Revista del Instituto Español de Estudios Estratégicos
ISSN-e: 2255-3479

Daniel Terrón Santos
Director de la Cátedra Almirante Martín Granizo USAL-CESEDEN

Correo: datersa@usal.es

Inteligencia artificial y sistemas biométricos de uso militar: redefinir la inteligencia artificial desde la ética y el derecho

Artificial intelligence and biometric systems for military use: redefining artificial intelligence from an ethical and legal perspective

Resumen

Se aborda en este artículo la necesidad de prevenir, sin limitar o haciéndolo lo justo, en aras de la que inteligencia artificial, como herramienta casi plenipotenciaria, aporte su inmenso grano de arena en las aplicaciones de seguridad y defensa. En particular, el uso de sistemas biométricos supone todo un mundo de oportunidades en el sector de la seguridad, pero que, al mismo tiempo, confronta con derecho y garantías irrenunciables. Esta situación de confrontación es preciso resolverla desde la ética y el derecho, sin que en ningún caso suponga determinar un límite absoluto al uso de una herramienta decisiva, con aplicaciones militares básicas e imprescindibles.

Palabras clave

Tecnología, Aplicaciones militares, Biometría, Ética, Control.

Abstract

This article addresses the need for prevention, without limiting or doing it just enough, so that artificial intelligence, as an almost plenipotentiary tool, may make a major contribution to security and defence applications. More specifically, the use of biometric systems opens up a whole world of opportunities in the security sector, but at the same time it clashes with inalienable rights and guarantees. This confrontation must be resolved on the basis of ethics and law, without in any case setting an absolute limit to the use of a decisive tool, with basic and essential military applications.

Keywords

Technology, Military applications, Biometrics, Ethics, Control, Standards.

Citar este artículo:

Terrón Santos, D. (2024). Inteligencia artificial y sistemas biométricos de uso militar: redefinir la inteligencia artificial desde la ética y el derecho. Revista del Instituto Español de Estudios Estratégicos. 24, pp. 183-205.

1 Presentando la inteligencia artificial predictiva

Predecir equivale a poder anticiparse con la finalidad de prevenir aquello que no es deseado o, por el contrario, aprovechar las oportunidades que el conocimiento previo puede suponer obteniendo unos resultados óptimos. A la predicción, siguiendo la definición que del término facilita el DRAE, se puede llegar por revelación, conocimiento fundado, intuición o conjetura. Esto supone, tras eliminar la opción de la intuición, que se parte de indicios u observaciones, que la predicción, en definitiva, es un análisis de datos que se han obtenido de distintas formas. En cierto modo, todo depende del significado, dentro de los que están aceptados, que se le quiera dar al concepto inteligencia. Este concepto es coincidente con el de predicción, por cuanto ambos suponen el análisis de datos para llegar a procurar conocimiento que, toda vez aplicado, puede servir para resolver problemas, para la toma de decisiones, etc.

A estas alturas, casi para nadie es extraño el término inteligencia artificial (en adelante, IA), por lo que todo lo dicho acerca de ese concepto es perfectamente transponible en este momento. Sin embargo, se debe abordar el mismo poniéndolo en relación con el análisis predictivo. Ahora bien, al acercarse al concepto de inteligencia artificial, en particular a ese carácter ajeno al comportamiento humano, se deberán utilizar algoritmos que sean capaces de predecir un comportamiento, sobre todo cuando las variables no permitan, por volumen, la decisión humana o hagan esta ineficaz por el tiempo preciso para lograr una predicción.

Así se podrá hablar de inteligencia artificial predictiva, cuando sea precisamente ese, no otro, el objetivo que perseguir: predecir. En general, la inteligencia artificial implica que sea la máquina la que piense, incluso que esta pueda aprender a pensar, *machine learning*, pero la inteligencia artificial predictiva va más allá. El *machine learning* supone que una máquina pueda aprender a clasificar, agrupar y en especial calcular desarrollando algoritmos formados a partir de un conjunto de datos otorgado antes. Pero el análisis predictivo interpreta y, sobre todo, lleva a la práctica todos los resultados obtenidos —información— a partir de los datos facilitados y ese es el momento preciso donde la predicción se transforma en inteligencia.

La inteligencia artificial predictiva (en adelante, IAP) es un método de análisis de datos que, a diferencia de la IA, permite predecir y anticipar, en la medida en que se ha simulado todo el conjunto de escenarios posibles, a partir de toda la información actual y pasada, recopilada dentro de un determinado ámbito. Esto supone que sin esos datos sería imposible modelar predicciones útiles y efectivas. En ningún caso es posible reducirla a un mero proceso de pronóstico, por cuanto los volúmenes de datos incorporados a la ecuación de pronóstico que permite la generación de predicciones más precisas emplean incluso datos obtenidos en tiempo real. También aquí se puede diferenciar la IA de *software* tradicional, donde se procesan datos con un programa previamente escrito —algoritmo— que genera resultados de los procesos de *machine learning* en los que se genera el programa a partir de los datos y de resultados históricos de esos mismos datos, en una aprendizaje automático y continuo de la máquina. Con ello se logra que lo aprendido se reutilice, proyectando hacia futuro lo que sucederá a

continuación, lo que permite plantear actuaciones concretas y adecuadas para lograr optimizar resultados o lograr objetivos.

El volumen y variedad de datos, como la automatización del aprendizaje, incrementan el grado la precisión, lo que genera un sistema robusto. Entra en juego en este punto un factor, no abordado hasta este momento, pero que tiene un protagonismo innegable en todo este sistema: la calidad del dato. De todo punto es legítimo pensar en la existencia de factores adyacentes que distorsionen los datos, incluso de forma intencionada, por ejemplo, recurriendo a la contrainteligencia. Si bien es cierto que cuantos más datos —internos o externos— se empleen, más precisas serán las predicciones de IAP, la calidad del dato puede distorsionar los resultados, aunque se hayan tomado en consideración y generado escenarios específicos en función de los mismos.

Hasta aquí poco se ha avanzado en el estado de la cuestión primigenia de este trabajo, o al menos así puede parecerlo. Sin embargo, quizá no se está tan lejos como pudiera parecer. De sobra es conocido que, hoy en día, casi cualquier aplicación de seguridad usa sistemas biométricos en la protección de personas e instalaciones. Las fuerzas armadas de todo el mundo emplean la biometría para llevar a cabo controles de acceso seguro a instalaciones, equipos y sistemas de tecnología de la información principalmente. Más avanzada es la tecnología biométrica utilizada para identificar adversarios en el campo de batalla o en terrenos operativos. Incluso la tecnología biométrica de vanguardia excede el ámbito más tradicional de la defensa e interactúa en la esfera de la inteligencia. Luego está claro que la biometría va mucho más allá del que ha sido su objetivo principal: reconocer la identidad de una persona al efecto de prohibir o restringir el acceso de un individuo, dependiendo de si está autorizado, su nivel de autorización o, si no estuviera autorizado, la prohibición en relación con los espacios físicos dentro del área física o de un entorno virtual que esté confinada (Illanas, 2024).

Estos sistemas de biometría se definen por la toma de datos de personas con el objetivo de llevar a cabo un reconocimiento inequívoco, aplicando de manera automática una serie de técnicas sobre los rasgos físicos o de conducta propios de cada persona. Pero, ¿acompaña la legislación al avance de la técnica de los sistemas de biometría? Claramente, la respuesta, como ya viene siendo habitual cuando se trata de tecnología y datos personales, es negativa.

La situación actual podría decirse que es de total inseguridad jurídica y las razones no son otras que la peculiaridad de la materia por tratar, la diversidad de autoridades y «voces» que emiten su opinión y doctrina sobre dicha materia y el imparable avance de la tecnología cada vez más sofisticada e intrusiva con la privacidad de las personas.

La ciencia y la tecnología han puesto sobre la mesa la cuestión de la posibilidad y viabilidad científico-técnica para que, en tiempos de metaverso y avatares[1], no

1 Según la Real Academia Española (RAE 4. m. Inform.), RAE en adelante , un avatar sería la representación gráfica de la identidad virtual de un usuario en entornos digitales. Por su parte, el diccionario Oxford English Dictionary (abreviado OED), 1.1992–Computing (originally Science Fiction), reconoce al metaverso como un término coloquial utilizado para

extrañe, incluso puede parecer hasta obsoleta, la irrupción definitiva de los robots. Pero, más allá de la automatización que supone su recurso a las tareas de carácter físico o mecánico[2], interesa particularmente referirse a la vertiente no mecanizada de estos, la que supone su capacidad de tomar decisiones en lugar de tenerlo que hacer quienes recurren a estos. Luego interesan los robots inteligentes.

