

DERECHO E INTELIGENCIA ARTIFICIAL EN LA UNIÓN EUROPEA. ¿CENTRALIDAD U OBSOLESCENCIA DEL SER HUMANO?

Law and artificial intelligence in the European Union.
Centrality or obsolescence of the human being?

VICTORIANO GALLEGO ARCE

Centro Universitario San Isidoro

vgallegoa@gmail.com

Cómo citar/Citation

Gallego Arce, V. (2024).

Derecho e inteligencia artificial en la Unión Europea.

¿Centralidad u obsolescencia del ser humano?

Revista Española de Derecho Constitucional, 132, 169-197.

doi: <https://doi.org/10.18042/cepc/redc.132.06>

Resumen

La inteligencia artificial se presenta como un prometedor campo de beneficios y posibilidades, pero también como el escenario en el que aparecen riesgos y vulnerabilidades para el ser humano que pueden socavar su esfera de derechos y libertades. No se trata de algo forzado o violento. La lesión a la dignidad y a los derechos inviolables de la persona se produce por la vía del consentimiento inconsciente que dejamos tras un rastro de datos que, debidamente tratados y procesados, pueden llegar a manipular y condicionar la vida y el pensamiento de las personas.

Parece existir la unanimidad global de que los algoritmos deben servir a la humanidad y no al contrario. De ahí la necesidad de una regulación global y centrada en la persona.

La Unión Europea está dando pasos pioneros en esa dirección mediante la aprobación de una norma general y vinculante para todos los Estados miembros. ¿Será suficiente?

Palabras clave

Inteligencia artificial; regulación; derechos fundamentales; centralidad; Unión Europea; riesgos.

Abstract

Artificial Intelligence is a promising field of benefits and possibilities. And at the same time, it is the scenario in which risks and vulnerabilities appear for human beings, which can undermine the sphere of human freedoms. It is not something forced or violent. The injury to the dignity and inviolable rights of the person occurs through the unconscious consent that we leave behind a trail of data. Thus these are properly treated and processed, they could manipulate and condition lives and thoughts.

It is unanimously accepted that algorithms should serve humanity and not the other way around. For this reason, global and person-centered regulation is necessary.

The European Union is taking pioneering steps in this direction by approving a general and binding regulatory text for all member states. Will it be enough?

Keywords

Artificial intelligence; regulation; fundamental rights; centrality; European Union; risks.

SUMARIO

I. INTRODUCCIÓN. II. EL PROBLEMA DE LAS CONSECUENCIAS IMPREDECIBLES: 1. ¿Qué vulnerabilidades presentan los sistemas de IA para los derechos fundamentales de las personas? III. EL ESTUDIO DE LA CENTRALIDAD DE LA PERSONA. IV. LA INICIATIVA LEGISLATIVA DE LA UNIÓN EUROPEA: 1. Aproximación inicial y antecedentes. 2. Aspectos más importantes del futuro reglamento. 3. Novedades introducidas en la tramitación en el Parlamento Europeo. V. A MODO DE CONCLUSIÓN. BIBLIOGRAFÍA.

I. INTRODUCCIÓN

Estamos asimilando con tanta rapidez este fenómeno disruptivo y distópico de la inteligencia artificial que nos referimos a él con el término familiar de «la IA», aunque todavía no hayamos tenido el tiempo suficiente para percibir con claridad el alcance de esta revolución y, mucho menos, ponernos a resguardo de las vulnerabilidades que brotan de estas asombrosas capacidades de hacer, decidir, crear o generar. Unas capacidades que se incrementan de manera exponencial proyectándose sobre todos los ámbitos y sectores de la vida pública y de la vida privada.

Es difícil encontrar en la historia de la humanidad una fuerza transformadora tan amplia, transversal y a la vez tan atractiva como la IA, porque nos seduce con la ilusión de encontrar cómodas y rapidísimas respuestas a los problemas diarios, simulando el proceder inteligente del ser humano y, en ocasiones, retándonos a distinguir entre la mente humana o la máquina, sometándonos a un cada vez más difícil test de Turing.

Conocemos, por su notoriedad, las opiniones optimistas o catastrofistas en sus diversos grados que, en los últimos años, han expresado personalidades de la más variada solvencia y que abarcan toda una amplia gama de gustos¹.

¹ Entre los muchos ejemplos, el del ilustre científico ya desaparecido Stephen Hawking. Este afirmó, entre otras cosas, que «el desarrollo de la inteligencia artificial podría significar el fin de la raza humana». Lo dijo en una entrevista en la BBC el 11 de diciembre de 2014. Unos meses antes de su fallecimiento, en noviembre de 2017, en una entrevista con *Cambridge News*, manifestó que próximamente la inteligencia artificial alcanzaría un nivel que llegaría a ser «una nueva forma de vida que superará

No son más que las voces de expertos que nos alertan y las de otros, igual de expertos, que nos tranquilizan sobre el avance de estas impactantes tecnologías.

No obstante, casi a diario, se incorporan nuevas aplicaciones de estos sistemas que emiten inquietantes señales de despersonalización y que despiertan la lógica atención (o alerta) en el jurista, porque parecen desvirtuar la esencia misma del ser humano. Nos referimos a esa tecnología robotizada capaz de decidir, mostrar empatías, emociones o incluso capacitada para expresar sensibilidades sociales, como si se tratara de un ser habiente de conciencia e identidad. Pero desprovista de la capacidad de ser consciente (Schneider, 2021).

