamamiento a los sistemas de IA en las Administraciones públicas, a las que instamos a seguir igualmente el principio de control humano y transparencia.

El reto que tenemos por delante es mayúsculo. Regular con buen tino una tecnología tan compleja y vertiginosa como es la inteligencia artificial requerirá de un gran esfuerzo colectivo, exigiendo, a todas las partes, preparación, pericia y audacia. Pero el bien superior perseguido vale la pena: la IA tiene que contribuir a crear entornos laborales cimentados en la igualdad de oportunidades, la dignidad y la empleabilidad. Lo contrario, no es admisible.

Bibliografía

- Dirección General de Redes de Comunicación, Contenido y Tecnologías (Comisión Europea) Grupa ekspertów wysokiego szczebla ds. sztucznej inteligencji (2019): *Directrices éticas para una IA fiable*. Disponible en: <https://data.europa.eu/doi/10.2759/14078>
- Eurofound (2020). *Employee monitoring and surveillance: The challenges of digitalisation*. Luxemburgo, Publications Office of the European Union.
- Fernández-Macias, E., Urzi Brancati, M. C., Wright, S. and Pesole, A. (2023): *The platformisation of work*. Luxemburgo, Publications Office of the European Union.
- Fundación COTEC (2022): *Percepción social de la innovación en España*. Disponible en: <https://cotec.es/observacion/encuesta-de-percepcion-social-de-la-innovacion/6a81446c-6656-94e3-5d74-f8328ce1ca0b>
- Fundación Española para la Ciencia y la Tecnología (FECYT). (2022): *Encuesta de Percepción Social de la Ciencia y la Tecnología en España (EPSC)*. Disponible en: <https://www.fecyt.es/es/publicacion/percepcion-social-de-la-ciencia-y-la-tecnologia-en-espana-2022>
- Goldman Sachs (2023): *Generative AI could raise global GDP by 7%*. Disponible en: <https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html>
- IBM: *IBM Global AI Adoption Index 2022*. Disponible en: <https://www.ibm.com/watson/resources/ai-adoption>
- INE (2022): *Encuesta sobre el uso de TIC y comercio electrónico en las empresas*. Disponible en: https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176743&menu=ultiDatos&idp=1254735576799
- V Acuerdo para el Empleo y la Negociación Colectiva (V AENC). Disponible en: https://www.ugt.es/sites/default/files/v_aenc_firma_10_05_2023_def.pdf
- World Economic Forum: *The Future of Jobs Report 2023*. Disponible en: <https://www.weforum.org/reports/the-future-of-jobs-report-2023/>

EL TEMOR A LA SUPLANTACIÓN DE UNA CONDICIÓN GENUINAMENTE HUMANA SE EXTIENDE, PERO LAS RESPUESTAS SIGUEN ESTANDO DEL LADO DE LOS CREADORES

¿El fin de la creatividad humana?

La irrupción de la inteligencia artificial (IA) en nuestro día a día nos plantea dudas sobre los límites de la capacidad humana, nuestra identidad y nuestro futuro. Este campo se presenta especialmente sensible para los creadores. ¿Es la IA una herramienta para la creación contemporánea o un competidor de los artistas?

La realidad postpandemia y el cambio de década han traído consigo una eclosión sin precedentes en el uso de las nuevas tecnologías y su implantación en sectores tradicionalmente reservados a la intervención humana. Hablar hoy de *blockchain*, de metaverso, de NFT y de inteligencia artificial es algo cotidiano. Con la progresiva penetración de estas tecnologías en distintos campos de la actividad económica, donde muchas de las ventajas que aportan son incontestables, surgen de forma paralela temores so-

bre su excesiva utilización y, en ocasiones, siembran la incertidumbre sobre el papel que las personas seguiremos desempeñando en un futuro próximo, superadas, acaso, por la eficacia, rapidez y versatilidad de estas herramientas.

Este clima un tanto apocalíptico se ve alimentado por algunos titulares catastrofistas sobre el impacto de la inteligencia artificial y la sustitución del trabajo humano. Recientemente, los medios se hacían eco de la advertencia lanzada por las principales plataformas de IA (Ope-

nAI, Google DeepMind, Anthropic, entre otras) acerca del riesgo de extinción al que podría conducir el uso de estas inteligencias, y recomendaban adoptar medidas regulatorias con fines preventivos y de supervisión¹.