Después de tanto tiempo con ellos, casi de repente se ha dado cuenta el ser humado de que el hecho de que un robot pueda «pensar», aunque sea de forma artificial, es responsabilidad del algoritmo, en ningún caso recae esta tarea en la máquina que meramente sirve de soporte, con independencia de su configuración. La historia contempla al algoritmo desde hace tres mil años, lo que precisamente deja claro que no puede haber ausencia de intervención humana en su origen. Es obvio que, en el momento en que un ser humano establece un conjunto de reglas que, aplicadas sistemáticamente a unos datos adecuados, resuelven un problema cierto, se está ante el nacimiento del algoritmo. Así ha transcurrido su devenir por los siglos hasta fechas recientes. Ahora, además, ese algoritmo se incorpora a una herramienta informatizada capaz de asumir un volumen de gestión de la información y tratamiento de datos inasumible para un ser humano.

El desarrollo de la ciencia informática junto con la variable que supone el algoritmo ha terminado por generar la denominada IA. Una vez formulado, el desarrollo tecnológico permite que este algoritmo, cada vez más complejo —consecuencia de la propia circunstancia técnica—, sea capaz por sí mismo, siempre en función de su configuración inicial, incluso de crear otros algoritmos. Si uno se detiene en este estado inicial de la cuestión, está claro que no se está ante un comportamiento ajeno al ser humano y, por tanto, sujeto a reglas determinadas por este. Incluso la posibilidad avanzada implica que el factor humano ha previsto la posibilidad, se ha mostrado conforme, por lo que su voluntad de «transferir» la decisión a la máquina está patente. Lo que se verá más adelante determinará incluso la posibilidad de atribución de responsabilidad al autor material del algoritmo primigenio, por cuanto si bien fue capaz, no evitó la contingencia, lo que deja en manos del algoritmo de manera deliberada decisiones que a él correspondían.

2 Aproximación a la biometría desde la IA

No sería muy arriesgado afirmar que la IA existe por cuanto hay máquinas capaces de percibir su entorno y, en consecuencia, llevar a cabo acciones que maximizan sus

describir una representación de la realidad llevada a cabo mediante programas de realidad virtual, acepción que ha sido validada por la RAE, sin incorporar todavía al DRAE, aceptando que es comprensiva de universos basados en entornos virtuales.

2 Con automatización se hace referencia a aquellos procesos por los que se pueden controlar de manera externa diferentes procesos de producción, gestión de datos, instrumentación, construcción, etc. En definitiva, se trata de la implementación de máquinas, sistemas informáticos, robots o programas de inteligencia a los diferentes campos de producción anteriormente citados.

posibilidades de éxito en algún objetivo o tarea. Sin detenernos en la exhaustividad de un análisis desde la perspectiva del lenguaje, el término «inteligencia artificial» se emplea para referirse a una máquina que imita las funciones cognitivas propias de los humanos, quienes, a su vez, las asocian con otras mentes humanas, como puedan ser percibir o razonar, pero, sobre todo, la capacidad para aprender y resolver problemas. Un concepto más elaborado la contemplaría como aquella capacidad reconocida a un sistema que le permite interpretar de forma correcta datos externos, aprendiendo de dichos datos y empleando esos conocimientos adquiridos para realizar tareas y alcanzar metas concretas a través de la adaptación flexible. Es esta última la nota característica de la IA (Kaplan y Haenlein, 2018).

Que las máquinas se vuelvan cada vez más capaces supone que tecnología que alguna vez se pensó que requería de inteligencia quedaría fuera de la definición, de igual modo que existen distintos tipos de percepciones y acciones que pueden ser obtenidas y producidas, respectivamente, por sensores físicos y sensores mecánicos en máquinas, pulsos eléctricos u ópticos en computadoras, tanto como por entradas y salidas de bits de un *software* y su entorno que, por haber entrado a formar parte de la rutina humana, pudiera parecer que se alejan de la consideración de IA pero, sin embargo, son perfectos ejemplos de la misma. El control de sistemas, la planificación automática, la habilidad de responder a diagnósticos y a consultas de los consumidores y el reconocimiento de escritura, del habla y de patrones están integrados en la realidad del día a día. Campos como la economía, la medicina, la ingeniería, el transporte, las comunicaciones y la defensa, entre otros, no se entienden sin la presencia de la IA. Incluso el ocio se ha visto imbuido de esta a través de distintas aplicaciones de *software* presentes en juegos de estrategia o incluso el propio ajedrez[3].

Luego de la IA no se puede negar su carácter inteligente, pero no hay que errar en compararla con la inteligencia humana, simplemente porque son en esencia distintas, al menos por el momento. Sin ir más lejos, la IA no tiene unos valores fijos ni es unidimensional, lo que supone que no pueda ser objeto de comparación[4]. Sin

3 Es necesario plantearse que los avances hacia modelos de aprendizaje basados en patrones no humanos aleja a la IA del comportamiento que sería propio de un humano. Esto sin duda es una ventaja, pero no carente de problemas éticos por cuanto puede generar resultados que sean inaceptables para el hombre cuando, sin embargo, son óptimos para la máquina. Conocido es el caso del robot AlphaGo Zero, que en 2016 adquirió conocimiento mediante un método conocido como *Reinforement learning*, donde no se necesita aprender a partir de comportamiento humanos, sino que es capaz de generar conocimiento desde cero. Estos patrones de aprendizaje permiten a las máquinas acumular el equivalente a miles de años de conocimiento en cuestión de horas, lo que al caso supuso que AlphaGo Zero fuera capaz de vencer, por cien partidas a cero, al campeón mundial del juego Go. Si se habla de otros de los grandes retos de la IA, como es la conducción autónoma, evitar tomar en consideración el factor humano supondrá que decisiones de IA puedan llegar a ser incompatibles con el pensamiento humano y, por tanto, no ser aceptado por este. Conviene no olvidar que ni la ética ni la moral están en la esencia de la IA, sino que se incorporan a través del aprendizaje basado en comportamiento humano.

4 Resulta curioso que, al hablar de inteligencia, no sea posible aplicar el cociente intelectual (en adelante, CI) a la IA. El CI compara las capacidades de aprendizaje, de comprensión, de formación de conceptos, de análisis de información, de aplicación de lógica y razonamiento de un sujeto en relación con el resto. Tiene carácter informativo, por lo que refleja la influencia de características no comunes que solo en sociedad se pueden determinar. Así, desde la perspectiva de las relaciones se puede conocer la existencia de aislamiento, rechazo o distanciamiento e incluso ponderarlos al tiempo que se pueden apreciar las facilidades o dificultades encontradas en las actividades intelectuales. Esto llevaría a entender que unos logros adquiridos por individuos exigen menor esfuerzo cognitivo que otros, o que los mismos logros no requieren el mismo nivel de esfuerzo para todos los sujetos. Es de todo punto correcto afirmar que el CI en el fondo es una fórmula para estimar

embargo, en el contexto de la IA, por tratarse de sistemas de IA distintos, no cabe comparación entre ellos, como no cabe comparar desde perspectivas o ejes únicos objetos sustancialmente distintos. De igual modo carece de sentido alguno inquirir qué algoritmo es más inteligente ni tampoco lo tiene preguntarse si un objeto es mejor que otro cuando ambos son distintos. El carácter estrecho que se propugna de la IA desvirtúa ese cuestionamiento[5].

El uso del término IA adolece del rigor propio del que es profano en una materia cierta. Su empleo en términos de generalidad no es posible, precisamente porque no hay una solo IA, por lo que solo será correcto expresar la presencia cierta de IA, pero de forma parcializada, en lugar de refrendar que la IA ha realizado esto o aquello.

En este sentido, el uso de la IA en el ámbito militar permite la creación de herramientas capaces de recoger, procesar y almacenar indicadores biométricos, huellas dactilares, escaneos oculares y parámetros faciales, de manera que puede establecer, con la ayuda de bancos de datos y de la IA, indicadores de coincidencia a la hora de identificar a un individuo hostil en potencia. Sin embargo, el sistema debería estar sujeto a supervisión y depender del factor humano (este decidirá qué criterios va a seguir el sistema para identificar sujetos, en qué aspectos centrar su búsqueda y podrá, además, controlar y desactivar el sistema en caso de error.

Si bien los sistemas de IA han convertido aparatos como drones y vehículos no tripulados en unidades autónomas (usadas en especial en el ámbito militar), desde Naciones Unidas y la Unión Europea se recomienda que estén siempre condicionados al factor humano, de manera que no cuenten con el total de la autonomía en la actuación[6]. Además, los sistemas de identificación biométricos actuales incluyen sistemas de IA en conjunto con satélites que albergan sistemas de defensa o preparados para la recogida y almacenamiento de dichos datos biométricos, gracias al uso del *big*

la inteligencia que mayoritariamente no responde a un único factor. De hecho, se utilizan distintos test estandarizados para su determinación, por lo que la medición de la inteligencia es de por sí cuestionable, aunque en todo caso se le debe reconocer un valor predictivo aplicado a determinadas funciones o comportamientos.