En otras ocasiones escuchamos términos y conceptos como la integración hombre-máquina, la singularidad o el transhumanismo², como si tuviéramos que asimilar que la mente humana no es más que un programa de *software* o

a los humanos». «Temo que la inteligencia artificial podría reemplazar a los humanos por completo [...]. Si las personas diseñan virus informáticos, alguien diseñará una IA [inteligencia artificial] que mejore y se replique a sí misma [...]. Esta será una nueva forma de vida que superará a los humanos. [...] El punto de no retorno ya lo hemos traspasado. El planeta se está volviendo demasiado pequeño para nosotros, el crecimiento de la población es alarmante y corremos el riesgo de exterminarnos entre nosotros mismos». Disponible en: <https://tinyurl.com/mrxzuc4p>.

En parecidos términos, Nick Bostrom: «Debemos ser capaces de saber cómo controlar la inteligencia artificial». Disponible en: <https://tinyurl.com/4c38bcwn>.

También han realizado manifestaciones parecidas Bill Gates o Steve Wozniak. *Vid.* <https://tinyurl.com/bdebfftv>.

Vladimir Putin afirmó, el 30 de mayo de 2019, que el país que logre el monopolio de las tecnologías de inteligencia artificial será dueño del mundo. *Vid.* <https://tinyurl.com/28pxarucl>.

Como ejemplo de opinión distinta, Yann LeCun manifiesta que la gente está «hiper-ventilando sobre el riesgo de la IA» al creer que la puesta en marcha de los sistemas de IA se producirá de una manera repentina, sin que podamos protegernos de sus consecuencias. Él considera que tendremos tiempo de adaptarnos a la nueva tecnología de manera ordenada. Disponible en: <https://tinyurl.com/yfyv2uxa>.

En otro orden de cosas, en marzo de 2023, el propietario de Tesla, Elon Musk, el cofundador de Apple, Steve Wozniak, y un gran número de CEO, CTO e investigadores de IA firmaron una carta para solicitar una moratoria de seis meses en el desarrollo de la IA al considerar que estos sistemas suponen un profundo riesgo para la humanidad y para la sociedad. Disponible en: <https://tinyurl.com/ybzd4xc3>.

² El 30 de enero de 2024, el empresario Elon Musk hizo pública la implantación de un chip en el cerebro de una persona llevada a cabo por su empresa Neuralink. Un controvertido proyecto que el magnate ha bautizado como *Telepathy* (telepatía). Véase <https://tinyurl.com/yp252264>.

que la tecnología, mediante un ejercicio de simulación, puede crear cosas como lo hace la mente humana.

Los avances tecnológicos siempre nos remiten, a modo de advertencia, al denominado «principio de Skolnikoff» (Skolnikoff, 1994) según el cual los usos que hacemos de una determinada tecnología no siempre coinciden con el propósito para el que fue creada o se había previsto originalmente en su diseño. Si se nos permite el símil, recordemos que el avión de pasajeros nunca fue diseñado ni pensado para ser empleado en la destrucción de un rascacielos y causar la muerte a miles de seres humanos.

Análogamente, mapear las conexiones neuronales de una persona puede permitir resolver enfermedades mentales, lesiones funcionales u otros problemas complejos que hoy hacen difícil la vida de muchas personas, pero, de la misma forma, también podría ser utilizado para acceder al conocimiento de las ideas y pensamientos de otras muchas.

Por otra parte, a la hora de hablar de los beneficios y vulnerabilidades de los avances tecnológicos, también resulta obligado detenerse en el conocido «dilema Collingridge», según el cual, en las primeras etapas de la innovación tecnológica, las consecuencias de su implantación no pueden predecirse de forma clara y fehaciente, porque existe un problema de falta de información. Pero, cuando ya disponemos de esos datos, de esa información palpable que nos advierte de las consecuencias negativas que genera, dicha tecnología está tan implantada que ya es muy difícil regularla y controlarla (Collingridge, 1980: 13-21).

Estas novedades tecnológicas nos aportan y aportarán soluciones; decisiones aparentemente acertadas y prácticas para nuestros problemas cotidianos. Eso puede inducirnos a pensar en una vida de absoluta despreocupación ante los problemas, porque, en ese presente que se aproxima, las máquinas atienden y satisfacen todas nuestras necesidades al momento. Otra cosa será cuando los humanos pasemos a ser piezas de museo para ellas.

Los sistemas de IA no pueden concebirse ya como objetos o herramientas para usar por las personas, es decir, como instrumentos neutros, porque estos sistemas deciden, crean, discriminan, observan y utilizan a la persona a través de la huella digital; clasifican y ordenan la sociedad en una especie de guetos virtuales como son las burbujas digitales o son capaces de interferir en la vida política de un Estado. Todo esto se complica por el potencial de autoaprendizaje o aprendizaje automático de estos ingenios que les permiten innovarse a sí mismos.

Por lo tanto, aunque percibamos que estos sistemas son solo en un artificio que «ayuda» al ser humano en su progreso, con luces más largas, puede observarse que tienen capacidad para incidir en las decisiones y para transformar a la persona y a la sociedad.

Al mismo tiempo se percibe que, en todas las estrategias, nacionales e internacionales, así como en los embriones regulatorios sobre la IA, se apela reiteradamente a la ética y al derecho como herramientas necesarias para pilotar este tren de gran velocidad. Pero, sobre todo, se insiste en que esta tecnología debe estar «centrada en el ser humano». ¿Por qué esta insistencia? ¿Tal vez porque somos conscientes de que el uso de estos sistemas puede producir deshumanización? Esta es la piedra angular de todos los problemas ético-jurídicos que plantean estos desarrollos robóticos.