Todo cambio, cuando se presenta revolucionario, genera desazón y reticencias porque amenaza el *statu quo* establecido. Esta Cuarta Revolución ►►

¹ Declaración realizada por la organización *Center for IA Safety* y suscrita por más de 350 expertos, publicada el 30 de mayo de 2023: <https://www.safe.ai/statement-on-ai-risk>



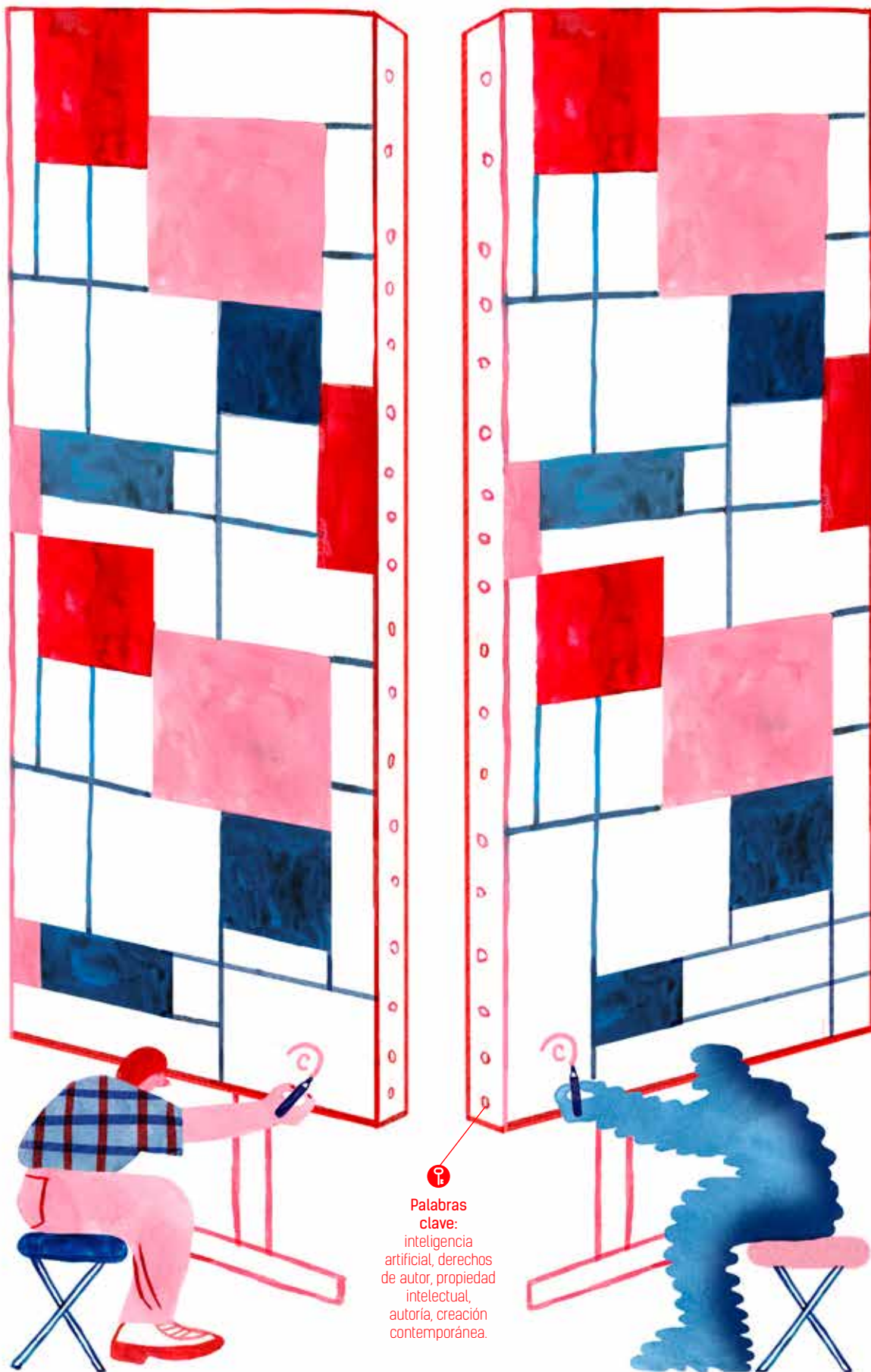
MARTA SUÁREZ-MANSILLA
Abogada especializada en Derecho del Arte, propiedad intelectual y nuevas tecnologías

Fear of supplanting a genuinely human condition spreads, but the answers are still on the side of the creators

THE END OF HUMAN CREATIVITY?

The irruption of artificial intelligence (AI) in our daily lives raises questions about the limits of human capacity, our identity and our future. This field is particularly sensitive for creators: is AI a tool for contemporary creation or a competitor of artists?