5 En términos comprensibles, no es posible comparar un coche con una casa. Se pueden valorar por separado y siempre en comparación con objetos similares, pero entre sí no admiten comparación. De forma individual (ajena a comparación alguna) solo se podría decir si el objeto gusta o no, si se entiende que satisface su empleo, etc. No es que se pretenda huir de la eterna pregunta ¿coche o casa?, simplemente la inteligencia no puede comparar ambos objetos por las diferencias entre ellos. Obviamente, en función de la información acerca de una determinada situación, sí se podría dilucidar si para el caso concreto interesa más adquirir un coche o comprar una casa, incluso se podría determinar el sistema de compra idóneo y, solo desde esa perspectiva, se determinaría cuál es idóneo, no mejor. Esto se acentúa más en el caso de la IA, ya que su carácter estrecho deriva de la propia capacidad para resolver un problema por parte de esta. Esto lleva a afirmar que una determinada IA sea capaz de hallar la solución a un problema no presupone que tenga capacidad para resolver otro problema diferente. Solo sometiendo el nuevo problema a su consideración se podría averiguar esta capacidad, sin que el resultado —positivo o negativo—, en términos de reconocer la capacidad, permitiera a su vez predicar la misma respecto de un nuevo problema.

6 En el ámbito de la UE puede verse el documento 2018/2752(RSP) *Resolución sobre sistemas de armas autónomos*. De igual manera, el dictamen del Comité Económico y Social Europeo, de 31 de mayo de 2017, preconiza un enfoque de la inteligencia artificial basado en el control humano y la prohibición de los sistemas armamentísticos autónomos letales. Este ideario se puede resumir en las últimas declaraciones del secretario general de la ONU quien ha manifestado que «lo más grave es que la inteligencia artificial está erosionando el principio fundamental del control humano sobre el uso de la fuerza», mencionando la selección de objetivos a través de algoritmos para tomar decisiones de vida o muerte. Ver: https://news.un.org/es/story/2024/12/1535261

data; actividades todas ellas en las que debe intervenir el factor humano, tanto en la programación como en el posterior control de su ejecución[7].

Ya el Acta final de Helsinki de 1975 establece «la necesidad de contribuir a la reducción de los peligros de conflicto armado y de errores de interpretación o cálculo sobre actividades militares que puedan suscitar temores, particularmente en una situación en la que los Estados participantes carezcan de información clara y oportuna sobre la naturaleza de tales actividades». Sin hacer mención expresa a una realidad en ese momento inexistente, es perfectamente extrapolable esta reflexión de la Organization for Security and Cooperation in Europe (en adelante OSCE) a la situación actual si lo que se pretende es la reducción de los riesgos inherentes a los conflictos, incluso la mejora de la transparencia en el ámbito de la planificación y las actividades militares[8]. Obviamente no es una referencia expresa al empleo de IA en el ámbito armamentístico, pero la referencia continua a la existencia de controles, la necesidad de confianza y previsibilidad aleja la idea de una opción autónoma de toma de decisiones por parte de sistema armamentístico sin un último control humano.

Aquí se entraría en el ámbito de la discusión surgido en torno a los conocidos como sistemas de armas letales autónomos (conocidos por las siglas en inglés Lethal Autonomous Weapons Sistems, LAWS), sistemas de armas autónomos (AWS), armas robóticas o incluso como robots asesinos. De nuevo, la cuestión estriba en si es necesario, conveniente o irrenunciable la determinación de límites y controles a la autonomía de estos sistemas dados los efectos colaterales que dichos ingenios tecnológicos pueden provocar, aun cuando se encuentren bajo el control de operadores o controladores, incluso si se cumplen las normas del derecho internacional humanitario (Queirolo Pellerano, 2019).

Procede entrar en este momento a hacer una referencia, aún somera, a la diferencia entre IA fuerte y débil. Esta última está diseñada para realizar tareas específicas y fácilmente predecibles, mientras que la IA fuerte tiene la capacidad de aprender y adaptarse a nuevas situaciones. Expresado en otros términos, se puede considerar a la IA fuerte como aquella que es capaz de razonar y tomar decisiones por sí misma —autoaprendizaje—, mientras que la IA débil solo sigue instrucciones, por lo que el factor humano en esta última es determinante, tanto a la hora de realizar las indicaciones como, sobre todo, por la capacidad de control que no desaparece en ningún momento.

7 Recientemente, la Agencia Española de Protección de Datos ha propuesto una sanción económica (aún no es firme) por importe de 200 000 euros a GSMA LIMITED, organizadora del Mobile World Congress de Barcelona (principal congreso mundial de tecnología), por no informar con claridad a la ciudadana que asistió al evento en calidad de ponente de qué haría con esos datos biométricos que había obtenido de esta en el propio congreso. La resolución puede consultarse en: https://www.aepd.es/es/documento/ps-00553-2021.pdf

8 En el seno de la OSCE es fundamental el *Documento de Viena sobre Medidas Destinadas a Fomentar la Confianza y la Seguridad*, que promueve la confianza y la previsibilidad a través de medidas de verificación y en pro de la transparencia que abarcan las fuerzas armadas y los sistemas principales de equipos. El Marco para el Control de Armamentos, acordado en 1996, vino a reconocer que el control de armamentos, que incluye el desarme y el fomento de la confianza y la seguridad, forma parte del concepto de seguridad integral y cooperativo de la OSCE. Por su parte el Documento de Viena, el Tratado sobre Fuerzas Armadas Convencionales en Europa (FACE) y el Tratado de Cielos Abiertos constituyen una red de obligaciones y compromisos en materia de control de armamento, que se vinculan entre sí porque todos ellos que tienen como fin fomentar la previsibilidad, la transparencia y la estabilidad militar y reducen el riesgo de que se produzca un conflicto importante en Europa.

Un sistema armamentístico con una IA débil sería asumible por cuanto tanto la decisión como el control tienen su origen e implementación en un humano. Sin embargo, no sucederá lo mismo con un sistema de armas que emplee IA fuerte.

2.1 Máquinas inteligentes

En puridad, ni tan siquiera un algoritmo —por cuanto que es una fórmula, un método, incluso un procedimiento— en sí mismo tiene necesariamente que ser IA. Siempre que el resultado de su aplicación sea producto exclusivo de esa fórmula especificada por el usuario, no se estaría ante una manifestación de IA. Cuando en 1956 John McCarthy la definió, lo hizo como la «ciencia e ingenio de hacer máquinas inteligentes», esto es, una combinación de algoritmos planteados con el propósito de crear máquinas que presenten las mismas capacidades que el ser humano. Si no se completa esta combinación y el algoritmo se limita a la fórmula que procura resultados conforme su programación, se trata de una automatización, no de una IA[9].

En definitiva, la característica que define a la IA es que esta habilita la posibilidad de que las máquinas aprendan de su propia experiencia al tiempo que se ajustan a contribuciones nuevas y, al igual que los humanos, lleven a cabo tareas. Para ello tienen que concurrir tres procesos: el aprendizaje (consistente en la adquisición de información y aplicación de las reglas para el uso de la información)[10], el razonamiento (empleo de las reglas para llegar a conclusiones aproximadas o definitivas) y la autocorrección.

Estos procesos concurren en todos los tipos de inteligencia artificial que, como cada vez que se pretende categorizar algo, siempre hay que hacerlo teniendo en cuenta el criterio clasificador, cuando no el prisma con el que se observe o, lo que es lo mismo, la subjetivización del experto. Sin ánimo pretencioso alguno, solo con mero afán clarificador, se puede decir, con relativamente pocas probabilidades de equivocación, que en esencia se puede dividir en dos grupos las manifestaciones de IA. Por un lado, estarían aquellas que únicamente emplean la lógica, en contraposición de las que además de esta recurren a la intuición.

Las primeras, las más recurrentes, son las que aplican los algoritmos (principios racionales de pensamiento humano). Por su parte, el segundo tipo, al incorporar la intuición, recibe el nombre de «redes neuronales artificiales», cuyo precursor es Hinton. Al igual que las primeras, funcionan con algoritmos, pero estos están diseñados como si

9 No se puede dejar de referir a quien es considerado como el padre de la inteligencia artificial. En 1947, Alan Turing (2012), al responder a la pregunta «¿puede pensar una máquina?», sentó las bases de la IA. La publicación de esta respuesta en 1950 es la primera manifestación bibliográfica de la IA. Si bien la respuesta no es irrefutable ni absoluta, traza las líneas generales por las que debería discurrir una respuesta que aunara cierta precisión al tiempo que fuera fácil de usar. En estas líneas, el test de Turing conduce a concluir que una máquina piensa si un ser humano, que se comunica con la máquina y con otros seres humanos, no es capaz de diferenciar cuando su interlocutor es una máquina y cuando un humano.

10 Aquí, en el proceso de aprendizaje, es donde cobra trascendencia la cuestión relativa al uso de los datos y se manifiesta toda la normativa de protección de los mismos, en cuanto sea de naturaleza personal, lo que es una manifestación máxima de la interrelación existente entre la IA y el derecho.

se tratara de neuronas humanas para que la máquina aprenda por sí sola. Probablemente este grupo sea más conocido por su nombre coloquial *deep learning* —«aprendizaje profundo»[11]—. En definitiva, en el primer caso, el algoritmo, por así decirlo, opera de forma individual, donde el resultado que arroja es consecuencia directa de aquel. En cambio, la opción del *deep learning* contempla algoritmos múltiples, de tal forma que un algoritmo puede, porque está diseñado para ello, a su vez crear otros algoritmos, lo que contribuye a que, de este modo, se produzcan procesos de aprendizaje a partir de los resultados que obtienen unos y otros, en lo que es todo un proceso de entrenamiento[12].

