

Regulación de la inteligencia artificial

PERE BRUNET

El término inteligencia artificial (IA) se asocia actualmente a la inteligencia artificial basada en datos. Mientras que en las últimas décadas del siglo XX la IA se planteaba modelar el conocimiento humano mediante modelos informáticos y a partir del análisis de cómo las personas abordamos y resolvemos los problemas, la IA basada en datos se nutre del aprendizaje a partir de conjuntos masivos de datos.¹ Tras una fase de aprendizaje con algoritmos que van conformando la red neuronal del sistema de IA, durante su uso el sistema ya no utiliza ningún algoritmo. Simplemente difunde la información de la entrada por la red neuronal para obtener el resultado o respuesta del sistema.²

Estos sistemas de IA están cada vez más integrados en nuestras vidas. Los encontramos en los sistemas de traducción automática, en la mejora automática de las fotos que realizamos en nuestros móviles, en los sistemas publicitarios que nos sugieren otras posibles compras, en los anuncios que nos proponen viajes y hoteles. Pero también en los sistemas generativos que crean texto e imagen, en el diagnóstico médico, en informaciones falsas que nos llegan, en mensajes electorales hechos a medida, en sistemas de control social, en sistemas de armamento e incluso en armas autónomas que pueden matar sin intervención humana. En algunos de estos casos, sus posibles errores son poco preocupantes y fáciles de corregir. En otros pueden afectar a personas concretas en situaciones críticas con

¹ Ramón López de Mántaras y Pere Brunet (en este mismo número) «¿Qué es la inteligencia artificial?»

² Podemos considerar, esquemáticamente, que una red neuronal ya entrenada con datos es como una inmensa hoja de cálculo en que cada celda es una neurona y cada columna es una capa. Para cada celda/neurona de una determinada capa, su valor se calcula combinando con distintos pesos los valores de múltiples celdas de la capa anterior (multiplicando los valores de estas celdas por determinados pesos y sumando todos estos resultados parciales). A continuación, cada celda aplica una función no lineal a su valor para mantenerlo dentro de límites aceptables, la columna/capa queda ya calculada, y el proceso se repite con las celdas de la siguiente columna. Cuando deseamos resolver un problema específico, por ejemplo traducir una frase o saber lo que hay en una imagen, nuestra información de entrada se traduce a valores numéricos, se copia en la primera columna/capa de la red neuronal, y esta va calculando secuencialmente los valores de sus celdas hasta llegar a la última columna, que, a falta de la transcripción semántica de sus valores numéricos, los ofrece la respuesta de la red neuronal a nuestro problema. La fase previa de aprendizaje consiste en ir ajustando los valores de los pesos de todas las neuronas en base a los datos de entrenamiento.

efectos ciertamente graves. Es por ello que la regulación de los sistemas de IA es apremiante, como indicamos en los siguientes apartados.

Regulación: una necesidad urgente

Virginia Eubanks, profesora de ciencias políticas en la universidad de Albany, observó que las nuevas herramientas de inteligencia artificial no son neutras, ya que castigan a los más pobres y acaban incrementando las desigualdades sociales. En su libro³ indicó que han llevado a un régimen de vigilancia en los Estados Unidos que incluye la creación de perfiles personales y que termina con medidas que significan castigo y exclusión para los que menos tienen. Gente castigada porque su delito es ser pobre, y por sistemas automáticos de IA basados en aprendizaje profundo que toman decisiones que aceptamos sin discutir ni analizar, y que acaban quitando derechos a quienes ya no tienen casi nada, incrementando automáticamente las desigualdades sociales.

Y es que además estamos entrando en la era del capitalismo de la vigilancia y del control. Todo el planeta se está viendo inmerso en un panóptico digital de dimensiones gigantescas, un panóptico al que nadie escapa. Tanto Google como las redes sociales adoptan formas panópticas que se presentan como espacios de libertad.⁴ En este capitalismo de la vigilancia, las experiencias vividas por la gente acaban siendo captadas por empresas privadas y convertidas en flujos de datos con valor comercial.⁵ La privacidad de unas personas acaba siendo un gran negocio para otras.

