

INTELIGENCIA ARTIFICIAL Y CAMBIO GLOBAL

Pere Brunet

RELATOS

TIEMPO
DE
TRANSICIONES

TIEMPOS DE TRANSICIONES

- Relatos -

INTELIGENCIA ARTIFICIAL Y CAMBIO GLOBAL

Pere Brunet



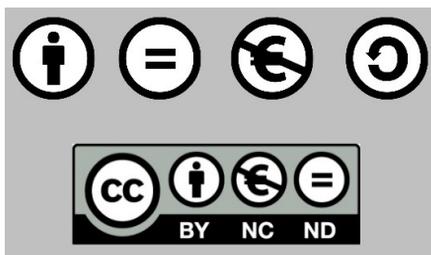
Primera Edición, 2024, Foro Transiciones.

Título: *Inteligencia artificial y cambio global*

Autoría: Pere Brunet

Diseño portada: Traficantes de Sueños [taller@traficantes.net]

Pere Brunet. Doctor y catedrático jubilado de la Universidad Politécnica de Catalunya. Divulgador científico e investigador del Centro Delàs de Estudios para la Paz, con interés en el estudio de las implicaciones sociales de la ciencia y la ingeniería y en el análisis del militarismo y terrorismo desde una perspectiva inspirada en bases científicas, así como en las aplicaciones de la ciencia en el ámbito de la paz.



INTRODUCCIÓN

Vivimos tiempos decisivos. Tiempos en los que las actuales generaciones tienen ante sí la responsabilidad de hacer frente a una crisis ecológica global, capaz de determinar el devenir de las sociedades humanas sobre el planeta.

Conscientes de que lo que está en juego son las mismas bases de la vida actual, las fundaciones Conama y Fuhem impulsaron en 2013 el Foro Transiciones, un *think tank* transdisciplinar y plural, con el objetivo de enriquecer el debate en torno al cambio de época y las temáticas que, desde el universo ecosocial, van a decidir el futuro de la humanidad.

El Foro ha tomado la iniciativa de impulsar la publicación de una serie de documentos que, bajo el lema “Tiempos de Transiciones”¹, ofrezcan análisis y propuestas para abordar procesos de cambio en nuestro país, tomando en consideración los marcos globales, especialmente el europeo. Los contenidos de la serie se orientan en tres líneas de trabajo: contribuciones generales a la construcción del relato sobre las transiciones; propuestas temáticas en cuestiones claves relacionadas con esas transiciones; y consideraciones en torno a temas de actualidad.

Esperamos que la iniciativa resulte útil para impulsar en la sociedad el debate sobre la importancia de los retos ecosociales para las actuales y futuras generaciones, porque, a pesar de nuestras lagunas de conocimiento, hemos de aceptar que sabemos lo suficiente para empezar a transformar una realidad en la que la vida, tal y como la conocemos, está en peligro por primera vez en la historia de la humanidad.

FORO TRANSICIONES

¹ La función del Foro Transiciones es auspiciar la publicación de trabajos que sean considerados de interés general, sin que ello signifique que, por su carácter plural, el Foro comparta colectivamente los contenidos que en cada caso expongan sus correspondientes autores.

INDICE

I. INTRODUCCIÓN.....	5
II. ALGORITMOS, AUTOMATIZACIÓN Y SISTEMAS HEURÍSTICOS.....	5
III. LA INTELIGENCIA ARTIFICIAL.....	9
IV. CARACTERÍSTICAS Y LIMITACIONES DE LA IA.....	12
V. LOS SISTEMAS CRÍTICOS Y LA IA MILITAR.....	18
VI. LA IA GENERATIVA.....	22
VII. EL COSTE ECOLÓGICO DE LA IA.....	26
VIII. HACIA UNA IA ECOLÓGICA Y HUMANA: REGULACIÓN Y REDISEÑO.....	34
IX. CONCLUSIONES.....	38

I. INTRODUCCIÓN

Los sistemas de inteligencia artificial han experimentado un auge espectacular en los últimos años, con aplicaciones que van desde el diagnóstico precoz en medicina hasta la robótica moderna, pasando por campos tan diversos como el de la traducción automática, los juegos por ordenador o la previsión del impacto del cambio climático. Los encontramos también 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 y también en un buen número de sistemas militares de combate.

En lo que sigue haremos primero un recorrido que nos llevará de los algoritmos a los sistemas automáticos y a los métodos heurísticos, para finalmente centrarlos en los sistemas de inteligencia artificial basados en datos (IA en lo que sigue). Hablaremos brevemente de sus características, de algunos de sus campos de aplicación y de los actuales sistemas de IA generativa, para terminar centrándonos en su huella ecológica, la necesidad de regulación y el camino hacia sistemas de IA basados en criterios éticos y ecológicos.

II. ALGORITMOS, AUTOMATIZACIÓN Y SISTEMAS HEURÍSTICOS

Un algoritmo no es más que un conjunto ordenado, finito y no ambiguo de reglas y operaciones que permiten resolver un problema o realizar una determinada actividad. Nuestra vida está llena de algoritmos, en plural. Las recetas de cocina y las

partituras musicales son algoritmos para preparar comidas e interpretar melodías. Y lo son las instrucciones de cualquier juego, los manuales de usuario y los protocolos que siguen los médicos en los hospitales. La palabra "algoritmo" viene de lejos. Proviene del nombre del sabio al-Khwârazm, que hace 1.150 años trabajaba en el observatorio de Bagdad y en la Casa de la Sabiduría de la misma ciudad. En su tratado de álgebra, Hissab al-jabr wa-l-muqabala explicó reglas y recetas (o sea, algoritmos) para repartir herencias, lo que en aquel tiempo y en el mundo árabe era complicado e implicaba muchos cálculos.

Por otra parte, vemos que los seres vivos pueden resolver problemas aparentemente muy complejos, y que los resuelven de manera óptima o casi-óptima. Con poca energía vital son capaces de encontrar soluciones que funcionan y que les ayudan a vivir. No hay más que observar a las hormigas o las abejas. La evolución ha ido cableando el algoritmo para la solución de sus problemas en su cerebro².

Los ordenadores y los teléfonos móviles funcionan en base a los algoritmos que algunos expertos antes han programado. Usamos algoritmos para mejorar fotos, para buscar información en internet, para saber la posición de los astros y para prever el tiempo que hará mañana. Disponemos de algoritmos que pueden encontrar el nombre de las músicas que escuchamos, y algoritmos que nos ayudan a traducir textos de un idioma a otro.

Vemos algoritmos también en los sistemas automáticos que

² Pere Brunet, *Ciencia, ética y paz: Historias del valle de los límites*, Dossier, FUHEM, 2023, pág. 128. Disponible en: <https://www.fuhem.es/2023/10/20/dossier-ecosocial-ciencia-etica-y-paz/>

realizan tareas rutinarias, previsibles y programadas desde su diseño y construcción: en los sistemas de climatización, en electrodomésticos, en el cobro mediante código de barras en las tiendas y en un sinfín de artilugios cotidianos.

En todo caso, debemos distinguir entre algoritmos exactos y algoritmos que podríamos denominar aproximados. El sistema GPS utiliza algoritmos específicos para calcular nuestra ubicación a partir de las señales que recibe de los satélites, algoritmos que son claramente fiables y robustos con un margen de error insignificante. Los algoritmos de los cajeros automáticos son asimismo fiables y exactos. Pero muchos de los algoritmos que nos rodean pertenecen a la categoría de aproximados. El motivo es que los problemas que tienen que resolver son demasiado complicados si tenemos en cuenta el tiempo que estamos dispuestos a esperar. Es lo que ocurre, por ejemplo, con los algoritmos de tratamiento de imágenes para la mejora de fotos. Cada vez que hacemos una foto con un teléfono móvil, ésta pasa por un proceso automático de mejora (de unas pocas décimas de segundo, y forzosamente sub-óptimo) antes de que la veamos.