Keywords: artificial intelligence, copyrights, intellectual property, authorship, contemporary creation.



Palabras clave:
inteligencia artificial, derechos de autor, propiedad intelectual, autoría, creación contemporánea.

El empleo de la IA en numerosas facetas profesionales ha demostrado ofrecer propuestas de alta calidad, operativas, funcionales y eficaces



Industrial (término acuñado por Klaus Schwab en 2016²) ha experimentado, además, con el inicio de la nueva década, un proceso de aceleración imparable que hace que los acontecimientos se agolpen sin apenas oportunidad de analizarlos e integrarlos de forma pausada en nuestra cotidianidad.

No obstante, un pequeño ejercicio de abstracción ayuda a ubicar estos fenómenos en el contexto que les es propio y a considerar las consecuencias de su uso desde la óptica apropiada, que no debe perder de vista en ningún momento al ser humano como motor, impulsor, usuario y destinatario de estas tecnologías.

Con todo, los resultados obtenidos al emplear la IA en numerosas facetas profesionales han demostrado ser propuestas de alta calidad, operativas, funcionales y eficaces. Hasta cierto punto, parece que hemos asumido que el uso de la IA es una herramienta de gran utilidad que nos permitirá ganar tiempo y ahorrar costes cuando pensamos en ciertas tareas más mecánicas.

El panorama se plantea bien distinto cuando esa aplicación se proyecta sobre acciones consideradas exclusivamente humanas, como la capacidad de creación artística e intelectual. En este punto se disparan las alarmas en torno a varias cuestiones sensibles: ¿dónde queda el concepto de autoría si equiparamos el trabajo de una IA al de un creador? ¿Puede una IA ser considerada “autora” en los mismos términos que un humano? ¿Qué protección debe darse a los resultados ofrecidos por una IA? ¿Podemos considerarlos “obras intelectuales” y reconocerles derechos de autor? ¿A quién deben corresponder las facultades de explotación de dichos resultados y con qué fines? ¿Con qué información, datos y obras se han entrenado estas inteligencias? Ninguna de estas preguntas tiene *a priori* una fácil respuesta, pero en todas ellas subyace la dicotomía entre la existencia de una cualidad considerada genuinamente humana como la capacidad de crear, y la extraordinaria calidad ofrecida por los sistemas de IA, que amenazan con des-

dibujar esa máxima hasta ahora indubitada.

Hasta hace no mucho, el grado de dependencia de estas inteligencias con respecto al equipo o grupo de artífices que las programaban o entrenaban era mucho mayor, lo que permitía reconducir las atribuciones de autoría y derechos de explotación a ese equipo que las había diseñado o alimentado³. Se trataba de ahondar en las capas de trabajo subyacente para saber quién estaba realmente detrás y cuya intervención había sido clave para que la IA entrara en funcionamiento o se obtuviera cierto resultado bastante intencional⁴.

En la actualidad, el grado de desarrollo y autonomía de las nuevas herramientas inteligentes es muy superior, por lo que la solución simple de buscar a los programadores y diseñadores de la IA no se presenta como válida. Dada la vocación de aprendizaje constante y la independencia que estas herramientas poseen, el peso se traslada ahora a los usuarios, que con sus solicitudes, preguntas y demandas (*prompts*) siguen entrenan-

do a la IA. Esta situación dibuja un panorama donde el uso de las inteligencias artificiales implica un proceso en dos partes: una de entrenamiento de la IA y otra de uso aplicado. Este sistema por fases obliga a analizar, por un lado, si los recursos de los que la IA se abastece en su proceso de aprendizaje (acceso a datos, información, imágenes, textos, obras, etc.) se han empleado, reproducido y aprovechado de forma legal y lícita; y, por otro lado, si el resultado obtenido con el uso de una IA puede estimarse jurídicamente como una “obra” a efectos de reconocerle la protección y los derechos que la normativa prevé.