En un limbo legal en cuanto al marco regulatorio completo que se aplica al uso de los sistemas biométricos en materia de privacidad, la utilización de aplicaciones de IA para la identificación biométrica viene siendo cada vez más frecuente. Solo desde la perspectiva de los datos se ha ido abordando la cuestión regulatoria de la «custodia» de los datos biométricos y los distintos tratamientos de los que pueden ser objeto que tienen un objeto variopinto. Desde aplicaciones para detectar comportamientos corporales que hagan prever la comisión de un tipo legalmente no admitido (comportamientos criminales o cuando menos ilícitos) al análisis de voz para sustituir al tradicional polígrafo en entrevistas de trabajo, sobre todo en puestos de alta dirección o con cualificación determinada, que exigen minimizar los riesgos inherentes a la persona que va a ocupar el puesto de trabajo al que aspira.

Lo cierto es que se está ante una cierta situación de legalidad reducida, por la única viabilidad cierta de poder aplicar la normativa de protección de datos (a nivel europeo, a diferencia de lo que ocurre en otros Estados como la RPC, EE. UU., Japón, etc.).

Esta situación ciertamente próxima a la alegalidad mantiene en vilo a las autoridades de control de toda Europa, que en ocasiones, lejos de mantener criterios comunes, abruman a la ciudadanía con discrecionales resoluciones bajo el lema de «en caso de duda la interpretación más favorable», y debe superar los juicios de «idoneidad, necesidad y proporcionalidad en sentido estricto», «una norma con rango de ley no es suficiente», «no se justifica la necesidad del tratamiento», «no cabe confundir necesidad con conveniencia», «la regulación actual es insuficiente» o el uso de esta tecnología debe considerarse «el último recurso»[13].

11 También conocidas como sistemas conexionistas, se caracterizan por que las neuronas, al conformar una red neuronal, están conectadas a su vez con otras a través de unos enlaces que determinan los efectos sobre las neuronas adyacentes. Así, estos sistemas aprenden y se forman a sí mismos, en lugar de ser programados de forma explícita, y sobresalen en áreas donde la detección de soluciones o características es difícil de expresar con la programación convencional (Hinton, 1912).

12 Uno de los ejemplos más conocido es el de los algoritmos de prueba y error. Entre ellos se puede encontrar el método de Newton-Rapshon. Es un tipo de algoritmo recursivo de búsqueda «hacia delante y hacia atrás», que se basa en aplicar de nuevo al resultado de haberlo aplicado previamente, lo que se puede hacer un número infinito de veces. Por ejemplo, es un sistema empleado en juegos como el ajedrez para aprender a jugar al mismo. De esta forma, se plantea un esquema (recursivo de prueba y error) donde las pruebas que realiza el ordenador construyen una serie de caminos de búsqueda de soluciones. Estos caminos de búsqueda generarán como resultado un árbol de subtareas, donde algunos caminos no llegan a la solución mientras que otros sí. En definitiva, es un ejercicio de tareas repetitivas que procura la mejora gradual del resultado a través de *deep layers* (capas profundas). Este es el funcionamiento de redes neuronales que son capaces de operar a través de capas de información cada vez más profundas para realizar predicciones, solucionar problemas o detectar características en objetos o situaciones.

13 Entre otros, ver COM (2021), 205-final, y COM (2024), 28-final.

Pero, si con la protección de datos todavía no se han solucionado los problemas legales que la globalización y la tecnología plantea, con los sistemas biométricos y la IA Europa no va a caminar mejor.

El *Libro Blanco sobre la inteligencia artificial –un enfoque europeo orientado a la excelencia y la confianza* (2020)[14] explica que la recopilación y el uso datos biométricos para la identificación remota entraña riesgos específicos para los derechos fundamentales y concluye que, a fin de abordar las posibles preocupaciones sociales con relación al uso de la IA y con el objetivo de evitar la fragmentación del mercado interior, la Comisión abrirá un debate europeo sobre las circunstancias específicas, si las hubiera, que puedan justificar dicho uso, así como sobre las garantías comunes.

2.2 *Respuesta jurídica*

Como ya se ha expuesto anteriormente, el derecho no puede conformarse con ir detrás de la realidad, reflejando el contexto, sino que debe poder responder de manera adecuada a los retos que implican los nuevos escenarios. El derecho debe comprender sus implicaciones, funcionamiento y utilidades. Pero, por la naturaleza de la nueva realidad, se tienen que anticipar y aventurar aplicaciones futuras derivadas de la capacidad de computación, cálculo y uso de la IA en todo aquello que suponga una ayuda para la toma de decisiones, cuanto más si se decide delegar en la IA la propia toma de decisiones, a medida que la misma vaya evolucionando y aproximándose a la *ratio decidendi* humana (Hoffmann, 2018).

Luego lo anterior, la mera ejecución de tareas que se desarrollan de forma sistemática, reproduciendo mecanismos, donde la capacidad de decisión está coartada por la especificidad de los límites que rigen su funcionamiento, es lo que se conoce como automatización, en ningún caso IA. Ciertamente, ejemplos como el Centro de Denuncias Automatizado de la Dirección General de Tráfico (ESTRADA), que gestiona de forma automática e inmediata las infracciones captadas por los radares o cámaras instaladas en las carreteras españolas, donde el sistema selecciona las imágenes captadas por los radares, desechando aquellas que «a su juicio» no reúnen los requisitos para incoar con garantías el procedimiento sancionador[15], suponen manifestaciones muy simples de IA. La parte más interesante de estas, desde la perspectiva del presente

14 COM (2020), 65-final.

15 Según la propia Dirección General de Tráfico informa en su web (https://www.dgt.es/menusecundario/dgt-en-cifras/dgt-en-cifras-resultados/dgt-en-cifras-detalle/Evolucion-de-Denuncias-e-Ingresos.-Denuncias-en-funcion-de-la-provincia-y-el-precepto/), el Centro Nacional de Tratamiento de Denuncias Automatizadas gestiona anualmente una cifra aproximada de 1,5 millones de denuncias por infracciones de tráfico, de la cuales se suelen descartar en torno al 40 % debido a que la imagen captada puede provocar dudas o no cumplen con los requisitos exigidos. El sistema funciona a partir de las imágenes tomadas por el radar o la cámara, que detecta un exceso de velocidad u otra infracción de la normativa de tráfico (como la ausencia de cinturón de seguridad o no respetar la señalización como «saltarse el semáforo», *stop*, etc.), A continuación, la información se envía vía satélite al centro de tratamiento, donde se procede a identificar al infractor mediante consulta al registro general de vehículos; el propio sistema genera la notificación de la denuncia y, también de manera automática, se procede a la notificación al titular del vehículo.

trabajo, es aquella que plantea la posibilidad de que un robot fuera capaz de adoptar sus propias decisiones o incluso, yendo un paso más allá, sus propias ideas hasta el punto de que cabría preguntarse qué titularidad tendrían estas ideas, si se pueden entender las mismas como consecuencia directa de la programación inicial, que depende de personas o si, por el contrario, al menos en parte, es fruto de su propio aprendizaje mediante el procesamiento de información.

Siguiendo con el ejemplo mencionado con anterioridad sobre los sistemas de IA usados en el ámbito militar, cuando se usan para la defensa activa, pueden coordinar y llevar a cabo ciertos ataques preventivos de manera prácticamente autónoma y reconocer dispositivos, elementos militares e individuos.

Habría que determinar hasta qué punto los avances tecnológicos y la creciente autonomía de estos sistemas, en detrimento del factor humano (en la recogida y procesado de datos, toma de decisiones), afectan ahora al control del armamento por parte de los Estados. La capacidad de monitorizar y procesar datos mediante sistemas de IA facilita la supervisión y verificación de los límites que establece el derecho internacional o los acuerdos entre Estados a la vez que aumenta y mejora la calidad de los datos almacenados.

El factor humano se reduciría a la supervisión y toma de decisiones en los últimos niveles del sistema, así como a controlar el flujo de información derivado de los sistemas implementados.

3 La ética en la IA. El ser y el deber ser

Es irrefutable que la inteligencia artificial compone unas de las tecnologías más avanzadas pero, al mismo tiempo, comprometedoras de los últimos tiempos. Las numerosas aplicaciones que ayudan a mejorar el sistema social, económico, automovilístico, sanitario y educativo entre otros muchos no deben ocultar la cantidad de riesgos que entrañan por cuanto la máquina, que no deja de ser eso, tiene un carácter automático, sobre todo a la hora de aplicar los derechos propios de la intimidad y relativos protección de datos, tan protegidos por la normativa en los últimos años[16].