Estos algoritmos aproximados que dan soluciones aceptables pero no óptimas se denominan también algoritmos heurísticos. Tienen varios parámetros ajustables, y su idoneidad en el contexto de una aplicación determinada dependerá justamente del ajuste de sus parámetros. La configuración de los parámetros, que nunca será óptima, en todo caso debería ser aceptable. Por ejemplo, las aplicaciones de mejora de fotos nos proponen determinados "filtros" que llevan a resultados específicos. No está claro cuál de ellos es mejor ni cuál es

óptimo, ya que todo depende del criterio de la persona usuaria. La diferencia entre unos y otros, además de determinados rasgos algorítmicos, es el valor de los parámetros que utilizan para filtrar la foto en cada caso. Son algoritmos heurísticos.

III. LA INTELIGENCIA ARTIFICIAL

Al hablar de IA nos referimos a aplicaciones informáticas diseñadas con el fin de actuar de manera similar a las personas humanas en ámbitos específicos.

Observemos que se trata de actuar, no de pensar o razonar, y de hacerlo en ámbitos muy concretos. Los sistemas de IA no piensan ni razonan, y tampoco anticipan. Además, una IA diseñada para mejorar fotos no sirve en traducción, ni viceversa. Es por tanto mucho más limitada que la inteligencia humana³.

Tras un auge inicial de los algoritmos de IA basados en conocimiento y reglas, durante las últimas décadas la inteligencia artificial se ha ido materializando básicamente en nuevos sistemas denominados de aprendizaje profundo o IA basada en datos, que son los que trataremos en lo que sigue. En otras palabras, hemos pasado de algoritmos y reglas que alguien programaba, a sistemas automáticos de IA que se

³ De hecho, es dudoso que “inteligencia” sea un término adecuado para definir este tipo de sistemas.

autoconfiguran en base a un volumen ingente de datos, el llamado *Big Data*⁴.

Esta IA basada en datos trabaja en dos fases: la primera de aprendizaje o entrenamiento y la segunda, de uso (en algunos casos, ambas fases interactúan de manera que los sistemas continúan aprendiendo durante su uso). La primera es altamente costosa y laboriosa e implica algoritmos que requieren una gran potencia de cálculo, mientras que la segunda puede ejecutarse en ordenadores personales o teléfonos móviles y es eficiente y rápida. La fase de aprendizaje es invisible, lenta y desconocida, mientras que la de uso puede parecerse casi trivial si olvidamos el volumen de trabajo preparatorio de entrenamiento.

La estructura de las redes neuronales de los sistemas actuales de IA basados en datos se asemeja a una hoja de cálculo gigante, en la que las neuronas se corresponden con las celdas de la tabla y las capas de la red neuronal con sus columnas. En ella, y en la fase de uso del sistema de IA, cada celda se calcula a partir de una combinación de los valores de las celdas de la columna anterior según los valores de ciertos parámetros que configuran la fórmula asociada a la celda⁵. En todo caso, y como

⁴ Ramon López de Mántaras y Pere Brunet, «¿Qué es la inteligencia artificial?», en revista *Papeles*, num. 164, *¿Quién teme a la Inteligencia Artificial?*, FUHEM, 2023:

<https://www.fuhem.es/2024/01/29/papeles-164-quien-teme-a-la-inteligencia-artificial/>

⁵ Véase, por ejemplo, Pere Brunet, Tica Font y Joaquín Rodríguez, *Robots Asesinos: 18 preguntas y respuestas*, Centro Delàs de Estudios para la Paz, 2021, páginas 02.1 y 02.2. Disponible en: https://centredelas.org/wp-content/uploads/2021/12/RobotsAsesinos_18PreguntasYRespuestas_DEF.pdf Ver también: <https://centredelas.org/robots-asesinos-18-preguntas-y-respuestas/?lang=es>, Cada neurona (celda) de la red neuronal pondera los

veremos más adelante, el número de parámetros es ingente, por lo que debemos considerar que las redes neuronales de la IA son sistemas masivamente heurísticos. La fase de entrenamiento utiliza algoritmos que van ajustando los parámetros de manera que la red neuronal se comporte de manera cercana a la esperada cuando trabaja con los datos que la van entrenando. Pero el aprendizaje es esencialmente imperfecto, porque los datos de entreno no pueden predecir los futuros datos que tendrá que tratar el sistema de IA durante su uso, y porque el ajuste de los miles de millones de parámetros de su red heurística sólo puede abordarse de manera muy aproximada.

Una vez entrenada, la red neuronal del sistema de IA se convierte en una gran tabla de neuronas con parámetros para obtener sus valores en función de las neuronas de la capa anterior, y su uso es rápido y eficiente. Solo hay que alimentar la primera capa con la información de entrada (una imagen, un texto, una melodía o lo que sea) codificada numéricamente, transmitir la información a través de las capas de neuronas, y obtener la respuesta del sistema en las neuronas de la capa final. Durante esta fase de uso, no tiene sentido hablar de algoritmos: el sistema de IA trata los problemas aplicándoles algo muy similar a una inmensa hoja de cálculo⁶.

valores de las neuronas de la capa (columna) anterior mediante una fórmula del tipo $F(p_1*c_1 + p_2*c_2 + \dots + p_n*c_n)$ donde los P_k son parámetros y los c_k son valores de las neuronas de la capa anterior, y en la que la función no lineal $F()$ evita la saturación de valores a la vez que mantiene la contribución individualizada de cada uno de los parámetros. La red neuronal contiene por tanto tantos parámetros como relaciones entre neuronas.

⁶ En este sentido, más que hablar de los algoritmos o “del algoritmo” de la IA, podría tener más sentido referirse a redes entrenadas.

Uno de los primeros desarrollos sorprendentes de esta IA basada en datos nos llegó de la mano de los traductores automáticos. El investigador Franz Josef Och fue pionero en diseñar los primeros algoritmos en 2003, y luego Google los incorporó entre 2005 y 2007⁷. Estos traductores, como ya hemos mencionado, funcionan tras aprender de ingentes cantidades de datos, presentando actualmente una fiabilidad bastante aceptable y en todo caso mucho mayor que la de los sistemas anteriores, basados en el análisis sintáctico de los textos.

IV. CARACTERÍSTICAS Y LIMITACIONES DE LA IA

Muchos de los artilugios y aplicaciones que usamos a diario incluyen sistemas de IA⁸. Los encontramos en las cámaras digitales de los teléfonos, en las predicciones del tiempo y en los sistemas de diagnóstico precoz en medicina. Pero también los podemos encontrar en la predicción del plegado de las proteínas y en la robótica moderna, pasando por campos tan diversos como los juegos por ordenador, la previsión del impacto del cambio climático, los coches autónomos o incluso el juego del tenis. Aunque limitados y exentos de una fiabilidad absoluta, los actuales sistemas de IA nos son útiles en estos y otros muchos campos. Sus redes neuronales, inspirados en las

⁷ Franz Josef Och, «Statistical Machine Translation: From Single-Word Models to Alignment Templates», Technical Report, RWTH Aachen, Department of Computer Science, 2023:

<http://www-i6.informatik.rwth-aachen.de/publications/download/520/OchF.J.-StatisticalMachineTranslationFromSingle-WordModelstoAlignmentTemplates--2002.pdf> - Ver también su presentación de 2005, ya como empleado de Google: «Machine Translation», Summit 2005, Phuket: <http://www.mt-archive.info/MTS-2005-Och.pdf>

⁸ Ramon López de Mántaras, *100 cosas que cal saber sobre intel·ligència artificial*, (en catalán), Cossetània, Barcelona, 2023.

capas de neuronas de nuestro cerebro, nos facilitan la solución de problemas a la vez que han heredado un defecto muy propio de los humanos: se equivocan. Es algo que iremos analizando en lo que sigue y que debe tenerse muy en cuenta cuando se plantea su uso.