¿Existe unanimidad de pensamiento sobre esta hipotética «deshumanización»? Evidentemente no. En este aspecto resulta muy ilustrativa la clasificación que nos ofrece Gissel Velarde en torno a cuatro grupos de pensamiento: pesimistas, optimistas, escépticos y pragmáticos (2021: 179-184). Dicho de otra manera: los que consideran que esto es el fin, los que dicen que es el futuro liberador, los que no se lo creen y, por último, quienes consideran que hay que aprovechar las ventajas que ofrecen los algoritmos inteligentes siempre que seamos capaces de mantenerlos bajo control a través de una regulación efectiva.

Este último enfoque, pragmático, es la base de partida de nuestro estudio, que dividimos en dos etapas. En un primer momento nos cuestionamos de qué manera esta vorágine tecnológica está afectando (y afectará) a todo el conjunto de derechos y libertades de la persona y que se contienen en el superior concepto de la dignidad humana. Y, en una segunda fase, analizaremos si las iniciativas regulatorias sobre esta materia permiten salvaguardar suficientemente este aspecto central del ser humano.

II. EL PROBLEMA DE LAS CONSECUENCIAS IMPREDECIBLES

Hasta ahora, la humanidad había creado utensilios (útiles para hacer algo) que servían para manufacturar (hacer con las manos), para hacer o facilitar la realización de la tarea que quiere realizar el usuario que maneja la herramienta.

La historia del ser humano también es la de su propio progreso y la de la consecución de hitos alcanzados, haciendo realidad sus sueños y superando, a veces, sus propias ficciones. El humano siempre ha deseado volar, ser el más rápido de los animales o cubrir grandes distancias submarinas. Pero, consciente de que no puede conseguirlo por sus propios medios, ha construido artefactos más o menos complejos que perfecciona continuamente y que le permiten aproximarse al hecho de volar, desplazarse con rapidez o navegar. Son máquinas construidas por la inteligencia humana, fruto del avance tecnológico, manejadas y dirigidas por el propio ser humano.

Sin embargo, lo que ocurre en esta nueva revolución es que se construyen artilugios que superan o superarán la inteligencia humana, dejan de estar bajo control, y apartan la decisión o la creatividad humana, ofreciendo la atractiva posibilidad de no tener que pensar, porque el propio sistema de IA piensa, innova y ofrece soluciones más precisas y perfectas de las que podía ofrecer la mente humana.

Si el propio sistema inteligente es el que analiza la situación, valora, decide y ejecuta tareas, cada día más trascendentales para la persona, para la sociedad y para la seguridad de ambas, el ser humano corre el riesgo de convertirse en un simple beneficiario de la máquina que vive gracias a sus prestaciones. Entregarnos a esa comodidad liberadora tiene el precio de acabar dirigidos por unas máquinas inteligentes que aprenden cada día.

Dicho de otra manera, los humanos acabaríamos siendo pececitos pequeños, congregados en bancos, moviéndonos más o menos caóticamente, pero sin que la persona lidere, decida o exprese su voluntad en la elección del rumbo que debe seguir el conjunto. Esto supone el palmario menoscabo de su autonomía, pilar fundamental de su libertad, así como el riesgo de abandonar aquella religión unamunoniana de «buscar la verdad en la vida y la vida en la verdad».

En definitiva y utilizando las palabras de la UNESCO (2021), «a largo plazo, los sistemas de IA podrían disputar al ser humano el sentido especial de la experiencia y la capacidad de actuar que le son propios». Esto supone un claro empobrecimiento del ejercicio de los derechos y libertades fundamentales si, finalmente, la tecnología olvida que el motivo de su existencia es servir al ser humano sin sustituir su libre albedrío.

1. ¿QUÉ VULNERABILIDADES PRESENTAN LOS SISTEMAS DE IA PARA LOS DERECHOS FUNDAMENTALES DE LAS PERSONAS?

Los estándares de seguridad que demandan estas aplicaciones de la IA deben dirigirse a la protección de la persona frente a amenazas modernas y hasta ahora desconocidas, como son la manipulación cognitiva conductual, la categorización biométrica, el reconocimiento y control de emociones en determinados lugares, la vigilancia policial predictiva o la detección y clasificación de las personas respecto a la ideología política, creencias religiosas u orientación sexual, entre otros, y que suponen vulnerabilidades graves para los derechos y libertades fundamentales.

La desinformación o la difusión de noticias falsas no es causa de los sistemas de IA. Muy al contrario, constituyen un clásico de la manipulación. Lo que ha

ocurrido en las últimas décadas es que la digitalización, las redes sociales y ahora la IA han planteado cambios en la manera de difundir y, sobre todo, de consumir información. La abundancia y el fácil acceso a la información también generan efectos negativos, como son las nuevas formas de manipulación y de censura. Las denominadas *fake news*, los conocidos bulos, la posverdad, el relato que llega y que «le gusta» a cada persona, no pueden entenderse como desinformación, sino, más bien, como deformación mental de la persona³.

Bajo el título de *Infocracia*, Byung-Chu Han (2022: 40-41) nos muestra una situación en la que la información se utiliza como arma, donde las campañas electorales no están basadas en argumentos y razonamientos, donde la verdad y la veracidad pasan a un segundo plano y donde no cabe el discurso porque prevalecen los algoritmos inteligentes y la democracia se hunde en una jungla de información.