Todo el planeta se está viendo inmerso en un panóptico digital de dimensiones gigantescas, un panóptico al que nadie escapa

³ Virginia Eubanks, *Automating inequality*, Macmillan Publishers, Londres, 2018. Libro ganador del premio «Lillian Smith Book Award» en 2019: <https://us.macmillan.com/books/9781250074317/automatinginequality> - A partir de un trabajo de campo en el que la autora realizó un total de 96 entrevistas, el libro analiza tres casos: el del sistema de salud del estado de Indiana, el de la asignación de vivienda social en Los Ángeles, y el del algoritmo Allengheny de vigilancia parental en Pittsburgh, con un total de 29 entrevistas en Indiana, 34 en Los Ángeles, 33 en Pittsburgh. El tejido del libro son las desgracias de toda esa gente que conoció personalmente, casos concretos que Virginia Eubanks va visibilizando uno a uno. Ver también: <https://www.nytimes.com/2018/05/04/books/review/automating-inequality-virginia-eubanks.html> y, entre otros: <https://www.npr.org/sections/alltechconsidered/2018/02/19/586387119/automating-inequality-algorithms-in-public-services-often-fail-the-most-vulnerab?t=1551540438028>

⁴ Byung Chui Han, *La sociedad de la transparencia*, Herder Ed., Barcelona, 2013.

⁵ Shoshana Zuboff, *La Era del capitalismo de la vigilancia*, Paidós y Ara (en catalán), Barcelona, 2019, disponible en: https://www.ara.cat/opinio/shoshana-zuboff-capitalisme-vigilancia_129_2605741.html

Los sistemas de IA basada en el conocimiento son y nos podrán ser útiles, pero en la situación actual desregulada presentan un buen número de aspectos problemáticos que acaban afectando a personas concretas. En bastantes casos se utilizan datos privados (e información de documentos con derechos de autor) sin ningún consentimiento, no hay garantía de que los datos que nutren su aprendizaje estén suficientemente libres de sesgos –habitualmente raciales o de género–, algunas aplicaciones –de control social, de vulneración de derechos humanos, sistemas militares letales– atacan la dignidad de las personas y sus derechos con total impunidad y sin asunción de responsabilidad ni posibilidad de rendición de cuentas, no hay suficiente garantía de que personas expertas acaben supervisando los resultados para evitar el efecto de los errores de los que estos sistemas adolecen en mayor o menor grado (errores que pueden llegar a ser absurdos y muy relevantes), y la “caja negra” de sus redes neuronales hace impracticable la comprensión del por qué, en un caso concreto, el sistema ha llegado a resultados claramente erróneos (lo que se conoce como falta de explicabilidad).

En este entorno, muchos actores exigen una regulación efectiva que ponga límites a los peligros que estamos creando.

El problema es que la tecnología siempre va por delante de la legislación, aunque ello no sea excusa para no regularla. Hay que legislar e imponer obligaciones a estas empresas tecnológicas en lugar de ponerles tantas alfombras rojas. Ellas mismas, que han creado esta tecnología de IA, hablan a favor de una regulación cuando en realidad lo que quieren es que se regule a su manera⁶ evitando los sistemas de código abierto, que ven como una amenaza a sus intereses económicos.⁷

Estado actual: propuestas y sombras

Nos encontramos en un momento de actividad por lo que respecta a la regulación de la IA. Un ejemplo de ello es el documento final sobre IA elaborado en la reunión

⁶ Jordi Escudé, «Entrevista a Ramón López de Mántaras», *Regió 7*, noviembre de 2023, disponible en: <https://www.regio7.cat/bages/2023/11/25/ia-omnipresent-molts-casos-sols-94954203.html>

⁷ Victoria S. Nadal, «Compañeros de trabajo tiranos: los mecanismos psicológicos detrás de la falta de ética», *Retina*, *El País*, 25 de enero de 2020: «Hay tres dinámicas que conducen a cruzar las líneas rojas: el sentimiento de omnipotencia, el entumecimiento cultural y la negligencia justificada. Eso sucede también en los negocios y en las relaciones que se gestan en el trabajo. Aunque es difícil encontrar pruebas de la falta de ética en los líderes, en los últimos años se han hecho públicos algunos casos que ilustran esta situación». Disponible en: https://retina.elpais.com/retina/2020/01/30/talento/1580382197_448234.html