Las aplicaciones basadas en IA pueden clasificarse según dos ejes, el que va de los usos no críticos a los que sí lo son, y el que se mueve entre el uso de la IA en casos concretos y su utilización de manera generalizada con resultados que acaban siendo útiles en promedio. En los casos de aplicaciones no críticas en situaciones concretas (sistemas para la mejora de fotos, por ejemplo) es aconsejable que alguna persona revise los resultados, algo que se conoce como “post-supervisión”. Pero en el caso de aplicaciones críticas en situaciones concretas con resultados que en algunos casos puedan afectar a personas determinadas, la post-supervisión de los resultados se hace imprescindible desde un punto de vista ético⁹. Ello permite avalar la respuesta del sistema de IA y garantizar que, en caso necesario, alguien podrá rendir cuentas.

En este sentido, podemos distinguir diversos tipos de sistemas de IA¹⁰:

- Los que conforman las aplicaciones no críticas para casos concretos, con post-supervisión: se trata de sistemas para usos específicos, que requieren que la persona interesada revise los resultados y corrija

⁹ Esta post-supervisión de los resultados, imprescindible desde un punto de vista ético, es de desear que pase a tener, mediante normativas regulatorias, validez legal. Véase el apartado sobre regulación.

¹⁰ Pere Brunet, Tica Font y Joaquín Rodríguez, 2021, 04.1 – 04.2, *Op. cit.*

posibles errores. Un ejemplo típico serían los sistemas de traducción automática, ciertamente útiles cuando hacemos una revisión sistemática del texto traducido. Podemos hablar también de los sistemas que proponen posibles soluciones (interpretaciones de manuscritos, jeroglíficos o incluso sonidos de animales, contraejemplos en conjeturas matemáticas, posibles compuestos farmacológicos y proteínas) que luego los expertos se encargan de validar o descartar.

- Los de las aplicaciones no críticas que funcionan en promedio y sin post-supervisión: podríamos citar el ejemplo de los sistemas publicitarios que nos envían mensajes con recomendaciones personalizadas de todo tipo, basadas en nuestro comportamiento anterior. Son sistemas éticamente discutibles, invasivos para las personas, pero que pueden ser beneficiosos para las empresas anunciantes: aunque no incluyan post-supervisión, sus errores en casos concretos no impiden que, en promedio, estos sistemas acaben generando un aumento global de ventas.
- Los sistemas que vemos en aplicaciones críticas y casos concretos, con post-supervisión, un ejemplo de los cuales serían los sistemas de ayuda al diagnóstico médico, además de los sistemas militares de combate que trataremos más adelante. En este y otros casos similares existe consenso entre los expertos en la necesidad de post-supervisión, de manera que el diagnóstico final sea responsabilidad una persona que asume la posibilidad de rendición de cuentas por su decisión.

- Los sistemas que se utilizan en casos concretos y sin post-supervisión: Pueden encontrarse en aplicaciones que no son críticas, como los juegos por ordenador, o en sistemas autónomos para usos críticos, como es el caso, éticamente discutible, de las armas autónomas (apartado 5).
- Finalmente tenemos los sistemas de las aplicaciones críticas que funcionan en promedio y sin post-supervisión: En este caso podríamos mencionar diversos sistemas de control social, así como los métodos de propaganda electoral fragmentada y personalizada para enviar mensajes electorales a medida, en concordancia con los intereses concretos de cada persona y con lo que ésta desearía que le prometiesen. Como en el caso de los mensajes publicitarios, pueden ser efectivos a pesar de sus errores porque consiguen más votos de los que hacen perder. Pero son ética y democráticamente inaceptables, difundiendo mensajes y promesas imposibles de cumplir por el hecho de ser inconsistentes e incompatibles entre ellos mismos.

Los sistemas de Inteligencia Artificial con post-supervisión pueden ser útiles en muchos casos, como ayuda a la toma de decisiones de los expertos y en nuestra vida diaria (sistemas de reconocimiento de voz y música, sistemas aconsejan determinados productos, etc.). Pero su post-supervisión por parte de humanos es esencial. Por lo que respecta a las aplicaciones que funcionan en promedio y sin post-supervisión (críticas o no) deberían ser objeto de una regulación que garantice los derechos de todas las personas.

En todo caso, debemos tener en cuenta que,

- El aprendizaje en los sistemas de IA es imperfecto y sesgado, ya que es imposible disponer de conjuntos de datos totalmente imparciales.
- A menudo, los datos de entrenamiento son privados y se acaban usando de manera ilícita¹¹, sin permiso ni consentimiento de las personas afectadas.
- Por muy grande que sea el volumen de datos usado durante su entrenamiento, éste será siempre parcial porque los datos nunca podrán explicar los acontecimientos imprevisibles del futuro.
- El entrenamiento de la IA va consolidando el aprendizaje mientras ajusta los parámetros de su red neuronal, pero tiene una limitación esencial: el sistema no puede desaprender ni olvidar.
- Los sistemas de IA se equivocan. El porcentaje de error depende del tipo de problema, de la calidad de los datos de aprendizaje, de la estructura de la red neuronal y de la calidad del proceso de entrenamiento, pero nunca es nulo.
- Los errores, en el caso de usos sociales de la IA, a menudo acaban perjudicando a los más débiles¹², como analizó Virginia Eubanks.

¹¹ Naomi Klein, «El gran robo de la IA», *The Guardian*, 2023: <https://www.theguardian.com/commentisfree/2023/may/08/ai-machines-hallucinating-naomi-klein>

¹² Virginia Eubanks, *Automating Inequality*, St. Martin's Press, 2018: <https://virginia-eubanks.com/>

- Los resultados de la IA no son explicables, ni por parte de los expertos¹³. Al no serlo, la rendición de cuentas en el caso de ausencia de post-supervisión se hace casi imposible.
- Los sistemas de IA son inestables, algo que acaba agravando su carácter no explicable porque cambios muy sutiles en la información de entrada que el sistema debe clasificar pueden llevar, inexplicablemente, a decisiones extraordinariamente dispares¹⁴.
- A diferencia de la mente humana, que recibe constantemente informaciones del mundo exterior a través de los sentidos, los sistemas de IA adolecen de falta de referentes de la realidad, ya que la conexión con el mundo real solo pasa por los datos del proceso de aprendizaje, que no son actuales.
- No existe ningún sistema de IA capaz de contextualizar y de hacer el tipo de inferencias básicas que incluso un niño realiza sin esfuerzo¹⁵.

¹³ El comportamiento no explicable es consecuencia de la extraordinaria complejidad de las redes neuronales que conforman estos sistemas, de su carácter heurístico, y de la ingente cantidad de parámetros que los gobiernan. Pero también es consecuencia de la multiplicidad de sesgos que con toda probabilidad contenían los datos de aprendizaje, y de su carácter inestable. Como resultado de estos sesgos y de la enormidad de implicaciones que acaban teniendo en el proceso de ajuste de los parámetros de la red, los sistemas de IA es probable que tomen “decisiones” incorrectas a la vez que extrañas y por ello inexplicables.