El referéndum del *brexit* en el Reino Unido, las elecciones en EE. UU. de 2016, la crisis sanitaria generada por la covid-19 o los mensajes generadores de indignación e incitación al odio muestran continuos ejemplos de que la manipulación y el bulo funcionan mejor en escenarios de incertidumbre y desconfianza creciente en el sistema.

¿Por qué es importante detenernos en esta vulnerabilidad? La pregunta se responde a sí misma: la formación y la información del individuo son la esencia del desarrollo de su personalidad y la herramienta insustituible para crear ciudadanos libres.

El acceso a una información condicionada por el entrenamiento de un *chatbot*, nunca imparcial, se propaga con facilidad y se dirige a aquellos grupos digitales cerrados que son propensos a creer ese contenido informativo. Cuanta más actividad desempeña el usuario a través de estos ingenios, más datos obtiene el propio sistema para configurar su perfil y mejor seguimiento se logra de este; de tal forma que la burbuja en la que vive no le permite la necesaria tarea de contrastar y, consecuentemente, salir del error o confirmar su acierto. Sin información no existen el consentimiento ni la decisión libre.

Los sistemas de IA deben mejorar y facilitar la elección de alternativas, pero nunca tomar decisiones por el humano, sustituirlo o adoptar resoluciones,

³ Mediante adecuados sistemas de IA, pueden procesarse nuestras actualizaciones, notificaciones, búsquedas e inquietudes en la red compartimentado a las personas en las denominadas «cámaras eco» o en «enjambres» donde no cuentan la razonabilidad o racionalidad de lo percibido, sino el contenido que «me gusta» escuchar y percibir, no exento de parcialidad, que, finalmente, influye en las emociones, el conocimiento y la conducta de la persona. Todo un perfecto campo de cultivo de sesgos y manipulaciones.

porque eso supone la destrucción de la autonomía de la persona. La profesora Adela Cortina (2019: 388-389) explica que la autonomía de la persona es algo más que poder tomar decisiones, con criterio propio e independiente de otros (humanos y máquinas); es la capacidad de autonormarse, de determinarse, esto es, de actuar según la convicción y la convención de lo que está bien o no, de lo que se debe hacer o no se debe hacer.

Otro riesgo evidente que emana de las aplicaciones de IA es la posibilidad de reconocimiento facial y de otros métodos involuntarios de identificación a través del uso de datos biométricos recolectados de internet o de la amplia red de cámaras repartidas por los distintos espacios.

Estas herramientas de reconocimiento biométrico permiten identificar a la persona a través de su rostro, mediante el cotejo automático de miles de millones de imágenes almacenadas (Seibt, 2021)⁴, pero también permite identificaciones a través de la voz, realizar evaluaciones de la personalidad o detectar mentiras. Véase la polémica aplicación ClearviewAI, empleada por agencias y organismos gubernamentales de distintos Estados occidentales. Razonablemente, salvo que exista un control legislativo riguroso que permita el empleo supervisado, motivado y proporcionado de estas herramientas, es obvio que deben quedar prohibidas por la inmersión que suponen en el espacio íntimo de las personas.

Ya no es una ficción que China haya desarrollado su propio sistema de reconocimiento facial, que permite, al parecer, detectar y distinguir el rostro de una persona sobre la que pesa una orden de búsqueda y captura, a través de cientos de millones de cámaras, situadas en cualquier punto de cualquier ciudad. Es el sistema Dragonfly Eye. Un algoritmo de IA de la empresa china Yitu que puede reconocer fácilmente cualquier rostro entre los de 2000 millones de personas (Aldama, 2018).

De igual forma, el polémico proyecto piloto denominado «sistema de calificación social», puesto en marcha en algunas ciudades chinas o en el metro de Pekín, permite al ciudadano someterse, de forma voluntaria, a un procedimiento de «suma de puntos» que le posibilita, a su vez, obtener ventajas o beneficios en su vida cotidiana. Aunque hay quien niega su existencia, se trataría, en su caso, de un ensayo pionero con el que se motivaría a la ciudadanía a

⁴ «En unos segundos y con una foto robada, la aplicación [ClearviewAI] ayudó a un hombre millonario a encontrar la identidad de un desconocido que estaba cenando con su hija. También permitió a la Policía identificar a un depredador sexual, que no estaba en las bases de datos oficiales pero cuyo reflejo apareció en un espejo en una foto tomada en un gimnasio. Estos son solo dos ejemplos de uso de este controvertido servicio».

respetar las normas de manera voluntaria, de tal forma que aquellas personas que hacen deporte, que utilizan transporte con bajas emisiones contaminantes, donan sangre, devuelven puntualmente sus libros a la biblioteca, o incluso, depositan su basura de forma ordenada, obtienen ventajas fiscales o de otro tipo. En definitiva, las autoridades chinas quieren saber quiénes respetan las normas, quiénes son obedientes y quiénes se someten dócilmente a ser vigilados.

En el negativo de esa fotografía social se observa que quienes no se sometan a dicha vigilancia pagarán más impuestos, les resultará más costoso acceder a las ayudas o, simplemente, serán objeto de sospecha (Deutsche Welle, 2021). Consecuentemente, el ciudadano se ve obligado a revelar detalles de sus actos personales e individuales que, en el ejercicio de su libertad, tendría derecho a desarrollar sin la observación de nadie.