16 La propia Comisión Europea, en conjunto con los Estados miembros, llevó a cabo un plan para una correcta regulación de este tema en concreto, lo que desembocó en la publicación de las *Directrices éticas no vinculantes para una IA fiable* (Comisión Europea, Dirección General de Redes de Comunicación, Contenido y Tecnologías), *Directrices éticas para una IA fiable* (Oficina de Publicaciones, 2019, ver: https://data.europa.eu/doi/10.2759/14078). Por su parte, el Parlamento también ha aprobado algunas resoluciones dentro del ámbito de la inteligencia artificial, la más reciente *Resolución del Parlamento Europeo, de 3 de mayo de 2022, sobre la inteligencia artificial en la era digital (2020/2266(INI))*. Pero, sin duda, destaca la propuesta de un reglamento que establecería tanto los principios éticos como las obligaciones de carácter legal, tanto para el desarrollo como la implementación y uso de la inteligencia artificial, la robótica u otro tipo de tecnologías de carácter similar, propuesta de reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, por el que se establecen normas armonizadas en materia de inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión (COM(2021)0206). Está claro que existe la pretensión de regular las tecnologías que se consideran de alto riesgo, aquellas que se considera que tienen un riesgo significativo a la hora de causar algún tipo de daño a las personas o a la sociedad, sobre todo mediante la vulneración de los derechos fundamentales o las normas establecidas por la UE, entre las que indudablemente se incluye la IA.

La existencia cierta de algoritmos no conocidos, «*opacos, desregulados e irrefutables*» (O´Neil, 2016), está causando un importante retroceso de las libertades y de la igualdad real de los ciudadanos.

No se sabe cómo operan; en ocasiones ni siquiera se conoce su existencia, pero lo más llamativo es que también se ignora qué es lo que pueden llegar a hacer: la maximización del resultado que se le pide puede llegar a sorprendernos. Descubriendo en el ingente caudal de datos del que se alimentan conexiones y patrones que el humano no capta, coincidencias y regularidades de las que se prescindiría son ahora elevados a resultados que aciertan a predecir que aquellas coincidencias se darán en el futuro. Sirven entonces, no solo para explicar, sino también para «predecir» y orientar la decisión humana, cuando no la adoptan por sí mismos. Pero lo hacen con base en correlaciones y reglas de conducta humanas que se conocen como derecho y que se asientan en la causalidad y responsabilidad.

Por otra parte, los datos con que se alimentan tales sistemas están trasladando información real que por tanto corresponde a los sesgos que presenta la sociedad en su estado actual, en el momento presente. Si se deja que el sistema de IA interprete tales datos como los únicos que debe considerar, ¿no se forzará aún más la reproducción de las desigualdades, de los sesgos de injusticia que ofrecen los datos actuales?[17]

La tecnología se ha ocupado de ofrecer buenas y eficientes soluciones mediante análisis de datos masivos, computación y algoritmos, pero ¿se ha ocupado el ser humano de la adecuación a sus propios valores de dignidad de la persona, libertad y responsabilidad, igualdad y solidaridad que debieran ser los principios de toda *ius algoritmia*? Y, aún más, ¿se ha ocupado de que tales soluciones respondan a *reglas humanas*? Porque, finalmente, de esto se trata: las soluciones técnicas, que producen excelentes resultados, se asientan en una consideración no causal (*no responsable*) de la conducta humana. El derecho y las reglas de este en el momento presente estaban construidas sobre el presupuesto de la responsabilidad y la causalidad. ¿Va el ser humano hacia un derecho algorítmico de la simple *correlación*?

Como mínimo, será posible generalizar algunos principios jurídicos generales de la *iusalgoritmia*. Los humanos tienen derecho a conocer la existencia de sistemas de IA que influyen sobre su posición jurídica. No importa ahora que funcionen o no estableciendo «perfiles» en sentido estrictamente normativo. Toda decisión que se fundamente en esencia o de forma exclusiva en el uso de sistemas de IA debe tener como presupuesto de su admisibilidad jurídica la posibilidad de que pueda ser discutida la lógica humana en que se fundamenta. Por ello, es requisito imprescindible que la existencia de aquel sea pública. Ni tan siquiera sería suficiente con la posibilidad de

17 El filósofo Daniel Innerarity (2022) no ha dudado en señalar que «los algoritmos son conservadores y nuestra libertad depende de que nos dejen ser imprevisibles». No se trata de renegar del resultado de aquello que, en definitiva, es obra nuestra, simplemente hay que adaptarse a las nuevas realidades al tiempo que se debe hacer lo propio con procedimientos e instituciones. Todo ello sin renunciar a la esencia humana, por lo que la tecnología facilitará las tareas y también la determinación de los fines de estas.

conocer la existencia; en tanto no se positive, no hará posible que pueda ser objeto de examen[18].

Tras la transparencia se revela el principio de cautela (evaluación previa) totalmente consolidado por el derecho comunitario, que lo declaró, con carácter general, como un principio general de este (sentencia del Tribunal de Primera Instancia, de 26/11/2002, Artegodan y otros/Comisión)[19]. En su virtud, no todo algoritmo será admisible en derecho por que, por ejemplo, su empleo puede terminar por dar lugar a prácticas restrictivas de la competencia. Así, este principio operaría de una manera muy sencilla, que gráficamente se podría resumir con un algoritmo básico: incertidumbre científica + sospecha de daño = acción precautoria. No es desconocida la cuestión para Cotino Hueso (2019a) ni para Fernando Pablo y Terrón Santos (2019), que consideran la ética como una herramienta definitiva para afrontar los riesgos de la IA con un argumentario parecido.

Este principio exige su aplicación atendiendo a las particularidades propias de esta realidad, donde la tecnología sin el derecho se encontrará con serios problemas para su implementación por el riesgo inherente a su uso del que cada vez se es más consciente.

18 Incidiendo en la transparencia, la *Resolución 200/2017, de 21 de junio, de la Comisión de Garantía del Derecho de acceso a la información pública de Cataluña, en relación con el derecho de acceso a un software utilizado para designar a los miembros de los comités encargados de evaluar los exámenes de admisión a las universidades públicas*, vino a reconocer el derecho de la persona reclamante para acceder al código fuente del programa informático empleado por el Consejo Interuniversitario en la designación de los miembros de los tribunales correctores de las PAU, por razones similares a las que llevaron a la resolución de esta Comisión de 21 de septiembre de 2016 a estimar la Reclamación 124/2016 y declarar el derecho a conocer el algoritmo matemático implementado por dicho programa informático.

En esos términos, manifiesta que:

> «[…] el código fuente de un programa informático utilizado por la Administración en la designación de los miembros de tribunales evaluadores constituye información pública a los efectos del artículo 2.b de la Ley 19/2014, del 29 de diciembre, de transparencia, acceso a la información pública y buen gobierno (LTAIPBG). Según este precepto, se entiende por información pública "la información elaborada por la Administración y la que ésta tiene en su poder como consecuencia de su actividad o del ejercicio de sus funciones, incluida la que le suministran los otros sujetos obligados de acuerdo con lo establecido en esta ley". Esta definición incluye toda la información que la Administración elabore o tenga en su poder en ejercicio de sus funciones, con independencia del lenguaje o forma en que se exprese. La información pública incluye así, no sólo aquella que se expresa en lenguaje natural (en palabras, que es la más habitual), sino también la expresada mediante fotografías, vídeos, planos, señales, etc. o mediante otros lenguajes, como el matemático o, en este caso, el informático. El artículo 19.1 LTAIPBG confirma esta noción amplia de información pública cuando dispone que "el derecho de acceso a la información pública incluye cualquier forma o soporte en que esta información haya sido elaborada o en que se conserve". El mismo se desprende del artículo 13 de la Ley básica estatal 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno, cuando establece que "se entiende por información pública los contenidos o documentos, cualquiera que sea su formato o soporte, que estén en poder de alguno de los sujetos incluidos en el ámbito de aplicación de este título y que hayan sido elaborados o adquiridos en el ejercicio de sus funciones"».

19 STPI de 26 de noviembre de 2002 Artegodan y otros/Comisión, asuntos acumulados T-74/00, T-76/00, T-83/00 a T-85/00, T-132/00, T-137/00 y T-141/00. La controversia se suscitaba por la retirada de la autorización de un medicamento de uso humano con efecto antianoréxico. El apartado 184 de la sentencia establece:

> «De lo anterior resulta que cabe definir el principio de cautela como un principio general del Derecho comunitario que impone a las autoridades competentes la obligación de adoptar las medidas apropiadas con vistas a prevenir ciertos riesgos potenciales para la salud pública, la seguridad y el medio ambiente, otorgando a las exigencias ligadas a la protección de estos intereses primacía sobre los intereses económicos. En efecto, en la medida en que las instituciones comunitarias son responsables de la protección de la salud pública, de la seguridad y del medio ambiente en todos sus ámbitos de actuación, cabe considerar el principio de cautela como un principio autónomo que se desprende de las mencionadas disposiciones del Tratado».