plenaria número 49 del G-7, que se celebró en Hiroshima en mayo de 2023. Los líderes de los países de este grupo (Alemania, Canadá, Estados Unidos, Francia, Italia, Japón y Reino Unido), con un poder económico y militar indiscutible, se consideraron «más unidos que nunca en nuestra determinación de hacer frente a los retos mundiales de este momento y sentar las bases para un futuro mejor».⁸ Uno de sus acuerdos comportó la elaboración de un código internacional de conducta para organizaciones que desarrollan sistemas avanzados de IA, código que salió a la luz unos meses más tarde, en octubre de 2023.⁹ Incluye propuestas para identificar, evaluar y mitigar los riesgos a lo largo del ciclo de vida de la IA y para identificar patrones de uso indebido durante el uso y la comercialización, así como para mitigar los riesgos sociales y de seguridad. Pero todo ello no son más que recomendaciones para las «organizaciones que puedan respaldar este Código de Conducta», propuestas que «proporcionan orientación voluntaria para las organizaciones que desarrollan los sistemas de IA más avanzados». Estamos hablando de códigos voluntarios, pero no de regulaciones efectivas.

Por otra parte, la Unión Europea aprobó su ley regulatoria de la IA (*IA Act*), en diciembre de 2023. El Parlamento Europeo la aprobó provisionalmente en junio de 2023,¹⁰ siendo luego enmendada por parte de la Comisión y el Consejo, haciendo caso a diversas presiones por parte del *lobby* empresarial. La ley clasifica y considera los sistemas de IA en base a su riesgo:

- Sistemas de riesgo inaceptable: Incluyen la manipulación cognitiva y conductual de personas o grupos vulnerables (incluyendo niños), la clasificación de personas según su comportamiento, estatus socioeconómico o características personales, así como algunos sistemas de identificación biométrica (incluyendo los de reconocimiento facial). Estos sistemas serán prohibidos
- Sistemas de alto riesgo (sistemas de aviación, coches, ascensores, dispositivos médicos y hospitalarios, biometría, educación, juguetes, sistemas legales, gestión de personal en empresas, migración, control de fronteras y otros): En estos casos los sistemas deberán homologarse, evaluándose antes de su comercialización y también durante todo su ciclo de vida.

⁸ Ver por ejemplo: <https://www.consilium.europa.eu/es/meetings/international-summit/2023/05/19-21/>

⁹ El código puede descargarse de esta página web de la Comisión Europea: <https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-code-conduct-advanced-ai-systems>

¹⁰ Parlamento Europeo, AI Act junio de 2023, disponible en:

[https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI\(2021\)698792_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf) - Ver también: <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

- **Sistemas de riesgo limitado:** incluye los sistemas de inteligencia artificial que generan o manipulan contenidos de imágenes, audio o vídeo, pudiendo generar información falsa. En estos casos, los sistemas deben cumplir con unos requisitos mínimos de transparencia que permitan a los usuarios tomar decisiones informadas, para decidir si quieren seguir usándolas. Debe garantizarse que los usuarios sean conscientes de que interactúan con sistemas de IA.
- **Sistemas de IA generativa (texto, imágenes, etc.):** A estos sistemas se les pedirá que indiquen en cada caso que el contenido ha sido generado mediante IA; las empresas deberán publicar información sobre qué datos usados en el aprendizaje estaban sometidos a derechos de autor, y los sistemas deberán estar diseñados de manera que eviten la generación de contenidos ilegales. En todo caso, este es un aspecto no cerrado a nivel europeo.

Según la ley, la Comisión Europea creará un organismo externo e independiente de control, y además podrá prohibir por decreto los nuevos usos de la IA que vayan apareciendo, sin que deba acordarlo con el Parlamento Europeo o con los Estados miembros. Y las empresas tendrán unos dos años para adaptarse después de que la ley entre en vigor, cosa que podría suceder en 2025.

Algunas opiniones indican que la clasificación basada en riesgos de la ley es compleja, siendo siempre discutible. Por otra parte, uno de los atractivos empresariales que puede tener la *AI Act* es que si las empresas cumplen con las reglas asociadas a la categoría de riesgo de sus sistemas de IA, quedarán protegidas legalmente contra la responsabilidad por los daños que podrían provenir de sus productos.¹¹

Estados Unidos, sede de muchas empresas líderes en IA, hasta ahora ha sido el país menos intervencionista. En octubre de 2022, la Oficina de Política Científica y Tecnológica (OSTP) de la Casa Blanca publicó un documento técnico que describe cinco principios destinados a guiar el uso de la IA. El documento indica que estos sistemas deben ser seguros y eficaces, no discriminatorios, protectores de la privacidad de las personas y transparentes. Pero no es lo mismo un plan de intenciones que una ley regulatoria. Es por ello que las principales empresas (Amazon, Anthropic, Google, Inflection, Meta, Microsoft y OpenAI) han hecho promesas vagas que no es claro que se vayan a plasmar en resultados concretos. Mientras