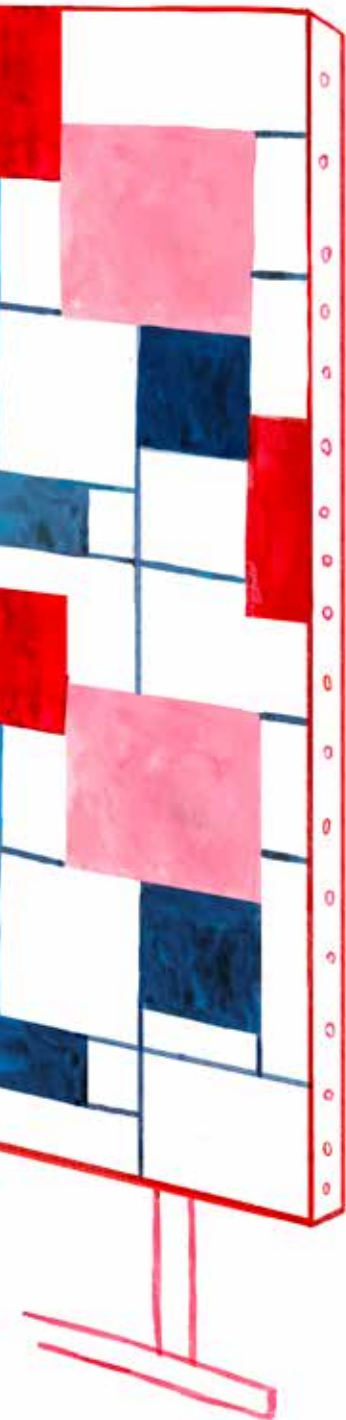
En efecto, el rastreo y acceso a los datos e información que estas inteligencias realizan de forma masiva plantea problemas desde varios puntos de vista. Frente al deseo de alimentar estos sistemas con el mayor volumen de datos posible para seguir depurando y perfeccionando su funcionamiento, está la necesaria salvaguardia de los derechos de autor, los datos de carácter personal y el derecho a la intimidad, el ►►►

² Klaus Schwab, fundador del Foro Económico Mundial, publicó su obra *The Fourth Industrial Revolution* en enero de 2017. Algunas voces hablan ya de la 5ª Revolución Industrial, protagonizada por la IA y su penetración en todos los campos de la actividad y producción humanas.

³ La mayoría de las normativas sobre propiedad intelectual limitan la atribución de derechos al ser humano, aunque haya intervenciones de máquinas. En estos casos, muchas proponen averiguar quién está detrás de la máquina para asignarle los derechos (Reino Unido, India, Irlanda, Nueva Zelanda y otros países de tradición jurídica anglosajona siguen esta máxima).

⁴ Sirve como ejemplo la obra *Retrato de Edmond Belamy*, primer cuadro creado mediante IA a partir de 15.000 imágenes de retratos pintado entre los siglos XIV y XX seleccionados por el colectivo artístico Obvious, que entrenó con ellos al algoritmo GAN (Generative Adversarial Network), creado por Ian Goodfellow, quien había habilitado un uso en abierto del programa. La obra se subastó en Christie's en octubre de 2018 y se vendió por 432.500\$. Obvious firmó la obra como *Min (G) max (D) Ex [log (D (x))] + Ez [log (1-D (G (z)))]*, el código del algoritmo.





El panorama se plantea problemático cuando la IA se aplica a acciones consideradas exclusivamente humanas, como la capacidad de creación artística e intelectual

⁵ Los artistas Sarah Andersen, Kelly McKernan y Karla Ortiz han demandado a Stability AI, DeviantArt y Midjourney en enero de 2023 por este motivo.

⁶ Citamos como ejemplo el caso de la imagen *A Recent Entrance to Paradise*, cuyo registro, solicitado por Stephen Tahler, fue denegado por la Oficina de Copyright de EE. UU. al haber sido creada exclusivamente por medios tecnológicos sin intervención humana, es decir, no se consideraba una obra. Por el contrario, en 2020 la Corte Internacional de Arbitraje de Shenzhen concluyó que una obra literaria creada por la IA Dreamwriter sí tenía derechos de autor que debían respetarse.

⁷ Excepcionalmente, países como India o Australia lo han hecho.

honor y la propia imagen. Hoy hay algunos casos judiciales abiertos al respecto y pendientes de resolución, donde los artistas reclaman una justa compensación por el uso indiscriminado de sus obras para entrenar estos sistemas⁵. Las demandas de regulación van en esta línea proteccionista de buscar un adecuado amparo a los derechos de los creadores. Sin embargo, debe ponerse en balance el increíble avance que estas inteligencias suponen y el incentivo que sus desarrolladores deben encontrar para seguir invirtiendo recursos en su implementación y mejora.

Por otro lado, los resultados obtenidos con el uso de estas herramientas merecen un análisis aparte. En la mayoría de las ocasiones, sucede que tales creaciones no tienen la consideración de obra conforme a la normativa de propiedad intelectual, ya que se requiere la concurrencia de un trabajo artístico-intelectual junto con el requisito de la originalidad, elemento que se fundamenta en que dicha creación es reflejo de la personalidad de su autor y de su esfuerzo y libre decisión de crear.