Ya se ha dicho que, conscientes o no, con consentimiento o sin él, vamos dejando rastros digitales —huellas digitales— en casi todo lo que hacemos: en el hogar a través de los asistentes virtuales, en el coche, en nuestras compras, en nuestros trabajos, en nuestras búsquedas, en nuestras comunicaciones, incluso, en nuestro simple deambular. Esta inmensa cantidad de datos, *big data*, alimenta los distintos sistemas de IA.

De esta forma, se va marchitando el concepto mismo del derecho a la intimidad, porque se va estrangulando progresivamente el flujo de savia que le da vida al propio derecho y que es la protección de su propia individualidad, de lo más interno de la persona y de la libertad de sentir, o no, lo público. El ámbito privado se va disipando sin que vislumbremos qué nuevo concepto vendrá a sustituirlo. Se debilita la percepción de que el derecho a la intimidad otorga a su titular el poder de resguardar ese espacio reservado, personal y también familiar⁵, e incluye la prohibición a terceras personas de divulgar detalles de dicho ámbito reservado sin el beneplácito del titular.

Para que exista vulneración, no se exige que el titular del derecho conozca que su intimidad ha sido vulnerada, pero el daño a la intimidad es mayor si se produce tal desconocimiento⁶. Así, si se concibe la intimidad, entre otras

⁵ SSTC 231/1988, de 2 de diciembre, 197/1991, de 17 de octubre, o 115/2000, de 10 de mayo.

⁶ «El derecho a la intimidad se traduce en un poder de control sobre la publicidad de la información relativa a la persona y su familia, con independencia del contenido de aquello que se desea mantener al abrigo del conocimiento público. Lo que el art. 18.1 CE garantiza es un derecho al secreto, a ser desconocido, a que los demás no sepan qué somos o lo que hacemos, vedando que terceros, sean particulares o poderes públicos, decidan cuáles sean los lindes de nuestra vida privada, pudiendo cada persona

facetas, como una reserva de conocimiento de un ámbito personal y que administra su titular (ámbito «privado»), tal administración y tal reserva se devalúan si el titular del ámbito de intimidad desconoce las dimensiones de este, porque desconoce la efectiva intromisión ajena (Pérez Royo, 2016: 293-294). Tal devaluación es correlativa a la de la libertad y a la de la «calidad mínima de la vida humana» (STC 231/1988, de 2 de diciembre, FJ 3), que posibilita no solo el ámbito de intimidad, sino el conocimiento cabal de este (STC 89/2006, de 27 de marzo, FJ 5).

El tiempo apremia en la resolución de estos problemas jurídicos porque los avances científicos y tecnológicos no esperan y porque ya se vislumbran situaciones futuras (y no tan futuras) a las que el derecho y la ética tendrán que dar una pronta respuesta. Se anuncia que, muy pronto, las personas pueden estar conectadas a internet directamente con el cerebro, mediante gorras o diademas capaces de inyectar en el pensamiento todo el flujo de información que se distribuye por la red (Ansedo, 2022: 27)⁷, ayudándolas a razonar con la misma facilidad que hoy nos ayuda el corrector ortográfico. De ser cierto, supondría que el mismo canal que deposita información en la mente de la persona podría extraerla. Es decir, el acceso a un dominio auténticamente íntimo de la persona al que nadie, absolutamente nadie, había accedido hasta hoy.

Leer el pensamiento implica que la libertad ideológica y de creencias, así como las libertades a través de las cuales se manifiesta y se exterioriza dicha libertad ideológica, como son la libertad de expresión, la de información, la de cátedra o la de creación artística, entre otras, se queda huera. Se ha accedido, precisamente, al pensamiento, baluarte inexpugnable y paradigma de una libertad negativa que ningún grillete podía detener o ningún tercero podía intervenir. La que puede ser, a nuestro juicio, la mayor conquista de las *big techs*.

reservarse un espacio resguardado de la curiosidad ajena, sea cual sea lo contenido en ese espacio. Del precepto constitucional se deduce que el derecho a la intimidad garantiza al individuo un poder jurídico sobre la información relativa a su persona o a la de su familia, pudiendo imponer a terceros su voluntad de no dar a conocer dicha información o prohibiendo su difusión no consentida, lo que ha de encontrar sus límites, como es obvio, en los restantes derechos fundamentales y bienes jurídicos constitucionalmente protegidos» (STC 134/1999, de 15 de julio, FJ 5).

⁷ El neurocientífico Rafael Yuste, catedrático de la Universidad de Columbia, y el ingeniero Darío Gil, director mundial del área de investigación de IBM, alertaron en la residencia del presidente estadounidense de la inminente llegada de un mundo en el que los ciudadanos se conectarán a internet directamente con el cerebro, mediante gorras o diademas capaces de leer el pensamiento. En ese hipotético porvenir, un algoritmo podrá autocompletar la imaginación, como ya hacen los programas informáticos de procesamiento de texto con las palabras.

III. EL ESTUDIO DE LA CENTRALIDAD DE LA PERSONA

Cuando Isaac Asimov escribía aquel famoso cuento del que derivan sus tres conocidísimas leyes⁸, completada con la cuarta ley o ley Zeroth, es como si estuviera pensando en un posible empleo torticero de la IA o, en el mejor de los casos, en un funcionamiento defectuoso de aquella.