¹¹ Matthew Hutson, «Rules to keep AI in check: nations carve different paths for tech regulation - A guide to how China, the EU and the US are reining in artificial intelligence», *Nature, News*, 8 de agosto de 2023, disponible en: <https://www.nature.com/articles/d41586-023-02491-y>

tanto, las agencias y los estados están empezando a trabajar. La Oficina de Gestión y Presupuesto (OMB) ha publicado un nuevo borrador de política sobre la innovación y la gestión de riesgos en el uso de la inteligencia artificial por parte de las agencias.¹² En California hay un proyecto de ley que está proponiendo que las

Las normas regulatorias actualmente más avanzadas provienen de Europa. Pero todas ellas plantean casi tantos interrogantes como propuestas

empresas implementadoras registren sus sistemas de IA indicando sus objetivos y explicando cómo usarlos. Y en Illinois, una ley de 2020 exige que las empresas anuncien y expliquen el uso de la IA en las entrevistas de empleo. Además, el estado cuenta desde hace tiempo con una ley que permite a los ciudadanos denunciar abusos por datos biométricos, incluyendo reconocimiento facial. Pero

las grandes tecnológicas estadounidenses no se muestran todavía activas en este tema de la regulación de la IA, desde la perspectiva de que no hay ninguna legislación regulatoria en el horizonte.

En China, el Gobierno está tratando de equilibrar la innovación con el mantenimiento de su estricto control sobre las corporaciones y la libertad de expresión. Sus políticas intentan compaginar una clara voluntad de control social con la protección de la información privada frente a prácticas invasivas de las corporaciones. Es interesante observar que las leyes y recomendaciones de la Administración China del Ciberespacio (CAC) se dirige a los sistemas de IA desarrollados por las empresas, pero no a los del Gobierno. Las empresas deben ser transparentes, usar datos personales de forma no sesgada, y deben evitar la difusión de contenido falso, privado, discriminatorio, violento, o que socave los valores socialistas chinos. Por otra parte, los proveedores de servicios que sintetizan imágenes, videos, audio o texto deben registrar los resultados y contrarrestar cualquier información errónea que generen.¹³

Canadá está planteando una Ley de Datos e Inteligencia Artificial que promete exigir transparencia, no discriminación y medidas de seguridad para lo que llama sistemas de IA de “alto impacto”, algo que no queda bien definido. Y el Secretario General de las Naciones Unidas, António Guterres, ha sugerido que podría ser necesario un nuevo organismo de la ONU para gestionar y regular la IA.¹⁴

¹² OMB, noviembre de 2023. Ver: <https://www.whitehouse.gov/omb/briefing-room/2023/11/01/omb-releases-implementation-guidance-following-president-bidens-executive-order-on-artificial-intelligence/>

¹³ Matthew Hutson (2023), *op. cit.*

¹⁴ Matthew Hutson (2023), *op. cit.*

En todo caso, aparte de aspectos interesantes que encontramos en instrumentos legales de diversos países, las normas regulatorias actualmente más avanzadas provienen de Europa. Pero todas ellas plantean casi tantos interrogantes como propuestas. La misma ley europea habla de que los sistemas de IA generativa deben «estar diseñados de manera que eviten la generación de contenidos ilegales», algo esencialmente imposible en unos sistemas que tendrán un porcentaje no despreciable de fallos que no serán explicables. Por otra parte, la clasificación de los sistemas de IA, según los riesgos tal como quedan definidos en la *AI Act* europea, es discutible y ha surgido en parte de las presiones de *lobbies* empresariales.¹⁵ Además, el futuro de la regulación de la privacidad y derechos de autor es todavía incierta, la ley ignora prácticamente los sistemas militares, y un buen número de aplicaciones y usos no quedarán cubiertos.¹⁶ En concreto, el uso de IA para guiar armas autónomas es motivo de preocupación, y decenas de países han pedido a la ONU que regule los sistemas de armas letales autónomas.