¹⁴ Hay muchos ejemplos que han mostrado esta inestabilidad, por ejemplo, en los sistemas de clasificación y detección a partir de imágenes. De hecho, la modificación de unos pocos píxeles en la imagen de entrada, un cambio que es totalmente imperceptible para el ojo humano es suficiente para inestabilizar el sistema, haciendo que salte de una clasificación a otra. Un típico ejemplo muestra una fotografía de ImageNet con un autobús escolar; cuando se distorsionan de manera inapreciable algunos píxeles, el sistema de IA sorprendentemente la clasificó como un avefruz. Véase Brunet, Font y Rodríguez, 2021, p. 05.1, *Op. cit.*

¹⁵ Ramón López de Mántaras, «El traje nuevo de la inteligencia artificial», *Investigación y ciencia*, julio de 2020:

<https://www.investigacionyciencia.es/revistas/investigacion-y-ciencia/una-nueva-era->

- El aprendizaje de los sistemas de IA es muy costoso en términos ecológicos y de energía, como veremos en el apartado 7.
- La IA es una herramienta más, no es la solución a todos los problemas de este mundo que la actual civilización está llevando al desastre global, ecológico, climático y ambiental.

Y a pesar de todos los mensajes que nos llegan sobre las grandes virtudes de la IA, no debemos olvidar sus aspectos controvertidos y su uso en muchos casos de control social. Los verdaderos problemas están en la manipulación, en el uso ilícito de datos privados y en su privacidad, en la vigilancia y el control masivo de la ciudadanía, en la autonomía de sistemas que pueden usarse contra las personas (como las armas autónomas), en la confianza excesiva en las capacidades de la IA, en los sesgos de los algoritmos, en la imposibilidad de rendición de cuentas en el caso de funcionamiento erróneo, y en el excesivo poder que acumulan unas pocas empresas tecnológicas¹⁶.

V. LOS SISTEMAS CRÍTICOS Y LA IA MILITAR

En algunos casos, los posibles errores de la IA son poco preocupantes y fáciles de corregir. Pero en otros pueden afectar a personas concretas en situaciones críticas con efectos graves o letales y sin rendición alguna de cuentas. Como ya hemos comentado, los usos críticos de la IA requieren post-

[para-el-alzheimer-803/el-traje-nuevo-de-la-inteligencia-artificial-18746](#)

¹⁶ Ramon López de Mántaras, 2023, p. 52, *Op. cit.*

supervisión. Esto incluye por ejemplo los sistemas de ayuda al diagnóstico en medicina, a las decisiones jurídicas o a la priorización de personas candidatas a empleos, servicios o subsidios.

Los sistemas de control social son a menudo discutibles, carecen de las necesarias garantías jurídicas y no incluyen una adecuada post-supervisión. Son por tanto usos críticos que requieren una especial atención. Los encontramos en los sistemas de reconocimiento facial para el control policial, en las técnicas de comprobación de personas en zonas fronterizas y sensibles o en manifestaciones y actos públicos. Dado que se trata de usos críticos que afectan a personas concretas, la post-supervisión es imprescindible. En todo caso, los activistas han encontrado maneras de burlar estos sistemas de reconocimiento facial explotando la inestabilidad ya comentada de la IA: se pintan la cara con marcas para convertirse en invisibles para las redes neuronales de la IA¹⁷. Curiosamente, unas marcas faciales que no afectan en absoluto a nuestro sistema perceptivo (continuamos reconociendo perfectamente a estas personas) llegan al talón de Aquiles de la IA y perturban de manera absoluta sus resultados.

Los sistemas de IA pueden ser también una herramienta de guerra que elimina combatientes junto con civiles, mujeres y niños, algo que sabemos que Israel ha usado en Gaza en 2024 gracias al periodista Yuval Abraham¹⁸. En su reportaje explica

¹⁷ Ver por ejemplo: https://i-d.vice.com/en_uk/article/jge5jg/dazzle-club-surveillance-activists-makeup-marches-london-interview

¹⁸ Yuval Abraham, «'Lavender': The AI machine directing Israel's bombing spree in Gaza», revista '972, 2024: <https://www.972mag.com/lavender-ai-israeli-army-gaza/>

que un oficial de inteligencia de Israel decía en la primavera de 2024 que «no debemos desperdiciar bombas caras con personas sin importancia, porque es muy caro para el país y las bombas son escasas». Otro afirmaba que en la guerra no hay tiempo para analizar bien todos los objetivos, y que por tanto hay que arriesgarse a atacar [gente inocente] por error, a los daños colaterales y a matar a civiles. Yuval Abraham detalla los sistemas de inteligencia artificial que Israel ha ido perfeccionando desde 2021 y que utiliza para detectar y localizar a presuntos miembros de Hamás en Gaza, sistemas que se basan como mínimo en tres componentes: el sistema «Lavanda» (*Lavender* en inglés), el «Evangelio» (*Gospel / Habsora*) y el «Dónde está el Papa» (*Where's Daddy*). El primero parte de información personal y datos sobre la actividad en internet y redes sociales de los más de dos millones de palestinos que viven en Gaza y acaba asignando a la mayoría de ellos un índice personal que estima la probabilidad de que sea un combatiente, con un valor entre 0 y 100. El objetivo del segundo, que utiliza, entre otros, información sobre llamadas de móvil, imágenes de satélite y de drones es el de detectar y marcar edificios sospechosos. El tercero parte de la lista de sospechosos dada por el sistema «Lavanda» y envía una alerta cuando la persona seleccionada entra por la puerta de su casa, dando luz verde para bombardear la casa. Los operadores militares analizan los resultados dados por estos sistemas y en unos 20 segundos deciden si se bombardea el edificio indicado, destruyendo y matando a las "personas sin importancia" que en él se encontraban. Se trata de un ejemplo macabro de sistemas de combate basados en IA y del tipo *on*

the loop, inaceptables desde un punto de vista ético¹⁹. Sistemas que están siendo probados “en combate” en las actuales guerras para luego ser exportados y vendidos a otras zonas en conflicto.

Por otra parte, los sistemas de combate militar basados en IA son cada vez más habituales en los campos de batalla. Los encontramos en misiles, drones, robots de combate y todo tipo de armas²⁰. Las razones para su uso se basan en que la sustitución de personal militar por máquinas conlleva una reducción de bajas propias, y en la reducción de costes logísticos asociados al despliegue de tropas. En definitiva, se supone que llevan a guerras menos caras y de menor riesgo, con menores sacrificios humanos propios y sin riesgo para los políticos, que esperan encuestas de opinión más favorables. Ello ha llevado a hablar de la «guerra asimétrica», que aparece entre ejércitos o fuerzas combatientes que tienen un profundo “gap” tecnológico entre ellas. Guerras sin víctimas militares propias, pero que evidentemente desprecian las vidas humanas de “los Otros” y los deshumanizan.

Tenemos finalmente los sistemas de combate autónomos. El armamento autónomo puede tomar forma de robot de combate, misil, dron u otros. Pero la clave reside en que este

¹⁹ Los sistemas *on the loop* requieren la aprobación de un operador militar, pero éste debe responder y actuar dentro de un período de tiempo que limitado (en el caso comentado de los sistemas de IA usados en Gaza en 2024, para evitar que el presunto combatiente vuelva a salir del edificio indicado). El principal problema ético de los sistemas *on the loop* es que sufren del llamado “sesgo de automatización”, que consiste en la tendencia natural que tenemos los humanos a dar por válidas las propuestas que nos llegan desde los sistemas automatizados y las máquinas.