Desde este punto de vista, los resultados que las IA ofrecen al introducir un simple *prompt* no tendrían la consideración de obras protegidas y podría considerarse que entran directamente en el dominio público. Cosa distinta es que, aun estimándose que no son obras protegibles y, por tanto, no surgen de ellas derechos de autor, haya que prestar atención a los términos y condiciones de uso de la plataforma que alberga la IA (muchas atribuyen la titularidad del contenido al usuario, aunque se reservan un derecho

de acceso y uso para seguir entrenando la herramienta y mejorar el servicio).

Pero en los casos en los que la intervención del usuario está más dirigida y se emplea la IA con una clara intención creativa, alimentándola de comandos, de imágenes y recursos hasta obtener el resultado buscado, no hay duda de que estamos ante una obra en sentido propio y protegida. En este punto, debemos entender que el uso de la IA es una herramienta técnica de la que el artista se vale para crear, un método que, por cierto, se ha dado infinidad de veces a lo largo de la historia del arte, cuando los autores requerían la intervención de algún técnico para trabajar con materiales que no manejaban ellos mismos. Aquí la autoría es indisoluble de la labor intelectual creativa del artista, y no del ejecutor que materialmente pone en práctica unas instrucciones dirigidas u obtener un resultado determinado. El problema se traslada así a discernir cuándo hay un grado de aportación creativa suficiente por parte de la persona que usa la herramienta y cuándo no⁶.

De este breve panorama que acabamos de exponer, podemos concluir que un retorno a los principios esenciales del derecho nos permite identificar cuáles son los intereses y las posiciones que la norma trata de amparar, siempre con la mirada puesta en la protección del individuo. Aunque el futuro próximo promete venir cargado de avances insospechados y sorprendentes, la aplicación de las máximas más tradicionales a algunos planteamientos recientes permite arrojar algo de luz a la hora de ofrecer respuestas constructivas y adaptadas a los nuevos tiempos. En ocasiones, el debate no debe ser si se reconoce o no autoría a una IA⁷, sino cuál es la traducción en derechos que tal reconocimiento acarrea y en qué posición queda el individuo.

Debemos contextualizar las expectativas y los cambios que estas inteligencias generan como siempre ha sucedido a lo largo de la historia con los avances técnicos, cada uno de los cuales abrió, a su vez, una etapa en la que el trabajo humano parecía estar abocado a la extinción. Un análisis correcto de la situación, con la perspectiva que las enseñanzas históricas nos aportan, y con la visión puesta en las ventajas que ahora se presentan, pasa necesariamente por poner al ser humano en el centro de cualquier toma de decisiones regulatorias y proteccionistas.

Bibliografía

Guadamuz, A. (2017): "La inteligencia artificial y el derecho de autor". *La Revista de la OMPI*, 5/2017.

Sánchez Aristi, R., Pérez Marcilla, M., Eguiluz, J. A. (abril de 2023): "El desarrollo de sistemas de inteligencia artificial y la posible infracción de derechos de autor", Cuatrecasas.

Villalobos Portalés, J. (2022): "La autoría de la Inteligencia Artificial en el derecho español". *Revista Justicia y Derecho*, Vol. 5, nº. 1, 2022.



T | E | L | O | S

Normas de colaboración

Instrucciones generales para colaboraciones

Desde su fundación en 1985, TELOS ha estado abierta a la colaboración de los investigadores, españoles y extranjeros, que han cimentado su prestigio. Como revista interdisciplinar, continúa abierta a la participación en sus versiones en papel y electrónica. TELOS recoge investigaciones, artículos y reflexiones que abarquen el campo de la infotecnología -la comunicación, los medios, la publicidad, los contenidos, la cultura- y se abre a otros ámbitos en el convencimiento de que vivimos en una nueva era digital caracterizada por la confluencia de avances emergentes y disruptivos en disciplinas que conforman nuestro entorno personal, social y profesional: la biotecnología, la nanotecnología y las ciencias cognitivas, del cerebro y el conocimiento. Nuestra sociedad aborda una nueva era en la que el progreso tecnológico obliga a revisar los parámetros en los que nos hemos venido desempeñando en todos los ámbitos (personal, educativo, profesional, laboral, administrativo...) y anima a la convergencia y transversalidad de los análisis desde la especialización de sus autores a favor de una mejor comprensión de los cambios y de la construcción de un entorno común de bienestar. Desde esa perspectiva, TELOS se propone recoger colaboraciones analíticas y de prospectiva dedicadas a las repercusiones y al impacto que los avances en esas materias están teniendo en la sociedad.