La progresiva toma de conciencia sobre los problemas vinculados a la IA produce sombras que llevan a replantearse no la utilidad —indiscutida—, pero sí la necesidad, que terminará siendo insoslayable de comprender el funcionamiento de esta realidad, para poder controlar en la medida de lo posible a los algoritmos y, cuando esto no sea posible, acotar el recurso a los mismos (Bostron, 2014). Si no se llega a comprender el funcionamiento de los mecanismos tecnológicos en que basan su funcionamiento los algoritmos, difícilmente se podrán justificar sus resultados. Esto lleva a observar que igualmente importante es identificar los objetivos que se quieren alcanzar recurriendo a los mismos porque solo así se comprenderá qué papel debe desempeñar la regulación. De otro modo, cualquier intento de limitar a través de la normativa el uso de la IA será un fracaso desde cualquier prisma con que se mire[20].

Sin ir más lejos, ¿qué pasaría si un robot —algoritmo de aprendizaje profundo— fuera diseñado para una determinada labor pero, consecuencia de esta, llevara a cabo una invención? El resultado de su labor ¿sería propiedad de quien adquiere el robot o del programador inicial? Está claro que el *software* permite que se produzcan dos titulares de derechos, los del fabricante o programador inicial, que lleva a cabo una programación inicial, y los del usuario, que obtiene esta inteligencia que puede llevar a cabo una programación más profunda. Está claro que, desde una perspectiva jurídica, la respuesta sale a la luz casi de forma inmediata: dependerá de lo acordado en el contrato de venta o cesión de uso. Así, se pueden dar situaciones que amparan tanto que el usuario solo obtiene una licencia de uso y no la propiedad del mismo algoritmo, como todo lo contrario. Pero también daría lugar a la problemática si se considerara que el robot tiene una capacidad jurídica propia. En tal caso, ¿quién tendría la titularidad del descubrimiento? A la respuesta solo es posible aproximarse desde una cuestión distinta, que es la de considerar si la IA goza de la titularidad de la persona jurídica. En este caso, la disputa será entre el fabricante, el usuario o el propio robot (Eidenmüller, 2017). Hasta que la legislación no clarifique de forma clara esta materia no se puede aseverar con rotundidad quién sería el verdadero titular del descubrimiento.

4 Límites a la inteligencia artificial

Esta cuestión ya tiene cierto recorrido, como demuestra que, a mediados de la primera década del siglo xxi —esto es, hace más de quince años—, ya se planteaba la necesidad de controlar la IA desde una perspectiva legal[21]. La evolución de esta materia se ha ido apartando

20 Opinión que es compartida por Boix Palop (2007: 145), que aboga por la correcta, además de consensuada, identificación de los que han de ser objetivos finales de la regulación, sin olvidar los valores a los que esta debe atender. Este razonamiento implica que solo a través del conocimiento de las implicaciones se podrá llevar a cabo una regulación que programe a las IA en función de lograr unos objetivos u otros. Por ello, ante esta nueva tercera revolución tecnológica y productiva, el papel de la ley ha de ser totalmente consecuente con esta necesidad de establecer objetivos y finalidades y tratar de, a partir de ahí, reorientar el funcionamiento de estos instrumentos.

21 Diversos medios de la época recogían la noticia de que, en países como Japón o Corea del Sur, la cuestión había sido objeto de un análisis más detallado, hasta el punto de que, primero en Japón y después en Corea del sur, se llegaron a redactar documentos que, recogiendo los principios básicos de Asimov, abogan por el claro propósito de que las máquinas

de esos pronunciamientos iniciales conforme los cuales la responsabilidad deriva del control férreo humano. Pero las circunstancias mandan y hoy en día se hace complicado pretender establecer una «responsabilidad algorítmica» en los desarrolladores de determinados programas, sobre todo en aquellos que emplean aprendizaje automático —*machine learning*— (ver: https://www.abogacia.es/2018/05/14/ese-algoritmo-el-discriminador/ - _ftn2). La ausencia de ese control pretendido empuja a buscar metodologías preventivas para evitar sesgos en la toma de decisiones automáticas que puedan afectar derechos fundamentales[22] y evitar que exista discriminación preprogramada y autoejecutada[23]. En definitiva, a evitar una responsabilidad «no deseada».

Ciertamente, las máquinas de aprendizaje plantean desafíos más desconcertantes que los sistemas expertos de la primera ola para los valores incrustados en el derecho administrativo. La transparencia, la responsabilidad, la previsibilidad, la igualdad ante la ley incluso la coherencia se cumplen en tanto imperen ciertas reglas, de modo que sería válida la tecnología específica implementada a su tenor. Pero la rápida evolución de la IA produce, obsérvese que se habla en términos de presente, que los algoritmos se integren cada vez más en procesos opacos de toma de decisiones, sin intervención humana, lo que hará que sea extremadamente difícil desafiarlos, con la evidente pérdida de garantías que ello supone. Esto lleva a que la decisión con respecto al grado de sofisticación de la automatización por implementar debe considerarse con cuidado entre los elementos a través de los cuales se diseña una regla específica que confiere poder de decisión[24].

Los sistemas de IA, ya sean usados por el poder público para la seguridad ciudadana o por el militar para garantizar la efectividad de sus actuaciones, han sido objeto de regulación internacional y de exigencia de límites a su capacidad de actuación. En

estén siempre bajo control humano, sin que exista posibilidad de atribuir de titularidad jurídica o de cualquier otro tipo a los robots. Entre otros, destaca el artículo de T. Delclos (2007).

Asimov (1942) fijó las tres leyes que deben regir la existencia de los robots: «Un robot no puede hacer daño a un ser humano o, por inacción, permitir que un ser humano sufra daño; un robot debe obedecer las órdenes de los seres humanos, excepto si estas órdenes entran en conflicto con la Primera Ley; un robot debe proteger su propia existencia en la medida en que esta protección no entre en conflicto con la Primera o la Segunda Ley». En conclusión, un robot hará lo que se le dice, pero no lo que se desea o necesita.

La propia resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de derecho civil sobre robótica (2015/2103(INL)), reconoce la existencia de precedentes previos conducentes a adoptar medidas normativas en el ámbito de la robótica y la inteligencia artificial, incluso con resultados positivos ciertos, hasta el punto de que determinados Estados miembros se habían aproximado, en modo de tentativa, a la elaboración o modificación de normas jurídicas al objeto de incluir las nuevas aplicaciones tecnológicas.

22 Crear sistemas que aprenden automáticamente —es decir, identificar patrones complejos en millones de datos— lleva a que el algoritmo revise esos datos y sea capaz de predecir comportamientos futuros, lo que los trasforma de reactivos a proactivos. Automáticamente, también en este contexto, implica que estos sistemas se mejoran de forma autónoma con el tiempo, sin intervención humana.

23 Un buen parámetro para el legislador sería tratar de revisar las exposiciones y materias abordadas en la Neural Information Processing Systems, una de las grandes reuniones internacionales sobre algoritmos y procesamiento de información. El establecimiento de límites en las solicitudes que realice el sistema, algo similar a lo que hace Amazon M. L. (ver: https://amzn.to/2TDQQPF), sería otra posibilidad para tomar en consideración.

24 A mayor abundamiento, ver: Pasquale, 2015: 147. Por su parte, el informe del año 2004 del Consejo Australiano de Revisión Administrativa (Australian Administrative Review Council) en particular subraya que, ya mucho antes del año 2000, los órganos administrativos solían recurrir a procesos automatizados de toma de decisiones relevantes, a menudo sin siquiera mencionarlo en la propia decisión.

2017, el Consejo de Seguridad de Naciones Unidas emitió la Resolución 2396/2017, de obligado cumplimiento para todos los Estados parte, que obliga a implementar sistemas de recogida de datos biométricos con el fin de implantar mejoras en las medidas de seguridad preventivas de ataques terroristas. Estos sistemas implican una mejor identificación de ciudadanos (escáneres faciales, huellas dactilares), así como el uso de sistemas de datos cruzados entre agencias y organizaciones de gobierno y seguridad, a nivel nacional y supranacional.

También en el ámbito europeo, la Unión Europea (en adelante, UE) se manifestó en 2021 a favor de una regulación estricta acerca de la captación de datos biométricos en espacios públicos, por considerarse una amenaza para la seguridad[25].

En Europa, las técnicas biométricas se evalúan por el Convenio Europeo de Derechos Humanos y Libertades Fundamentales de 1950 (en adelante, CEDH). Este tipo de tecnología presenta ciertas vulnerabilidades, como la duda razonable que surge en relación con sus posibles fallos, ya sean provocados o fortuitos. Además, se muestran las preocupaciones éticas inherentes al empleo de un sistema totalmente autónomo y las implicaciones que puede tener la identificación errónea de una persona (civil o militar). Por ejemplo, son habituales los fallos de IA a la hora de identificar determinados patrones raciales o sexuales o la relativa facilidad con la que se puede «engañar» a la IA a la hora de identificar los datos biométricos obtenidos[26].

4.1 La responsabilidad

Partiendo de la premisa cierta de que derecho y sociedad no son rivales —al contrario, son elementos simbióticos, inseparables—, se ha de señalar que el derecho es el medio necesario para el desarrollo y la evolución social. Luego no puede limitarse a dar respuestas, también ha de remover obstáculos y facilitar el avance científico y tecnológico, lo que solo desde la anticipación normativa es posible. Lograr esto es cuestión de flexibilidad más que de adaptación[27].