²⁰ Pere Brunet, Tica Font y Joaquín Rodríguez, 2021, *Op. cit* (nota 4).

puede operar sin controladores humanos involucrados en el ciclo de acción del arma. Su sistema de IA detecta el objetivo, decide proceder a su destrucción, lo sigue y finalmente lo ataca. Todo ello, sin intervención humana. Estos sistemas autónomos están empezando a proliferar en diversos conflictos armados.

La aparición de sistemas armados autónomos es ética y jurídicamente inaceptable porque delegar en una máquina las decisiones de matar va en contra de la dignidad humana y de los derechos de las personas. Hay que situar el concepto de dignidad humana como límite insalvable, marcando una línea roja en la que la autonomía en los sistemas de armas deja de ser aceptable. Las armas y los drones letales autónomos no podrán tomar decisiones éticas complejas en campos de batalla dinámicos, ni podrán distinguir adecuadamente entre soldados y civiles, y tampoco podrán evaluar el grado de proporcionalidad de un ataque.

VI. LA IA GENERATIVA

Durante la década de 2010, los sistemas de IA trabajaban a partir de determinados datos de entrada, procesándolos a través de sus redes neuronales para generar un resultado que era de tamaño similar o menor²¹.

²¹ Durante el uso de los sistemas de IA, tanto la información de entrada como la salida es numérica, y su tamaño puede medirse en bits. Las imágenes que deseamos procesar o los textos que queremos traducir se codifican numéricamente, se copian como datos de entrada a las neuronas de la primera capa de la red neuronal, se van transmitiendo por las capas de la red a la vez que los parámetros de las conexiones entre neuronas los van transformando, y el resultado, disponible en la última capa de salida de la

Los inicios de la década de 2020 nos han traído en cambio un nuevo tipo de IA, la generativa. Su principal característica es que el tamaño del resultado es mucho mayor que el de los datos de entrada, que se denominan *prompts*, usando habitualmente este término inglés que podríamos traducir por *sugerencias*, y que consisten habitualmente en textos cortos²². A partir de este texto, la red neuronal de la IA da una respuesta en base a todos los ajustes de aprendizaje que ha experimentado anteriormente.

Las aplicaciones de la IA generativa son muy diversas, yendo desde la ayuda en la generación de textos (periodísticos, escolares, informes e instancias, narraciones, descripciones, epistolares y otros), en la creación de programas informáticos, en la invención de imágenes y material gráfico e incluso en la creación de vídeos y animaciones. La invención de imágenes y falsas fotografías con métodos de IA generativa se denomina *promptografía*, y se hizo viral en 2023 gracias a Boris Eldagsen, que presentó una de estas imágenes a un concurso fotográfico y tras ganar un premio renunció al mismo sacando a la luz el engaño y denunciando la facilidad con que la *promptografía* podía suplantar fotografías reales²³.

red, es también numérico. Gracias a una descodificación final, el resultado vuelve a ser interpretable por nosotros.

²² Como ejemplos de *prompts* podríamos hablar de «escribeme una página sobre la guerra de Troya» o «dame un poema de amor de no más de 20 líneas».

²³ Boris Eldagsen consiguió confundir y engañar al jurado, que le otorgó el premio considerando que se trataba de una fotografía real. En una declaración posterior renunció al premio, explicando que en realidad se trataba de una *promptografía*. Véase la noticia en *The Guardian* (2023): <https://www.theguardian.com/artanddesign/2023/apr/18/ai-threat-boris-eldagsen-fake-photo-duped-sony-judges-hits-back>

Los sistemas de IA generativa heredan los problemas y limitaciones ya indicados en la IA clásica, pero en algunos aspectos, estos se amplifican. Su entrenamiento es altamente costoso en tiempo de cálculo, energía y requerimientos de refrigeración, se utilizan grandes cantidades de información privada sin consentimiento, no contextualizan, no pueden evitar los sesgos y requieren una post-supervisión para revisar el resultado y corregir errores. Puede ser discutible su uso como elemento de ayuda para la preparación de textos o programas informáticos, pero su empleo para la producción de escritos finales es reprobable, ya que el documento o bien queda como anónimo y sin autoría, o bien la persona que ha usado la IA se atribuye indebidamente esta autoría.

El debate ético entorno a la IA generativa, necesario y esencial, está ya generando deserciones laborales. Jan Leike, investigador clave en materia de seguridad en OpenAI y codirector con la responsabilidad de garantizar que los sistemas de inteligencia artificial potentes se ajustasen a los valores y objetivos humanos, dejó la empresa en 2024 afirmando que OpenAI debería invertir más recursos en cuestiones como el impacto social, la confidencialidad y la seguridad en su próxima generación de modelos. Le preocupaba que la empresa no estuviera haciendo pasos para lograrlo y que en cambio se centrara en la “brillantez” de sus nuevas aplicaciones²⁴.

²⁴ Dan Milmo, *The Guardian*, mayo de 2024:

<https://www.theguardian.com/technology/article/2024/may/18/openai-putting-shiny-products-above-safety-says-departing-researcher>

Noam Chomsky, en un artículo conjunto con Ian Roberts y Jeffrey Watumull de marzo de 2023, hace un análisis preciso y certero de los sistemas generativos de texto como ChatGPT²⁵. Explica que su defecto más grave es que no tienen lo que él califica como la mayor capacidad de cualquier inteligencia: la de decir no solo lo que ocurre o sucedió, sino también lo que no pasa y lo que podría pasar (o no). Y es que la verdadera inteligencia se demuestra en la capacidad de pensar y expresar cosas improbables pero esclarecedoras. Comenta además que ChatGPT «se niega a posicionarse, no alega solo ignorancia sino también falta de inteligencia, y se defiende diciendo que cumple órdenes pasando toda la responsabilidad a sus creadores». Y añade otra reflexión: la IA generativa puede encontrar correlaciones, pero no relaciones causa-efecto.

En septiembre de 2023, Julia Angwin explicaba que la organización NewsGuard había identificado 475 páginas web de noticias e información de “baja calidad” a gran escala²⁶, generadas mediante técnicas de IA generativa. Al cabo de un año, la página web de NewsGuard ya hablaba de 1.006 páginas web de noticias artificiales²⁷. Angwin comentaba además que el entrenamiento del chatbot Llama de la empresa Meta se basa en más de 170.000 libros pirateados, y que una investigación del *Washington Post* había descubierto que el aprendizaje de

²⁵ Noam Chomsky, Ian Roberts y Jeffrey Watumull, *The New York Times*, 2023: <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>

²⁶ Julia Angwin, *The New York Times*, septiembre de 2023: <https://www.nytimes.com/2023/09/23/opinion/ai-internet-lawsuit.html>

²⁷ Página web de NewsGuard: <https://www.newsguardrealitycheck.com/p/biden-is-dead-hoax-lives-on>

ChatGPT se ha realizado a partir de datos obtenidos de cientos de miles de páginas web sin consentimiento alguno.

VII. EL COSTE ECOLÓGICO DE LA IA

Hasta hace pocos años, uno de los secretos bien guardados por las empresas del campo de la IA era el de su consumo en términos de energía y agua. Se hablaba de las grandes virtudes de la IA, silenciando al mismo tiempo su huella ambiental.