La humanidad, a lo largo de los siglos, ha conseguido altos niveles de progreso tecnológico. Pero la protección, la garantía y la salvaguarda de los derechos humanos tampoco han estado exentas de grandes esfuerzos evolutivos. La tecnología, como actividad humana que es, debe ir de la mano del derecho y de la ética. Está obligada a diseñar un progreso que no limite ni restrinja, de forma real o virtual, la autonomía de la persona ni menoscabe la esfera de derechos y libertades de aquella.

Incluso el secretario general de la Organización de Naciones Unidas, en su discurso del 21 de enero de 2022, ante el Pleno de la Asamblea General, calificó las tecnologías emergentes como la cuarta prioridad del mundo, describiéndolas como un ámbito en el que «prácticamente no existe gobernanza mundial», y, considerando que «la humanidad debe estar en el centro de la tecnología», advirtió que esta no debe usar a las personas, sino a la inversa. La tecnología —afirmó— debería ser una herramienta benéfica para todos. En su discurso dejó patente que «nuestra información personal está siendo explotada para controlarnos o manipularnos, cambiar nuestros comportamientos, violar nuestros derechos humanos y socavar las instituciones democráticas».

Tampoco el Consejo de Europa puede estar ajeno a los efectos de esta revolución digital porque su razón de ser es la de garantizar la protección de los derechos humanos, la democracia y el Estado de derecho. También en este escenario digital. Es la razón por la que nace el Comité sobre Inteligencia Artificial, con la misión de elaborar un convenio [marco] sobre el desarrollo, diseño y aplicación de la IA, basado en los estándares del propio Consejo de Europa. Las negociaciones sobre este Convenio se iniciaron en septiembre de 2022, en

⁸ En mayo de 1942, Isaac Asimov publicó uno de sus cuentos de ciencia ficción titulado «Círculo vicioso» (*Runaround*). Las conocidas leyes de Asimov son las siguientes: primera, un robot no puede hacer daño a un ser humano o, por inacción, permitir que un ser humano sufra daño; segunda, un robot debe obedecer las órdenes dadas por los seres humanos, excepto si estas órdenes entrasen en conflicto con la primera ley; tercera, 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, y cuarta, un robot no puede dañar a la humanidad o, por inacción, permitir que la humanidad sufra daños.

enero de 2023 se presentó en Estrasburgo un proyecto marco⁹ y, finalmente, fue firmado el 5 de septiembre de 2024¹⁰.

Se trata del primer convenio internacional jurídicamente vinculante sobre los sistemas de IA, absolutamente compatible con el derecho de la Unión Europea y alineado con ella, que muestra la preocupación por los riesgos para el ser humano, su dignidad, sus derechos y su autonomía, a la vez que recoge el convencimiento de la necesidad de establecer un marco jurídico aplicable a nivel mundial¹¹ para garantizar el respeto de los derechos humanos, del Estado de derecho y las normas jurídicas democráticas en el uso de dichos los sistemas.

De la misma forma, la Organización para la Cooperación y el Desarrollo Económico (OCDE) ha elaborado el documento titulado «Recommendation of the Council on OECD Legal Instruments Artificial Intelligence», para la administración responsable de una IA confiable. Según la OECD¹², entre los cinco principios que han definido este documento, están los valores y la equidad centrados en el ser humano.

También la UNESCO (2021) ha dejado documentada esta idea básica de que cualquier marco normativo debe tener un enfoque global y estar centrado en la dignidad humana y sus derechos fundamentales. Incluso, como ejemplo de cooperación internacional, pueden citarse los resultados de la Alianza Digital Unión Europea-América Latina y el Caribe, como iniciativa conjunta para impulsar una transformación digital centrada en el ser humano (España Digital 2026, 2023).

En la misma línea se han manifestado las políticas públicas estatales. Así, la Casa Blanca propuso en octubre de 2022 el borrador que conducirá a una «Declaración de Derechos sobre Inteligencia Artificial» en cuyo texto se recogen cinco principios y prácticas asociadas con finalidad de proteger a los ciudadanos estadounidenses frente los riesgos que presentan el diseño, uso y despliegue de los sistemas de IA (The White House, 2022). Al igual que las estrategias nacionales de los Estados, desde la Estrategia Nacional de Inteligencia Artificial de España (2020: 3-5) hasta la Estrategia de la Inteligencia

⁹ Véase <https://tinyurl.com/yc6s7uaf>.

¹⁰ Véase <https://tinyurl.com/4445ayt8>.

¹¹ El convenio marco fue adoptado por el Comité de Ministros del Consejo del Consejo de Europa el 17 de mayo de 2024, para los 46 Estados miembros del Consejo de Europa, la Unión Europea y 11 Estados no miembros (Argentina, Australia, Canadá, Costa Rica, Estados Unidos de América, Israel, Japón, México, Perú, la Santa Sede y Uruguay)

¹² Véase <https://tinyurl.com/yeyvdt6v>.

Artificial de Japón (García, 2021: 6) entienden que esta explosión de desarrollo tecnológico debe estar centrada en el ser humano.

Finalmente, es preciso citar que, el 10 de enero de 2023, representantes de las tres religiones abrahámicas¹³ asistieron en el Vaticano al evento «AI Ethics: An Abrahamic commitment to the Rome Call», que condujo a la posterior firma de una declaración conjunta en la que, sucintamente, puede decirse que se llama a todas las partes interesadas a un desarrollo de la IA basado en la ética (algorética), en una regulación legal internacional de esta novedad tecnológica y en la protección de la dimensión humana (antropocentrismo).