25 C(2021) 32 Decisión de Ejecución (UE) 2021/27 de la Comisión de 7 de enero de 2021, sobre la solicitud de registro de la iniciativa ciudadana europea titulada «Iniciativa de la sociedad civil para la prohibición de las prácticas de vigilancia biométrica masiva».

26 Basta realizar una simple prueba para tomar razón de esto. Si en cualquier buscador se solicita la búsqueda de una imagen de un hombre con éxito o de una mujer hermosa, los resultados generados muestran con mucha mayor profusión imágenes de personas de raza caucásica, excluyendo así la diversidad racial. Sin ir más lejos, es un mero ejemplo, pero significativo, de cómo los algoritmos muestran preferencias hacia ciertos grupos raciales y discriminan o minusvaloran a otros.

27 «El Derecho ha de desarrollarse en mayor proporción al progreso de la técnica, aumentando su cometido y su responsabilidad. Pues el dilema derecho-técnica puede aportar factores de amenaza a la divergencia –ya actual- entre desarrollo técnico-económico y jurídico. Mostrar esto ha sido uno de los fines de este trabajo que concluyo con la conciencia de ser sólo una advertencia y no una solución». Con estas palabras, Villar Palasí (1975) hace referencia a la necesidad de que el derecho recurra a la lógica para ser capaz de no quedar fuera de la relación que le une al desarrollo, dando por sentado que el mismo se va a producir, incluso con un marco jurídico que no sea el idóneo para ello, si bien se puede ver ralentizado, con el evidente perjuicio social que ello supondría. Luego, muy al contrario de lo que opinan quienes solo ven el derecho como el perseguidor, la ciencia jurídica ha de ser el promotor del desarrollo científico-técnico, por extensión económico, en beneficio de la sociedad.

Normas laxas, basadas en la lógica, en principios generales actualizados ofrecerán este resguardo a la sociedad para que esta, en su avance, no pierda las garantías jurídicas que necesita para su normal desenvolvimiento.

Entre estas garantías jurídicas se encuentra obviamente la sujeción al principio de legalidad, que no deja de ser la base de los prominentes del derecho, al igual que la equidad, la transparencia, la racionalidad y la responsabilidad.

Este artículo ya se ha centrado en la cuestión ética, acompañada de la transparencia y la racionalidad desde la perspectiva de la ética. Luego toca detenerse en la cuestión de la responsabilidad aplicada a la IA. Conviene recordar que los algoritmos sirven no solo para explicar, sino también para «predecir» y orientar la decisión humana, cuando no la adoptan por sí mismos. Pero lo hacen a base de correlaciones y las reglas de conducta humanas que, habiéndolas llamado derecho, se asientan en la causalidad y responsabilidad.

Tradicionalmente, la tecnología se ha ocupado de ofrecer soluciones que en términos generales son válidas y positivas, a través del análisis de datos masivos, computación y algoritmos. Sucede que, al hacerlo, se ha aparcado la adecuación de estas soluciones a los valores de dignidad de la persona: libertad e igualdad, solidaridad y, por supuesto, responsabilidad, que debieran ser los principios de toda *ius algoritmia* (Terrón, 2022). Yendo más allá, es posible afirmar con poco margen de error que el ser humano se ha despreocupado de que tales soluciones respondan a sus propias reglas, que debería haber sido el argumento principal que esgrimir. Porque, finalmente, de esto se trata cuando las soluciones técnicas, que producen excelentes resultados, se asientan en una consideración no causal (no responsable) de la conducta humana. El derecho, las reglas del momento presente, estaban construidos sobre el presupuesto de la responsabilidad y la causalidad, pero se ha avanzado (¿sin ser conscientes?) hacia un derecho algorítmico de la simple correlación.

Aunque se esté hablando de IA, no se debe perder de vista que no se está ante un comportamiento extraño al ser humano, tal y como ya se ha puesto de manifiesto anteriormente. Esto supone que es este quien fija las reglas —diseña el código fuente— y, por tanto, el algoritmo motor de la IA desde un primer momento se encuentra sujeto a las determinaciones del hombre. Es cierto que existe la posibilidad de «transferir» la decisión a la máquina, pero siempre dependerá de la voluntad que se tenga de hacerlo. Cuestión que guarda relación directa con la posibilidad de atribuir la responsabilidad al autor material del algoritmo primigenio, puesto que será este el responsable de la contingencia, que con su decisión no evitó pudiendo hacerlo. La deliberación implica la atribución de responsabilidad, que incluso puede entenderse como una manifestación de mala fe al intentar disolver la responsabilidad conocida a través de la atribución de la decisión a terceros, el algoritmo en este caso. No sería un supuesto de dejación de funciones, ni tampoco en puridad se estaría ante una delegación competencial; más bien se trata de un recurso particularmente dirigido a evitar que el decisor será responsable aun debiendo serlo.

Si se estuviera ante meras manifestaciones de automatización, sin duda la cuestión es mucho más sencilla, por cuanto en ese caso se determina una continuidad en la

responsabilidad del órgano sobre sus actos, incluyendo los automatizados, que además deben sustentarse en un procedimiento concreto y conocido. Pero, dado que no es posible —tampoco inteligente— equiparar automatización a discrecionalidad, toda vez que el recurso tecnológico que supone la automatización queda restringido tan solo a supuestos claramente reglados y muy limitados en cuanto a posibles variaciones, la IA supone todo un reto desde la perspectiva de la responsabilidad.

No corresponde en este momento analizar toda la manifestación de responsabilidad que es posible en cuanto a la IA se refiere. Pero sí es posible avanzar que, en aras de la neutralidad tecnológica, conviene despejar, siquiera a grandes rasgos, la cuestión por cuanto es clara la afectación que la IA tiene sobre la sociedad. En ese sentido, es prioritario reforzar los mecanismos de responsabilidad y rendición de cuentas al tiempo que se garantiza la protección de los derechos y libertades de los ciudadanos, sean personas físicas o jurídicas en sus distintas manifestaciones[28]. Esto se puede alcanzar en una única forma, que no es otra que el establecimiento de limitaciones a la configuración de los algoritmos. A la hora de la elaboración de códigos algorítmicos, fórmulas como la auditoría algorítmica, la concienciación de los desarrolladores tecnológicos —humanismo digital— y la evaluación pública pueden servir como mecanismos para lograr la explicabilidad del algoritmo, sin necesidad de acudir al extremo de la liberalización del código algorítmico para descanso de la propiedad intelectual del mismo.

Entender el código algorítmico es determinante de la atribución de responsabilidad, ya que solo cuando quede claro que no ha existido una dejación de funciones, por cuanto no hubo voluntad alguna de trasferir la adopción de la decisión que genera la circunstancia, se podrá eludir la responsabilidad sobre la misma, de modo que esta recae en el algoritmo en sí, lo que se podría equiparar a una causa de fuerza mayor elusiva de la misma. Es decir, la determinación del fin es algo que debe ser inherente y exclusivo de la esencia humana. El fin podrá ser más o menos ambicioso y, si en él se encuentra «utilizar» al algoritmo mediante la no determinación de fines, que deberían haber sido determinados por el programador, la responsabilidad seguirá siendo de este último, por lo que no serviría aquello de «echémosle la culpa al algoritmo» (Innerarity, 2022).

5 Conclusiones

Lo aquí aportado no es fruto de la casualidad ni de la inventiva. No son pocos los organismos internacionales (Naciones Unidas, OSCE, etc.) que han mostrado su preocupación por el uso de sistemas biométricos y su aplicación militar, incluso sin desconocer intentos, de momento vanos, de regular este recurso. La principal preocupación se encuentra en cuanto se acercan los aspectos derivados de sistemas de IA autónomos, basados en la recogida de indicadores biométricos, como el

28 A estos efectos puede verse la *Guía de auditoría algorítmica*, elaborada por Eticas Research and Consulting S. L., disponible en: https://bit.ly/3hBvtX9

reconocimiento facial. Los conocidas como LAWS centran la mira de la sociedad civil y las organizaciones políticas vinculantes, que, como la UE, están a la procura de una legislación que regule el uso de las LAWS y de cualquier otro tipo de armamento autónomo basado en la coordinación multinivel por medio de la IA. En todos ellos, el uso de datos biométricos como medio de identificación de objetivos es un elemento imprescindible a la par que no exento de polémica.

Europa y la UE representan un espacio en el global, sin duda, pero no es uno de los principales actores del teatro geopolítico mundial. Por más que pueda ser arriesgado hacer esta afirmación para quien suscribe, lo cierto es en tanto que Estados Unidos, China, Rusia, incluso siendo atrevidos, se puede añadir a Corea del Norte, Irán o el propio Israel, no muestren interés —parecen todos ellos lejos de hacerlo, más allá de algún brindis al sol— en desarrollar una verdadera legislación con respecto a las LAWS o drones autónomos, el problema no va a desaparecer[29].

Tampoco es lícito negar que existen iniciativas en el Congreso de Estados Unidos destinadas a evaluar los riesgos que comporta el desarrollo de este tipo de sistemas autónomos, pero no van más allá de un análisis de riesgo, que distan mucho de ser una norma de aplicación y regulación del uso de estos sistemas[30].