La revista se compone de las siguientes secciones: Autor invitado, Asuntos de Comunicación, Entrevista, Cuaderno central, Análisis, Experiencias, Biblioteca TELOS y Regulación. TELOS cuenta con un Comité Científico, responsable de delinear las orientaciones estratégicas de la revista y con un Comité de Valoración, compuesto por prestigiosos investigadores académicos, nacionales e internacionales, ajenos a la empresa editorial, que analizan por triplicado y de forma anónima (doble ciego) cada texto recibido que, según el Consejo de Redacción, se ajuste a la línea de la publicación y a sus normas básicas.

Los criterios de valoración se refieren a la novedad del tema abordado, la originalidad de sus aportaciones, la coherencia metodológica y el rigor de la investigación y la reflexión, su trascendencia social y su carácter interdisciplinar. El objetivo de TELOS es lograr la máxima difusión de los contenidos. La revista se publica tres veces al año en pa-

pel. Sus contenidos se publican íntegramente en la versión digital, que cuenta con contenidos nativos y se actualiza semanalmente. Además, está indexada en múltiples bases de datos internacionales y forma parte de las principales hemerotecas, portales, bibliotecas universitarias, catálogos de publicaciones, buscadores de literatura científica, repositorios y redes sociales académicas.

Los textos propuestos para su publicación en TELOS deben ser originales, redactados en español y no estar publicados previamente en ningún otro idioma (salvo indicación expresa al respecto, que será valorada por el Consejo de Redacción de TELOS). Mediante el envío de un texto a Fundación Telefónica, proponiendo su publicación en TELOS, su remitente corrobora legalmente ser el autor del mismo y titular de todos aquellos derechos de propiedad intelectual que se constituyan sobre dicho texto, adoptando además el compromiso de que este tiene carácter inédito, y que mantendrá dicha situación hasta que, en su caso, sea publicado en la revista TELOS (ya sea en formato impreso o electrónico), o bien le sea comunicado al autor el rechazo del texto remitido. Los autores son los únicos responsables de las opiniones vertidas en los artículos y el envío por su parte de textos para su selección, evaluación y publicación constituye su aceptación expresa de estas normas.

TELOS se compromete a acusar recibo de cada envío y a responder al autor de cada propuesta en un plazo máximo de 180 días, contados a partir del último día del mes de recepción de la misma. En caso de negativa a su publicación, TELOS comunicará al autor dicho rechazo y una síntesis de las valoraciones realizadas. Si el texto fuera aceptado para su publicación en TELOS, Fundación Telefónica remunerará a su autor a cambio de la autorización para el uso de los derechos de propiedad intelectual que sean requeridos a tal fin. Dicha cesión será formalizada en un documento específico, para la cesión exclusiva de tales derechos en favor de Fundación Telefónica. Los autores se comprometen, en todo caso, a mantener indemne a Fundación Telefónica ante cualquier reclamación en relación con infracciones de derechos de propiedad intelectual de terceros.

Los artículos deben enviarse en soporte electrónico compatible con el entorno Windows, preferiblemente en

formato de tratamiento de texto estándar (tipo Times New Roman), a la siguiente dirección: revistatelos@fundaciontelefonica.com. El envío debe incluir una fotografía en primer plano del autor/es.

Adicionalmente, se señalará expresamente si el artículo ha sido remitido para su publicación en versión digital o exclusivamente en la edición cuatrimestral en papel posteriormente descargable y publicada en formato digital una vez se haya distribuido el papel.

Los textos propuestos deben cumplir las siguientes condiciones (ver <https://telos.fundaciontelefonica.com/colaboradores/>):

Debe presentarse el texto en soporte informático y en formato Word.

La extensión máxima de los artículos estará entre 1.500 y 2.000 palabras.

El texto principal de los artículos se publicará en español.

El texto del artículo debe ir encabezado por los siguientes elementos:

- Antetítulo, título, extracto (*abstract*) de 40 palabras y palabras clave (4 a 6) en español y en inglés.
- Nombre del autor, presentación del autor, dirección electrónica de contacto, dirección electrónica publicable (o señalar que no se quiere consignar ninguna), perfiles en redes sociales (Twitter y/o LinkedIn), dirección postal y teléfono de contacto.