Actualmente, los actores involucrados todavía discuten exactamente qué es lo que hay que controlar, cuáles son los riesgos que implica la IA y qué es lo que realmente debe restringirse. Y de hecho, mientras OpenAI y otras empresas están pidiendo públicamente más supervisión, estas mismas empresas se han resistido a algunos de los controles propuestos por la UE y están abogando por compromisos voluntarios, en lugar de leyes regulatorias.¹⁷

Como bien dice Matthew Hutson, sería mejor dejar de hablar de los futuros apocalípticos de la IA y hablar de sus riesgos actuales. Las administraciones deberían regular escuchando más las necesidades de las personas que los intereses de las grandes empresas. Analizando los riesgos que pueden acabar afectando a sus ciudadanos en vez de escuchar falsos relatos y visiones interesadas supuestamente utópicas.

¹⁵ La empresa OpenAI presionó para evitar que la IA generativa quedase clasificada como de alto riesgo, y lo consiguió, quedando finalmente en una categoría específica y distinta de las relacionadas con el riesgo.

¹⁶ Matthew Hutson (2023), *Op. Cit.*, “Las herramientas de generación de imágenes y los LLM, por ejemplo, pueden dar lugar a una proliferación de contenidos nocivos como “pornografía de venganza”, malware, estafas e información errónea y, en última instancia, podrían socavar la confianza de las personas en el sistema social. Una preocupación importante es qué tipo de transparencia debería exigirse a tales herramientas y si será posible hacerla cumplir. Por otra parte, como que estos sistemas están entrenados con inmensas cantidades de texto y arte generados por humanos, la violación de los derechos de autor también es un problema sin resolver”

¹⁷ Matthew Hutson (2023), *op. cit.*

La regulación que deseáramos

Nos encontramos en una encrucijada. Los sistemas actuales de IA pueden sernos

**Las personas, como
potenciales afectadas
por estos sistemas de IA,
deben ser los agentes
activos de su regulación**

útiles en muchos campos, pero el descontrol regulatorio nos ha llevado a sistemas pensados desde el negocio y no desde la gente. Steven Pinker decía que las distopías sobre la inteligencia artificial provienen de «una psicología parroquial alfa-masculina aplicada al concepto de inteligencia», pero que era posible concebir una IA para las personas, basada en valores femeninos.¹⁸ Ciertamente revelador.

Las actuales iniciativas regulatorias se enfrentan a dos importantes problemas. En primer lugar, son locales –o regionales– cuando lo que intentan regular son productos de grandes corporaciones internacionales que se mueven y actúan a nivel global. En segundo lugar, intentan nadar entre dos aguas, protegiendo tanto a las personas como al negocio de estas grandes empresas. El tuit de Ursula Von der Leyen tras la aprobación de la *AI Act* europea fue clarificador,¹⁹ como lo fueron las opiniones editoriales de diversos medios de comunicación.²⁰ En el primer caso, se reconocía que había que velar por los derechos fundamentales tanto de las personas como de las empresas; en el segundo, de declaraba que “nos convenía” que la regulación no fuese más estricta de la cuenta. Pero, ¿hay que buscar un “equilibrio” entre las necesidades y derechos de las personas, por una parte, y el negocio empresarial por otro? ¿Partimos de la ética y los derechos de la gente o del negocio?

Las iniciativas regulatorias difícilmente serán éticamente aceptables si dependen de la colaboración voluntaria de los principales actores públicos (estados) y priva-

¹⁸ Cita de Steven Pinker en: Ella Rhodes, «What do psychologists think about machines that think?», *The Psychologist*, enero de 2015. Texto de la cita: «Las distopías sobre la inteligencia artificial provienen de una psicología parroquial alfa-masculina aplicada al concepto de inteligencia, pero es significativo que muchos de estos tecno-profetas no consideren la posibilidad de que la inteligencia artificial se desarrolle en base a valores femeninos, siendo totalmente capaz de resolver problemas, pero sin necesidad de aniquilar personas o de dominar civilizaciones», disponible en: <https://thepsychologist.bps.org.uk/volume-28/march-2015/what-do-psychologists-think-about-machines-think>

¹⁹ Tuit de Ursula Von der Leyen, Nuestra «Ley de IA es una primicia mundial. Un marco legal único para el desarrollo de la IA en el que podemos confiar. Que vela por la seguridad y los derechos fundamentales de las personas y las empresas. Un compromiso que asumimos en nuestras directrices políticas y lo cumplimos. Acojo con satisfacción el acuerdo político de hoy» X, 8 de diciembre de 2023, 23:50.