Las primeras advertencias llegaron alrededor del año 2020, con trabajos como el que Strubell, Ganesh y McCallum presentaron en el congreso internacional sobre inteligencia artificial de aquel año²⁸. Asimismo, David Patterson y otros, en un artículo riguroso fruto de la colaboración entre expertos de Google e investigadores de la Universidades de Berkeley y Cornell, indicaban que la información sobre la energía requerida por los sistemas de inteligencia artificial no se publicaba. Tampoco aparecía ni en los informes de las grandes empresas tecnológicas del sector ni en los de las administraciones. Comentaban que los modelos de IA van siendo más precisos y más capaces, pero que estos modelos, más y más complejos, implican un mayor esfuerzo de computación y en consecuencia, mayores demandas de energía²⁹. En su trabajo analizaron con

²⁸ Emma Strubell, Ananya Ganesh y Andrew McCallum, «Energy and policy considerations for modern deep learning research», en los Proceedings del congreso AAAI sobre inteligencia artificial, abril de 2020 (Vol. 34, No. 09, pp. 13693-13696): <https://ojs.aaai.org/index.php/AAAI/article/view/7123>

²⁹ David Patterson y sus colaboradores se centraron en los sistemas de procesamiento del lenguaje natural (PLN) por su importancia en los productos de Google y por el reciente desarrollo de modelos de PLN de gran tamaño. Véase el artículo «Carbon Emissions and Large Neural Network Training» (Universidad de Cornell, 2021) en: <https://arxiv.org/pdf/2104.10350> y <https://arxiv.org/abs/2104.10350>

detalle los requerimientos energéticos y la huella de carbono de los sistemas T5, Meena, GShard, Switch Transformer y GPT-3.

En pocos años hemos pasado de la IA generativa de sistemas como GPT-3 a las redes neuronales de GPT-4, con un número de parámetros que se ha multiplicado por un factor del orden de 570 respecto a los de su antecesor GPT-3³⁰. La consecuencia es que los procesos de aprendizaje, que deben ir ajustando estos parámetros en base a ingentes cantidades de datos, están aumentando su complejidad de manera exponencial.

En todo caso, el coste ecológico de la IA, consecuencia directa de su impacto ambiental en términos de requerimientos de energía y agua, puede desglosarse en tres partes diferenciadas:

- El coste asociado a los centros de datos que almacenan la información “en la nube” de personas, de las aplicaciones móviles que estas utilizan, y de todo tipo de empresas, corporaciones y administraciones. Estos centros se componen de ordenadores (servidores) conectados a la red, con discos de almacenamiento que funcionan de manera ininterrumpida, junto con sistemas automáticos para preservar la información mediante copias de seguridad.
- El coste de la ejecución de los algoritmos asociados al proceso de aprendizaje de los propios sistemas de IA, que se alimenta constantemente de

³⁰ Chris Stokel-Walker, agosto de 2023:

<https://www.theguardian.com/technology/2023/aug/01/techscope-environment-cost-ai-artificial-intelligence>

grandes cantidades de datos mientras va ajustando los valores de los parámetros de las redes neuronales de los sistemas de IA. Este mecanismo de aprendizaje, todavía muy ineficiente como veremos más adelante, requiere cantidades asombrosas de potencia de cálculo, con sistemas basados en placas gráficas (las llamadas GPUs³¹) que necesitan una continua refrigeración, consumiendo cantidades colosales de electricidad y agua³². Cabe recordar, además, que cada sistema de IA requiere un aprendizaje específico y adaptado a su objetivo concreto. El entreno de un sistema diseñado para sugerir piezas de música no tiene nada que ver con el de un sistema de reconocimiento facial, por ejemplo.

- El coste de la utilización de los sistemas de IA, cada vez que los usamos para hacer consultas, pedir que genere un texto, traducir, mejorar fotos o cualquier otro uso.

La gran complejidad y coste de los procesos de aprendizaje de los sistemas de IA, en particular la de los generativos, tiene que ver con la gigantesca sofisticación de sus redes neuronales. GPT-4 incluye ocho redes neuronales con 220.000 millones de parámetros cada una³³, lo que nos lleva a un total de 1,76 billones de parámetros³⁴ que el sistema de aprendizaje debe ir modificando y optimizando individualmente. Todo ello implica

³¹ Curiosamente, los procesos de aprendizaje de los sistemas de IA utilizan las mismas placas gráficas (GPUs) que se diseñaron y construyeron para la industria de los juegos por ordenador.

³² John Naughton, «The carbon footprint of server farms used to power generative AI could be problematic», *The Guardian*, diciembre de 2023: <https://www.theguardian.com/commentisfree/2023/dec/23/ai-chat-gpt-environmental-impact-energy-carbon-intensive-technology>

³³ Chris Stokel-Walker, 2003, *Op. cit.*

un aprendizaje que acaba siendo inevitablemente imperfecto, con algoritmos altamente lentos y complejos, excesivo consumo energético y un sistema final que ineludiblemente tendrá comportamientos inexplicables.

Volviendo a los centros de datos, que anteriormente a la era de la IA consumían el 1% de la energía eléctrica mundial, se prevé que lleguen hasta un consumo del 3,5% en 2030³⁵. Y ya en 2026, su consumo se considera que será de 114 Giga vatios (unos 1000 Teravatios hora), equivalente al consumo de Japón³⁶. Concretándonos a los Estados Unidos, el grupo de inversión Goldman Sachs predice que los centros de datos representaran el 8% de su consumo energético en 2030 frente al 3% actual³⁷, con un consumo total en 2030 que se cifraría en 35 Giga vatios (en 2022 fue de 17GW³⁸).

Sasha Luccioni explica que el uso conjunto de energía por parte de las principales empresas tecnológicas que ofrecen servicios y productos en la nube con inteligencia artificial (Google, Microsoft, Meta, Amazon y otras) se ha más que duplicado en

³⁴ Este valor es la suma de todas las conexiones de las celdas (neuronas digitales) de la red neuronal del sistema de IA entre ellas y las de la capa anterior. En otras palabras, es el número total de conexiones inter-neuronales de la red.

³⁵ Chris Stokel-Walker, 2003, *Op. cit.*

³⁶ Dan Milmo, *The Guardian*, 2024:

<https://www.theguardian.com/technology/article/2024/jul/02/google-ai-emissions>

³⁷ David Gelles, *The New York Times*, 2024:

<https://www.nytimes.com/2024/07/11/climate/artificial-intelligence-energy-usage.html>

³⁸ Según datos de Newmark, 15 de enero de 2024:

<https://www.datacenterdynamics.com/en/news/us-data-center-power-consumption/>

los últimos cinco años. Estas empresas se encuentran entre los mayores compradores de energía renovable en los Estados Unidos, habiendo comprado casi 50 GW de esta energía hasta la fecha, cifra que equivale a la capacidad de generación de Suecia³⁹. También comenta que la IA, que ahora consume alrededor del 10-20% de la electricidad en los centros de datos, según las previsiones llegará a consumir el doble de la cifra actual en 2030. Luccioni incluye en sus análisis todas las etapas del ciclo de vida de la IA. Además del proceso de entrenamiento, su implementación y los componentes que la hacen posible, considera tanto la extracción de materiales y la fabricación de los equipos como el uso de energía en la refrigeración, redes y almacenamiento.

Por otra parte, debemos citar los trabajos del TOI Tech Desk, un equipo de periodistas comprometidos que desde la India publican noticias recientes y relevantes del mundo de la tecnología. Recientemente han indicado que la energía requerida por la IA está creciendo exponencialmente a razón de entre el 26 y el 36% anual. Por ello, consideran que la potencia informática necesaria para impulsar la IA en 2028 igualará el consumo energético total de Islandia en 2021⁴⁰.