Las técnicas biométricas en Europa están recogidas bajo el paraguas genérico que supone el CEDH, que aun con eficacia limitada por el ámbito de su cobertura y aplicación relativa, dado que el convenio no recoge, mucho menos regula, el uso de armas que impliquen biometría para la toma de decisiones autónoma, pero es el texto que ampararía, a fecha de hoy, su uso. Los miembros del Consejo de Europa, incluidos los 27 Estados de la UE, Reino Unido y Turquía, son parte del CEDH, por lo que el nivel de protección pudiera parecer mayor, sobre todo por la presencia de Rusia hasta septiembre de 2022, pero realmente ni es, ni ha sido así.

Si fuera infalible, el problema sería menor o directamente no existiría, pero lo cierto es que la tecnología biométrica no está exenta de vulnerabilidades. En el mejor de los supuestos, cuando menos habrá de afrontar una duda razonable sobre la indemnidad de los sistemas. La posibilidad de un fallo es inherente a cualquier elemento tecnológico. Con independencia del origen, la consecuencia podría ser la misma, tanto si el fallo en los sistemas estuviera provocado por un mal funcionamiento o si se estuviera ante una manipulación malintencionada. Luego, si algo falla, mejor que sea el decisor humano. Incluso cuando a este se le pueda controlar estableciendo mecanismos tecnológicos como sistemas de doble control, necesidad de confirmación, etc., que como mínimo alejen el fallo no intencionado. Al humano, la ética se le puede exigir; de la IA solo se

29 En la sede que supone la Convención sobre Ciertas Armas Convencionales y el Grupo de Expertos Gubernamentales (CCW y GGE, por sus siglas en inglés) sobre sistemas de armas autónomas letales se presentó la ponencia *Principles and Good Practices on Emerging Technologies in the Area of Lethal Autonomous Weapons System*, suscrita por Australia, Canadá, Japón, Corea, Reino Unido y Estados Unidos, sin que en la misma se abogue por una prohibición total de las LAWS, donde se limita el uso de las mismas al derecho humanitario internacional.

30 Congressional Research Service Report R44466, Lethal Autonomous Weapon Systems: Issues for Congress.

puede esperar que la herede a partir de los datos que la alimentan. Por consiguiente, tienen que existir preocupaciones éticas inherentes al empleo de un sistema totalmente autónomo. Es más, se entraría en otra cuestión como es la responsabilidad derivada del error autónomo de la IA. ¿Quién es el responsable: el algoritmo; su autor, que quebraría en el caso del deep *learning*; el usuario del sistema? Muchas posibilidades para una respuesta que puede parecer sencilla, pero comporta implicaciones muy delicadas en el momento en que se recurre a herramientas donde se sabe que la ética no impera y se espera impunidad moral.

Además, los fallos de la IA no son excepciones a la regla general. Patrones sexuales, raciales, étnicos, etc., no son extraños a los sistemas biométricos automáticos, que pueden ser engañados tanto en la captación de indicadores biométricos como en la propia identificación de estos. ¿Aceptaría la sociedad un sistema de cuya capacidad de discriminación entre militares y civiles no tuviera certeza absoluta?

Todas estas cuestiones están más allá del dilema ético, o quizá no. Los continuos avances tecnológicos no pueden dejarse en su desarrollo al libre albedrío, confiando en un sistema que emplea datos, máxime cuando estos datos deben controlarse para garantizar la viabilidad del resultado precisamente por la calidad de la información que se ha suministrado. Quienes abogan por la intervención, la regulación, la ética, etc., difícilmente van a acercar posiciones con aquellos sujetos más laxos, de moral distraída, que no ven, o no quieren, el riesgo en la forma de control de armamento.

En un futuro ya presente, es una realidad que la tecnología emergente va a determinar la relación entre Estados, lo que hará preciso mecanismos y acuerdos sobre el control de armamento. Nuevas armas, mayor riesgo y, cuanto mayor sea el riesgo, más necesaria es la prevención y anticipación para logar la alerta temprana frente a ataques, que cada vez tienen más capacidad de desestabilizar a un Estado. Una solución pasaría por la limitación, cuando fuera posible, a usuarios no idóneos de este tipo de armas o de otros sistemas no convencionales. Esto se sabe que no funciona así. Cosas de la tecnología, que antes o después termina por ser accesible, quizá no la última, pero si la inmediatamente anterior. Incluso los propios datos biométricos serán objeto de transformación. La recogida de estos datos afinará su precisión, de modo que incrementará la fiabilidad de la propia IA, siempre que no desaparezca la transparencia, que, en el fondo, es lo que garantiza la seguridad.

Bibliografía

Asimov, I. (1942). *Runaround*. EE. UU., Astounding Science-Fiction.

Boyd, A. (2020). Intel Agencies Seek to perfect biometric recognition from drones. *Netxgov*.

Bostron, N. (2014). *Superintelligence. Paths, Dangers, Strategies*. Oxford, Oxford University Press.

Boix Palop, A. (2007). De McDonald's a Google: la ley ante la tercera revolución productiva. *Teoría y Derecho: revista de pensamiento jurídico*. 1, pp. 124-147.

Comisión Europea. (2019). *Directrices éticas para una IA fiable* [en línea]. Bruselas, Unión Europea. [Consulta: 2025]. Disponible en: https://data.europa.eu/doi/10.2759/14078

Consejo De Seguridad De Naciones Unidas. (2017). *Resolución 2396/2017*.

Cotino Hueso, L. (2019a). Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y *big data* confiables y su utilidad desde el Derecho. *Revista catalana de dret públic*. 58.

—. (2019b). Riesgos e impactos del *Big data*, la inteligencia artificial y la robótica: enfoques, modelos y principios de la respuesta del derecho. *Revista general de Derecho administrativo*. 50.

Delclos, T. (2007). Japón y Corea del Sur preparan leyes para regular la conducta de los robots [en línea]. *El País*. [Consulta: 4 de mayo 2022]. Disponible en: https://elpais.com/diario/2007/04/19/ciberpais/1176950126_850215.html

Eidenmüller, H. (2017). *The Rise of Robots and the Law of Humans* [en línea]. Oxford, University of Oxford. [Consulta: 2025]. Disponible en: https://www.law.ox.ac.uk/business-law-blog/blog/2017/04/rise-robots-and-law-humans

Fernando Pablo, M. M. y Terrón Santos, D. (2019). Sobre la gobernanza de la inteligencia artificial. En: Guayo Castiella del, I. y Fernández Carballal, A, (coords.). *Los desafíos del derecho público en el siglo XXI: libro conmemorativo del XXV aniversario del acceso a la Cátedra del Profesor Jaime Rodríguez-Arana Muñoz.* Insitituto Nacional de Administración Pública.

Gil, A. (2021). Bruselas quiere prohibir sistemas de identificación biométrica remota en espacios públicos` en su regulación de la inteligencia artificial. *El Diario*.

Hinton, C. H. (1912). *The fourth dimension* [en línea]. Kessinger Press. [Consulta: 2025]. Disponible en https://archive.org/details/fourthdimensionoohintarch/page/n5/mode/2up

Hoffmann-Riem, W. (2018). *Big Data. Regulative Herausforderungen*. Nomos, Baden-Baden.

Illanas García, L. (2024). Fundamentos históricos de la biometría aplicada a la defensa y sus planteamientos éticos. *Historia Actual Online*. 63(1), pp. 199-212.

Innerarity, D. (2022). Igualdad algorítmica. *El País*.

Kaplan, A. y Haenlein, M. (2018). Siri, Siri in my Hand, who's the Fairest in the Land? On the Interpretations, Illustrations and Implications of Artificial Intelligence. *Business Horizons*. Bloomington, Indiana University. 62(1), pp. 15-25.

Sayle, K. (2021). *Biometric technologies and global security*. Congressional Research Service.

O´Neil, C. (2016). *Weapons of Math Destruction, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy.* Nueva York, Broadway Books.

Parlamento Europeo. (2021a). *Report on artificial intelligence: questions of interpretation and application of international law in so far as the Eu is affected in the areas of civil and military uses and of state authority outside the scope of criminal justice*.

—. (2021b). *Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión (COM(2021)0206)*.

—. (2022). *Resolución de 3 de mayo de 2022, sobre la inteligencia artificial en la era digital (2020/2266(INI))*.

Pasquale, F. (2015). *The Black Box Society. The Secret Algorithms That Control Money and Information*. Cambridge, Harvard University Press.

Queirolo Pellerano, F. (2019). Sistemas de armas autónomos letales (LAWS). Reflexiones para un debate [en línea]. *Revista Política y Estrategia*. 134, pp. 147-170. [Consulta: 2025]. Disponible en: https://doi.org/10.26797/rpye.v0i134.790

Terrón Santos, D. (2022). *Administración inteligente y automática. Una visión más allá del algoritmo*. A Coruña, Colex.

Turing, A. M. (2012). *¿Puede pensar una máquina?*. Oviedo, KRK.

Artículo recibido: 10 de octubre de 2024
Artículo aceptado: 14 de enero de 2025