²⁰ Extracto del editorial del periódico *Ara*: “...la regulación europea sobre la inteligencia artificial que se está discutiendo estos días, por mucho que pueda parecer que nos queda lejos, es más importante de lo que parece. Y al mismo tiempo nos conviene que no sea más estricta de la cuenta”, *Ara*, 9 de diciembre de 2023, disponible en: https://www.ara.cat/editorial/paper-clau-etic-tecnologic-ia_129_4879766.html

dos (grandes empresas). Las administraciones tienen el deber de proteger y cuidar a sus ciudadanos, y ello implica que la regulación debe ser participativa, dinámica, adaptativa a los cambios sociales y tecnológicos, y democrática. Es en este marco regulatorio basado en principios éticos y de respeto a las personas, en el que las empresas deberían ejercer su labor.

Las personas, como potenciales afectadas por estos sistemas de IA, deben ser los agentes activos de su regulación. Esto comporta definiciones claras y no ambiguas de los riesgos, porque no es lo mismo un error en un sistema de traducción automática que una disfunción en sistemas en los que los errores pueden causar víctimas mortales, y no es lo mismo un error en un sistema publicitario que un mensaje electoral fraudulento.

En el caso de la IA generativa, los problemas son realmente complejos y de futuro poco previsible. Pero en este caso, una posible idea regulatoria sería plantear que los textos e informaciones que aparecen y circulan por internet y redes sociales llevaran una “etiqueta digital” con marca de agua que indicase fecha, procedencia y autoría, al modo de las etiquetas que llevan los alimentos procesados. En una primera fase, podría empezarse en campos como el educativo y medios de comunicación, e incluso como una recomendación voluntaria. La identificación de la autoría permitiría conocer la persona física que se hace responsable del texto, imagen, audio o vídeo a todos los efectos, y en el caso de informaciones sin etiqueta, los lectores sabrían a qué atenerse.

Los marcos regulatorios deberían considerar múltiples aspectos. Entre ellos, la privacidad de los datos que se usan para el aprendizaje y los posibles derechos de autor;²¹ también, la transparencia, la publicación de la autoría y la asunción de responsabilidades en caso de que sea necesaria una rendición de cuentas; los posibles daños debidos a la desinformación o información maliciosa; la falta de explicabilidad de los sistemas de IA;²² los problemas asociados a los inevitables

²¹ En lo que respecta a la protección de los derechos de las personas durante el procesamiento de sus datos para el aprendizaje de los sistemas de IA, las congresistas estadounidenses Anna G. Eshoo y Zoe Lofgren y la senadora Kirsten Gillibrand plantean propuestas muy interesantes. Entre ellas, derechos sobre nuestros datos privados (derecho a saber qué organizaciones disponen de nuestros datos, derecho al acceso a los mismos y a su eliminación, derecho a poder fijar el período durante el cual podrán usarse, incluso el derecho a recibir un porcentaje del beneficio económico que se genere con nuestras informaciones), junto con deberes para las empresas e instituciones (no vender ni hacer públicos los datos privados sin consentimiento, no usarlos de modo que violen los derechos humanos, no reidentificar las personas a partir de datos agregados). Asimismo, proponen la creación de una agencia federal que vele por la protección y cumplimiento de estos derechos. Ver: <https://lofgren.house.gov/media/press-releases/lofgren-eshoo-reintroduce-sweeping-privacy-legislation>

errores que estos sistemas pueden cometer, que implican la necesidad, en aplicaciones críticas, de una postsupervisión de los resultados²³ por parte de una persona experta que asuma la responsabilidad de la decisión final.

En este mundo globalizado, la regulación de la IA no será totalmente efectiva si no es global, debiendo basarse únicamente en principios éticos y de respeto a los derechos y a la dignidad de todas las personas. Una regulación que debe preceder al negocio.

Pere Brunet i Crosa es doctor y catedrático jubilado de la Universidad Politécnica de Catalunya, investigador del Centre Delàs d'Estudis per la Pau y divulgador científico.



²² La falta de explicabilidad significa que es imposible llegar a entender por qué el sistema ha fallado en un caso concreto y, por tanto, es imposible evitar que el fallo se vuelva a repetir.

²³ Si los sistemas altamente sofisticados de traducción automática cometen errores, que aceptamos como mal menor y corregimos sin problema supervisando sus resultados, ¿puede tener algún sentido que aceptemos sin supervisión lo que indica un sistema de diagnóstico médico o un arma autónoma? Es por ello que consideramos que todo sistema de regulación debe incorporar elementos de postsupervisión, clarificación de las responsabilidades humanas y rendición de cuentas.