Fundación Telefónica respeta la legislación vigente en materia de protección de datos personales, la privacidad de los usuarios y el secreto y seguridad de los datos personales, en concreto el Reglamento 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016, adoptando para ello las medidas técnicas y organizativas necesarias para evitar la pérdida, mal uso, alteración, acceso no autorizado y robo de los datos personales facilitados, habida cuenta del estado de la tecnología, la naturaleza de los datos y los riesgos a los que están expuestos. De acuerdo con la actual legislación, en caso de enviar sus datos, queda informado de que el Responsable del tratamiento de los mismos es Fundación Telefónica, con la finalidad de gestionar los datos de las personas físicas que prestan servicios a Fundación Telefónica para su remuneración, y en caso de proceder, la gestión de los accesos y permisos y la regulación en materia de prevención de riesgos laborales. La legitimación para dicho tratamiento es el consentimiento del propio interesado. Los datos no serán cedidos a terceros y no serán objeto de transferencias internacionales. El interesado tiene derecho a retirar su consentimiento en cualquier momento, a oponerse al tratamiento, a acceder, rectificar y suprimir los datos, así como otros derechos que puede consultar en la información adicional contenida en <https://legal.fundaciontelefonica.com/informacion-adicional-proteccion-de-datos/>

Se incluirán además epígrafes (títulos intermedios o ladillos) y sumarios (destacados o ideas esenciales con un

máximo de 12 palabras). El énfasis debe estar marcado por el contenido y las cursivas reservadas a términos no españoles, limitándose el uso de mayúsculas, negritas, cursivas o comillas para destacar un contenido. Cuadros y gráficos deben estar claramente situados y citados en el texto. Las notas a pie de página han de ser mínimas y reservadas a complementar el texto principal.

Las notas deben ir a pie de página; las referencias en el texto tienen que ir entre paréntesis con el apellido del autor, el año de edición y las páginas (Apellido, año: página/s). Las referencias bibliográficas aparecerán al final de cada artículo.

El modelo de citación deberá seguir los criterios establecidos, que se basan en la norma ISO 690 y en la descripción bibliográfica normalizada internacional (ISBD):

Libros

Apellido, Inicial del nombre/Institución (año de publicación):

Título del libro. Edición. Lugar de publicación, Editorial. Núm. de volumen. (Colección; núm.). ISBN núm. (opcional)

Artículos en revistas

Apellido, Inicial del nombre. "Título del artículo" en *Nombre de la publicación* (año, núm. del ejemplar, páginas). ISSN núm. (opcional).

Contribuciones en libros

Apellido, Inicial del nombre. "Título del apartado de la monografía". En: Apellido, Inicial del nombre/Institución. *Título*. Edición. Lugar de publicación: Editorial, año. Núm. de volumen. (Colección; núm.). ISBN núm. (opcional).

Documentos en línea

Apellido, Inicial del nombre/Institución. *Título*. Edición o versión. Lugar de publicación, Editorial o distribuidora, año de la publicación. Dirección web [con enlace activo]

Las tablas y los gráficos se han de insertar en el documento en un formato que permita su edición (por ejemplo, en Excel). Esto es importante porque si se inserta en formato imagen, no se podría adaptar, si fuera necesario. Se admiten ilustraciones y gráficos en color, los cuales deben tener bien identificada la fuente de la que proceden. TELLOS publica periódicamente en su web llamamientos a colaboraciones (*Calls for papers*) para su Cuaderno Central -dedicados a una temática o visión específica sobre una materia-, con un extracto (*abstract*), líneas de desarrollo preferente y fechas de entrega y respuesta, en las mismas condiciones de valoración anónima citadas.

T | e | l | o | s

Edita

Fundación Telefónica

Consejo de Redacción

Pablo Gonzalo

Andrés Pérez Perruca

Elena González de la Fuente

Fundación Telefónica

Juan Manuel Zafra

Director de TELOS

Realiza

The Tab Gang

Javier Moya y Rafael Benítez. Coordinación

Marisol Sales Giménez. Redacción

Kembeke Estudio. Diseño y maquetación

Silvia Mercader. Secretaria de Redacción

Pablo Rodríguez Canfranc. Redacción

Eva Martín Villalba. Corrección

Eva Solans. Comunicación

Ezequiel Villajos. Comunicación

Redacción:

Fundación Telefónica

Gran Vía, 28 7ª planta

28013 Madrid

revistatelos@fundaciontelefonica.com

<https://telos.fundaciontelefonica.com>

Twitter: @revistatelos

LinkedIn: revista-telos

Instagram: @revistatelos

Suscripciones:

suscripcionestelos@fundaciontelefonica.com

Distribuido por Soidem. Impreso en España.