Por lo tanto, existe una insistente y generalizada concordancia respecto a la necesidad de la puesta en marcha de límites éticos y normativos, del establecimiento de líneas rojas para que el desarrollo de esta revolución tecnológica siempre sitúe a los seres humanos en el centro de su ámbito de aplicación, y, tal y como la Eurocámara señala con frecuencia en sus escritos, la Unión y sus Estados miembros tienen una responsabilidad particular a la hora de velar por que la IA, la robótica y las tecnologías relacionadas —que pueden tener un uso transfronterizo— estén centradas en el ser humano, es decir, que se conciban básicamente para ser utilizadas al servicio de la humanidad y del bien común, a fin de contribuir al bienestar y al interés general de sus ciudadanos (2020/2013[INI]).

1. ¿CÓMO ENTENDEMOS LA CENTRALIDAD EN EL SER HUMANO?

La IA ha creado una realidad que opera en un escenario espacial que no tiene un «espacio», porque no existe un espacio físico, sino que se desarrolla en el ámbito cibernético por el que transitamos sin desplazamiento físico. La IA no trabaja con cosas ni con conceptos, sino con números y con cálculos. Para su funcionamiento necesita una materia prima necesaria que es la información y, para ello, tiene que cosechar datos.

Aquí está la base nuclear de la problemática que afecta a los derechos de las personas, que, por ser tan evidente, puede que se deje de tener en cuenta: todas nuestras actividades dejan un rastro digital (una huella). Así se constituyen inmensas cantidades de datos que el algoritmo procesa y, en ese procesamiento, se generan decisiones que acaban afectando directa o indirectamente, para bien o no tan bien, a más o menos personas. No necesariamente se trata de «datos personales» —a los que el derecho protege—,

¹³ Monseñor Paglia, presidente de la Pontificia Academia para la Vida, el rabino jefe, Eliezer SimhaWeisz, y el jeque Abdallahbin Bayyah.

sino de datos de la actividad humana que acaban influyendo en las decisiones de las personas.

¿Dónde están los grandes caladeros de datos? Evidentemente, en las redes sociales abiertas, en la interacción social que produce el intercambio de información masiva y cruzada, así como en el acceso a la información libre, que deja rastros valiosos sobre nuestros comportamientos, nuestras preferencias o repudios en los ámbitos económico, social, ideológico, religioso o político. Toda una descripción de nuestro ser y actuar que escanea con precisión, incluso, lo que no vemos o sabemos de nosotros mismos.

Evidentemente, el tratamiento de esos datos e información tendrá una repercusión de gran impacto en la autonomía, la libertad, la intimidad, la igualdad y, en general, los principios y valores que soportan el funcionamiento democrático del Estado de derecho (Tenorio Cueto, 2021: 121), porque mediante el análisis de esa información se puede predecir y conducir el comportamiento de una persona o de un grupo.

En nuestras zonas de retiro no virtuales disfrutamos de aquellos contenidos privados e íntimos que no queremos compartir ni publicar, porque sabemos protegerlos frente a los tradicionales métodos de invasión (cerrando la puerta o bajando las persianas). Esto es, mantenemos el adecuado control sobre los mecanismos de preservación de la intimidad y de la autonomía. Sin embargo, en los espacios virtuales asistimos a penetraciones vulneradoras de los derechos y libertades incluso con nuestro consentimiento, no siempre conscientemente otorgado. Todo ello, a pesar del importante peso que tuvo el principio del consentimiento en la constitucionalización europea de la protección de datos personales, que se encuentra limitado y en crisis en la era digital (de Giovanni, 2022: 51).

De esta forma, se libera la información relativa a las vidas de los ciudadanos, que, posteriormente tratada, conduce a que los algoritmos lleguen a predecir decisiones que deben adoptarse de conformidad con sus gustos, aficiones o preferencias de consumo, proporcionando a la persona una sensación de comodidad, seguridad e inmediatez (Tenorio Cueto, 2021: 121) sin percibir que, a cambio, entrega una parte esencial de la condición humana, que es el libre albedrío.

De ahí que el ser humano deba ser cuidadoso y advertir posibles riesgos para su libertad y su progreso. Unos riesgos que se inoculan a través del *like* o amén digital y que encadenan al ser humano sin que este se resista, porque estos sistemas inteligentes no utilizan la imposición, la coacción ni la represión, sino la persuasión y la permisión, para la dominación, sin que seamos conscientes de ello (Byun-Chul, 2021: 39-41).

En la larga historia de la humanidad, lo único que ha permanecido inmutable es la capacidad insaciable de progreso y evolución que siempre ha mostrado el ser humano. Pero sin desprenderse de su esencia y condición. La capacidad de progresar y de innovar de la persona es consustancial con su inteligencia y con la reflexión ética sobre la aplicación de dicha inteligencia en beneficio de la humanidad.

Pero, al hablar de ética, nos referimos a la ética que deben desplegar los humanos que participan en la implementación de estos sistemas inteligentes. Porque se ha llegado a decir que pueden introducirse en estas máquinas unos códigos éticos programados como la panacea para solucionar los riesgos o vulnerabilidades de la IA. El problema es que las máquinas no son éticas. Éticas solo pueden ser las personas, ya que tienen la exclusividad de tener la «conciencia» y la «consciencia» de saber y comprender por qué está bien lo que está bien y por qué está mal lo que está mal.