Y además del consumo energético, la refrigeración de los sistemas de aprendizaje de la IA requiere un consumo muy significativo de agua, que en 2027 podría llegar a una cifra de

³⁹ Sasha Luccioni, 2024: <https://huggingface.co/blog/sasha/ai-environment-primer>

⁴⁰ TOI Tech Desk, 2024: <https://timesofindia.indiatimes.com/technology/tech-news/chatgpt-google-search-need-power-to-run-heres-how-much-water-and-electricity-are-used-to-answer-questions/articleshow/111382705.cms>

entre 4.200 y 6.600 millones de metros cúbicos, casi la mitad de lo que gasta anualmente el Reino Unido⁴¹.

Las consultas a los sistemas de IA generativa tienen aparentemente un coste energético moderado, ya que la energía involucrada en una consulta a Chat GPT podría cifrarse en unos 2,9 vatios hora. Pero debemos observar que la energía involucrada en una de estas consultas oculta un gasto energético que puede ser del orden de 204 millones de veces más elevado⁴²: el del proceso previo de aprendizaje del sistema de IA que estamos usando. Según un informe del *Economic Times*, un año de uso de ChatGPT llega a consumir la energía que utiliza Nueva Zelanda en tres meses⁴³.

Se calcula que en 2022, los sistemas de almacenamiento en la nube de Amazon (AWS), Alphabet, Meta y Microsoft consumieron 90 teravatios-hora (TWh) de electricidad, lo mismo que todo el estado de Colombia. Y la situación actual está llevando incluso a que las grandes tecnológicas se planteen recurrir a la energía nuclear para alimentar sus sistemas de IA. En marzo de 2024, AWS pagó 650 millones de dólares por un centro de datos de 960 megavatios en Pensilvania con energía suministrada por un reactor nuclear. Asimismo, Microsoft ha llegado a un acuerdo con Constellation Energy para el suministro de energía nuclear a su centro de datos de Virginia, como refuerzo en caso en los casos en que las energías convencionales y renovables sean insuficientes⁴⁴. Todo ello

⁴¹ TOI Tech Desk (2024), *Op. cit.*

⁴² John Naughton, 2023, *Op. cit.*

⁴³ TOI Tech Desk, 2024, *Op. cit.*

⁴⁴ *La Vanguardia*, mayo de 2024:

<https://www.lavanguardia.com/vida/20240508/9619416/grandes->

mientras que Sam Altman⁴⁵ de OpenAI propone la energía solar y la fusión nuclear como posibles soluciones al problema de la energía que necesitarán los futuros sistemas de IA⁴⁶.

Las emisiones de Google el pasado año fueron un 48% más altas que en 2019, en gran parte debido a los centros de datos y al auge de la IA. Las emisiones de Microsoft también se dispararon, por las mismas razones, un 29% el año pasado respecto a 2020. Y las emisiones de Meta aumentaron un 66% de 2021 a 2023. Mientras tanto, Microsoft y OpenAI planean construir un centro de datos de 100.000 millones de dólares, según varios informes que sugieren que podría requerir cinco Giga vatios de energía, lo que aproximadamente equivale a cinco reactores nucleares⁴⁷.

Como indica Chris Stokel-Walker citando a Sasha Luccioni⁴⁸, hay un gran sinsentido en el hecho de que se estén usando métodos basados en IA para hacer un seguimiento de los procesos de deforestación, cuando vemos que las emisiones de los sistemas de IA acaban contribuyendo al calentamiento global y al aumento de incendios forestales.

Podemos concluir que el coste energético y ambiental de todo el ciclo de vida de los sistemas de IA, en términos de energía eléctrica y agua e incluyendo tanto los centros de datos como los sistemas de aprendizaje y el uso final de sus redes

[tecnologicas-recurren-energia-nuclear-alimentar-inteligencia-artificial.html](#)

⁴⁵ OpenAI es la empresa creadora de ChatGPT.

⁴⁶ TOI Tech Desk, 2024, *Op. cit.*

⁴⁷ David Gelles, 2024, *Op. cit.*

⁴⁸ Chris Stokel-Walker, 2003, *Op. cit.*

neuronales es, sin duda, desmesurado, con una tendencia expansiva que está haciendo saltar muchas alarmas.

Pero, sorprendentemente, los sistemas comerciales de IA tienen un coste monetario aceptable o nulo. El sistema Luzia⁴⁹ es gratuito, al igual que el Chat GPT 4o Mini. Y el precio del sistema Chat GPT 4.0 Plus se mantiene en márgenes asequibles⁵⁰.

En los actuales sistemas de IA vemos un abismo entre el precio que se ofrece al consumidor y su coste total (incluyendo los costes climáticos y ambientales que acabamos de indicar). Es un contrasentido que unos sistemas que suponen un inmenso dispendio energético y que acaban contribuyendo de manera significativa al deterioro ambiental se ofrezcan como servicios gratuitos.

La comparación, en términos de potencia, tamaño y consumo energético, entre los actuales sistemas de IA y el cerebro humano es clarificadora. Nuestro cerebro funciona con una energía de 20 vatios, lo que necesita una bombilla de bajo consumo. Como explica Vijay Balasubramanian, lo fascinante de los cerebros animales es que son muy pequeños y no requieren demasiado entrenamiento:

Las pocas sinapsis que hay entre mi ojo y la región del cerebro donde reconozco tu cara hacen que si vuelvo a verte mañana te reconozca sin necesidad de miles de ejemplos.

⁴⁹ Sistema de IA generativa que ofrece whatsapp

⁵⁰ El precio del sistema generativo Chat GPT 4.0 Plus era de 20 euros al mes en septiembre de 2024.

Ahora mismo las máquinas están superando límites porque vertemos más datos y las hacemos cada vez más potentes, algo que no podemos hacer con nuestros cerebros porque están limitados por el canal del parto.⁵¹

El problema energético y ambiental de la IA nos lo acaba mostrando con toda su crudeza Vijay Balasubramanian. El crecimiento voraz de los sistemas de IA no existiría si sus máquinas tuvieran que cruzar pasillos angostos como los que traspasan los niños al nacer y además tuviesen que consumir una energía similar a la que requiere nuestro cerebro. Habrá que encontrar nuevas maneras de diseñar y construir las siguientes generaciones de sistemas basados en IA, métodos que tengan en cuenta la sostenibilidad y los límites ecológicos y planetarios.

VIII. HACIA UNA IA ECOLÓGICA Y HUMANA: REGULACIÓN Y REDISEÑO

Como indica Sasha Luccioni, en estos momentos estamos inmersos en un gran debate sobre la IA, incluyendo aspectos tanto sobre cómo se debe abordar⁵², como sobre la privacidad de sus datos de entrenamiento, sobre sus efectos sociales y sobre sus consecuencias para el medio ambiente. En este contexto, aflora la importancia indudable de su regulación.

⁵¹ Vijay Balasubramanian, entrevista en Ara.cat (en catalán), 2023: https://www.ara.cat/ciencia-medi-ambient/moltes-coses-intel-ligencia-artificial-son-irritants-inutils_128_4820469.html

⁵² Sasha Luccioni, 2024, *Op. cit.*

La Unión Europea aprobó su ley regulatoria de la IA en diciembre de 2023. La ley clasifica los sistemas de IA en base a su riesgo: sistemas de riesgo inaceptable que deben ser prohibidos, sistemas de alto riesgo que deberán homologarse, sistemas de IA generativa (a los que se les pedirá que informen en cada caso sobre que el contenido ha sido generado mediante IA) y finalmente los sistemas de bajo riesgo o riesgo limitado como los que ya llenan nuestra vida cotidiana. En este último caso la norma exige unos requisitos mínimos de transparencia que permitan a los usuarios tomar decisiones informadas y que les permitan saber, en concreto, si desean seguir usándolas. En resumen, la *AI Act* de la UE dispone que debe garantizarse en todo caso que los usuarios sean conscientes de que interactúan con sistemas de IA^{53, 54}.