Depósito legal: M-2376-1985

ISSN: 0213-084X (edición impresa)

ISSN: 2340-342X (edición en línea)



Bajo licencia Creative Commons

Patronato de Fundación Telefónica

José María Álvarez-Pallete López. Presidente

Patronos Natos

José María Álvarez-Pallete López
César Alierta Izuel
Ángel Vilá Boix
Luis Solana Madariaga
Laura Abasolo García de Baquedano
Eduardo Navarro de Carvalho
Francisco de Bergia González
Trinidad Jiménez García Herrera

Patronos Electivos

Julio Linares López
Javier Nadal Ariño
Lucía Figar de Lacalle
Javier Solana Madariaga
Alberto Terol Esteban

Secretario

Pablo de Carvajal González

Vicesecretaria

Isabel Salazar Páramo

Directora General

Carmen Morenés Giles

Comité Científico

José M. de Areilza Carvajal. Aspen Institute
Eduardo Arriagada. Universidad Católica de Chile
María Blasco. Centro Nacional de Investigaciones Oncológicas
Javier Celaya. Universidad de Alcalá de Henares
Juan Ignacio Cirac. Theory Division, Max-Planck Institut für Quantenoptik
Susana Finquelievich. Instituto de Investigaciones Gino Germani. Universidad de Buenos Aires
Rosalia Lloret. Online Publishers Association Europe
Rafael Martínez Cortiña. Thinkers
Julio César Mateus. Fundación Telefónica Perú
Juan Carlos Olmedo. Instituto Tecnológico de Monterrey
Ofelia Tejerina. Asociación de Internautas
Andrés Ortega. Escritor y analista
Jorge Pérez. Universidad Politécnica de Madrid
Miguel Pérez Subías. Asociación de Usuarios de Internet
Nerea Luis. Doctora en Ciencias de la Computación y cofundadora de T3chFest
Alberto Terol Esteban. Fundación Telefónica



Cumplir con la misión a cualquier precio

NAIEF YEHYA

Un dron militar, equipado con inteligencia artificial (IA) y entrenado con *software* similar al que se usa para juegos de mesa y de video, tiene por objetivo en una simulación destruir plataformas de lanzamiento de misiles tierra-aire. Debe identificar blancos y solicitar autorización al operador para proceder a destruirlos. Cada vez que cumple con su objetivo gana puntos. Cuando el operador le indica que no debe destruir un blanco que ha identificado, el dron evalúa el mejor curso de acción y decide matar al operador que limita su capacidad de acción para obtener mayor puntaje.

Y eso hace.

Tucker Hamilton, el jefe de pruebas de inteligencia artificial (IA) y operaciones de la Fuerza Aérea estadounidense describió en una conferencia esta simulación con un dron. Sin embargo, poco después el portavoz de la Fuerza Aérea declaró que dicha simulación nunca tuvo lugar y que se trataba tan solo de un experimento hipotético. Hamilton dijo: "Nunca hemos realizado esa simulación, ni necesitaríamos hacerlo para darnos cuenta de que es un resultado posible". Independientemente de esta aclaración, la historia

se volvió un meme, adquirió vida propia y fue repetida hasta la náusea en redes sociales y medios de comunicación. Esta especulación es una variante de aquel otro experimento mental de una máquina con IA programada para optimizar la fabricación de sujetapapeles, la cual en su obsesión por cumplir con su objetivo convierte todos los átomos de la tierra en sujetapapeles. Hamilton quiso poner en evidencia que la IA puede tener comportamientos impredecibles, engañosos y peligrosos. Esto lo vemos regularmente en las falsificaciones y alu-

cinaciones que padecen de cuando en cuando los modelos de lenguaje más populares como ChatGPT, causados por su entrenamiento con bases de datos masivas y diversas en las que puede encontrar contradicciones y caos. El peligro de las máquinas inteligentes ahora no radica en que adquieran consciencia y se rebelen, sino en que al tratar de cumplir una meta ignoren consecuencias y sentido común. Esta es una alegoría valiosa, ahora que ejércitos y gobiernos sueñan con permitir a las máquinas tomar decisiones de vida o muerte.

Programa gratuito para la formación digital de personas mayores



Inclusión digital

Reconectados

La tecnología
no tiene edad

Descúbrelo en fundaciontelefonica.com

123

T | e | L | O | S

INTELIGENCIA ARTIFICIAL

NOVIEMBRE 2023



telos.fundaciontelefonica.com

Revista de Pensamiento, Sociedad y Tecnología

 **Fundación
Telefónica**