El descontrol regulatorio nos ha llevado a sistemas pensados desde el negocio y no desde la gente. Pero las administraciones tienen el deber de proteger y cuidar a sus ciudadanos, y ello implica que la regulación, en vez de realizarse únicamente con la colaboración de las grandes empresas transnacionales, debería ser participativa, democrática y adaptable además a los cambios sociales y tecnológicos. Es en este marco regulatorio

⁵³ La norma europea, *AI Act*, entró en vigor el 1 de agosto de 2024. En el momento de escribir este informe todavía no se aprecian signos de que los dispositivos informáticos y teléfonos informen a sus usuarios del hecho de algunas de las aplicaciones que usan están total o parcialmente basadas en técnicas de IA.

⁵⁴ Pere Brunet, «Regulación de la inteligencia artificial», en *Papeles*, núm. 164, FUHEM, 2023:

<https://www.fuhem.es/2024/01/29/papeles-164-quien-teme-a-la-inteligencia-artificial/>

basado en principios éticos y de respeto a las personas, en el que las empresas deberían ejercer su labor⁵⁵.

Los sistemas de IA nos serán útiles en la medida que sepamos que los estamos usando y que podamos (o debamos, en casos críticos) revisar, supervisar y corregir sus resultados. Sabemos de los problemas asociados a sus inevitables errores, y debemos ser conscientes de su falta de explicabilidad en casos de comportamiento inaceptable. Son cuestiones inherentes a la IA que les acompañarán durante mucho tiempo.

Pero estas técnicas de IA adolecen de otros defectos que requieren soluciones a corto plazo. Temas como el de los sesgos y la privacidad de los datos que se usan para el aprendizaje y sus posibles derechos de autor, porque no es ético valerse de datos personales sin autorización⁵⁶. O bien, la necesaria publicación de la autoría junto con una regulación sobre cómo proceder frente a posibles daños a personas o comunidades debidos a la desinformación, información maliciosa o control ilícito de personas por parte de los sistemas de IA, además de la asunción de responsabilidades en caso de que sea necesaria una rendición de cuentas.

Las personas, como potenciales afectadas por estos sistemas de IA, deben ser los agentes activos de su regulación. En este mundo globalizado, la regulación de la IA no será totalmente efectiva si no es global, debiendo basarse en principios éticos y de respeto a los derechos y a la dignidad de todas las personas. Una regulación que debe prevalecer sobre el negocio.

⁵⁵ *Ibidem*.

⁵⁶ Naomi Klein, 2023, *Op. cit.*

Por otra parte, los sistemas futuros de IA con toda probabilidad serán distintos a los actuales. Por una parte, nos encontraremos con aproximaciones híbridas que combinarán las técnicas actuales de IA basada en datos con técnicas de IA basada en conocimiento, que han demostrado ser muy eficientes en tareas que requieren razonamiento o planificación⁵⁷. Como ya hemos comentado y como bien explica Balasubramanian, habrá que encontrar nuevas maneras de diseñar y construir los futuros sistemas de IA, para que tengan en cuenta la sostenibilidad y los límites ecológicos y planetarios.

La construcción de un nuevo mundo más ecológico, sostenible y humano implica también un rediseño radical de la IA, pensándola desde la ética y las personas y para las personas, la biosfera, el planeta y nuestros descendientes. De manera que pasemos de las aplicaciones pensadas para controlar, manipular y restringir a resolver los problemas reales de la gente y los pueblos. Haciéndola más versátil e independiente de problemas específicos, intentando que adquiriera conocimientos de sentido común y viendo que pueda inferir relaciones de causa-efecto. Pero, además, aumentando su rendimiento energético y de todo tipo de recursos tanto materiales como de información, de manera que mejore su fiabilidad con menos cantidad de datos de aprendizaje, con menos energía y agua, con menos potencia de cálculo.

⁵⁷ Ramon López de Mántaras, 2023, pp. 48-50, *Op. cit.* (nota 13).

Con soluciones políticas y sociales que reduzcan el uso de las nubes de datos y, por tanto, el tamaño de los actuales centros de datos. Con sistemas que pasen del actual control privado al control democrático y participativo. Con una adecuada regulación de los equilibrios entre costes y precios, para evitar el sinsentido actual de las aplicaciones gratuitas que requieren sistemas que implican costes ecológicos desmesurados. Y en base también a los principios feministas. En la línea de las afirmaciones de Steven Pinker, que decía que las distopías sobre la inteligencia artificial provienen de «una psicología parroquial alfa-masculina aplicada al concepto de inteligencia», mientras que debemos comprender que es posible «concebir una IA para las personas, basada en valores femeninos»⁵⁸.

IX. CONCLUSIONES

En las páginas precedentes hemos analizado los actuales sistemas de inteligencia artificial basados en datos, exponiendo brevemente sus características y algunos de sus campos de aplicación. Hemos tratado asimismo de los actuales sistemas de IA generativa, para terminar centrándonos en su huella ecológica, la necesidad de regulación y el camino hacia sistemas de IA basados en criterios éticos y ecológicos.

La IA puede sernos útil si la usamos desde el conocimiento de sus luces y sombras. Debemos ser conscientes de que lo que tenemos es una herramienta tecnológica muy incipiente. Como herramienta, es algo de lo que deberíamos poder disponer para usarlo con responsabilidad cuando lo necesitemos, pero que

⁵⁸ Brunet, 2023, *Op. cit.*

debemos dejar en el cajón de sastre de los artilugios el resto del tiempo. Debemos controlar conscientemente su uso y evitar a toda costa que sea ella la que nos dirija. Su sofisticación no debe llevarnos a pensar que es esencialmente distinta a otros artefactos cotidianos como las tijeras o la nevera, que usamos solo cuando los necesitamos. Pero además, la IA es un conjunto de tecnologías totalmente incipientes. Los sistemas futuros de IA serán forzosamente distintos a los actuales, que podríamos comparar de alguna forma a los primeros aviones de hace un siglo.

Necesitamos un rediseño de la IA que se base en principios de respeto a las personas y al planeta. No todo es válido. Los creadores, diseñadores y tecnólogos deben proceder desde un imperativo ético, pensando no solamente en cómo abordar la creación y construcción de los sistemas futuros, sino sobre si éstos deben ser diseñados, desarrollados y ofrecidos a la sociedad, o si más bien deben ser repensados. Y las personas que queremos usar estos sistemas deberemos plantearnos críticamente si es ético utilizarlos o si lo mejor es dejarlos en el desván de las tecnologías que no debieron existir.

En todo caso, es también imprescindible educar a los ciudadanos –en particular a los jóvenes en las escuelas y universidades y a los políticos– sobre los beneficios y riesgos de estas tecnologías de IA. Los estudiantes de ciencias e ingeniería deberían recibir una formación ética que les permita entender las implicaciones sociales de estas nuevas tecnologías que en algún momento desarrollarán. Y los ciudadanos en general deben exigir estar mejor informados, dotándose de un espíritu crítico que les permita discernir, que les aporte mayor

capacidad para evaluar los riesgos tecnológicos y que lleve a hacer valer sus derechos. Finalmente, las administraciones deberían ser valientes para regular a favor de la gente y visionarias para invertir en una educación que capacite adecuadamente a sus jóvenes y ciudadanos.

Porque quien nos querrá controlar serán personas, no máquinas ni sistemas de IA.

COLECCIÓN
**TIEMPO DE
TRANSICIONES**