Tal y como nos enseña Hans Jonas (1995: 29), lo que presenta relevancia ética es el trato directo de la persona con la persona, aunque sea un trato consigo misma. Pero la actuación sobre los objetos no humanos no constituye un ámbito de relevancia ética, porque esta concepción de la ética es antropocéntrica (*ibid.*). De ahí que el problema al que debe hacer frente la regulación de estos sistemas sea la despersonalización del hombre por la máquina. El problema es que esta tecnología no se ha construido por la persona para la persona, sino para imitarla y para sustituirla.

IV. LA INICIATIVA LEGISLATIVA DE LA UNIÓN EUROPEA

1. APROXIMACIÓN INICIAL Y ANTECEDENTES

Europa ha afrontado el reto de poner en marcha una novedosa y pionera regulación, sabedora de que la IA es «una tecnología estratégica que ofrece numerosas ventajas a los ciudadanos, las empresas y la sociedad en su conjunto, siempre y cuando sea antropocéntrica, ética y sostenible y respete los derechos y valores fundamentales» (COM/2020/65 final/2. 30).

La Unión Europea, concedora de la posición estratégica que ocupan otras potencias en esta materia (25.4.2018 COM [2018] 237 final. 4-5)¹⁴,

¹⁴ El Gobierno de Estados Unidos presentó una estrategia sobre inteligencia artificial e invirtió cerca de 970 millones de euros en investigación no clasificada sobre inteligencia artificial en 2016. Con su «Plan de desarrollo de la inteligencia artificial de próxima generación», China ambiciona alcanzar el liderazgo mundial en 2030 y está

consciente de que o lidera o es liderada (Pedreño y Moreno, 2020: 23), trata de reforzar su posición en materia de IA a escala mundial, sobre la base de los principios que conforman su creación. En los últimos años, distintas instituciones de la Unión Europea han desarrollado un amplio conjunto de acciones, propuestas y resoluciones de *softlaw*, siempre centradas en el ser humano y con la finalidad de que la progresiva implantación de estas modernas tecnologías se desenvuelva en un clima de confianza en el derecho y en la tecnología; con seguridad, transparencia, protección contra el sesgo y la discriminación, garantizando, además, el derecho a la reparación, la responsabilidad y el respeto a los derechos fundamentales de la persona. Lógicamente, los frutos de estas iniciativas siempre procurarán (y necesitarán) la debida coordinación y armonización con todos los Estados miembros.

El inicio de esta hoja de ruta que ha seguido la Unión Europea podríamos situarlo en la iniciativa del Parlamento Europeo que condujo a la creación, en junio de 2020, de la Comisión Especial sobre Inteligencia Artificial en la Era Digital, a la que se le encomendaron unas misiones estrictamente definidas (2020/2684 [RSO])¹⁵. De entre esas misiones se subraya la presentación al Parlamento de una evaluación que definía los objetivos comunes de la Unión Europea a medio y largo plazo y que incluía los principales pasos necesarios para alcanzarlos. Todo esto se desarrollaría sobre la base de cuatro instrumentos no vinculantes, que son los siguientes: «Dar forma al futuro digital de Europa» (COM/2020/67 final); «Una estrategia europea de datos» (COM/2020/66 final); «Libro Blanco sobre inteligencia artificial: un enfoque europeo de la excelencia y la confianza» (COM/2020/65 final), y el «Informe sobre las implicaciones de seguridad y responsabilidad de la inteligencia artificial, la internet de las cosas y la robótica» (COM/2020/ 64 final).

A este cuarteto documental, que no agota, ni mucho menos, las aportaciones de la Unión Europea a la regulación de la IA¹⁶, se incorpora la pionera

efectuando ingentes inversiones. Otros países, como Japón y Canadá, también han adoptado estrategias en relación con la inteligencia artificial.

¹⁵ Se trata de una comisión de 33 parlamentarios para el estudio y el análisis de la inteligencia artificial en la economía de la Unión, debiendo presentar recomendaciones a través de las comisiones permanentes competentes del Parlamento.

¹⁶ Entre otros, destacamos a modo de ejemplo los siguientes textos:

- Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012), porque resalta la necesidad de formar en el impacto ético y jurídico de la IA, y se recoge

propuesta legislativa presentada por la Comisión Europea, denominada Reglamento del Parlamento Europeo y del Consejo 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, de 21 de abril de 2021 (COM [2021] 206 final), que, sin duda, contribuirá a dibujar un marco regulatorio horizontal en la Unión sobre la IA, así como a mejorar las consideraciones comunes que ya existen sobre la materia, porque entre los fines de esta propuesta están la garantía de la protección de los derechos fundamentales y la seguridad de los usuarios, a fin de que haya confianza en el desarrollo y adopción de estos sistemas de IA.

Obsérvese que el instrumento jurídico que se eligió fue el reglamento, que, como sabemos, presenta un alcance general, uniforme, obligatorio y directamente aplicable en todos los Estados miembros (TFUE, art. 288), lo que viene a indicar, a nuestro juicio, la preocupación de la Unión Europea por dar un paso más decidido en aras de abordar (más allá de los principios éticos y sin separarse de ellos) las ventajas y los riesgos que conlleva la IA para proporcionar seguridad jurídica y confianza en la inteligencia artificial, desde la base del respeto a los derechos fundamentales y valores de la Unión (COM [2021] 206 final.4) y desde una perspectiva basada en el uso de estos sistemas y en los riesgos asociados.

El 14 de junio de 2023, el Parlamento Europeo adoptó su posición sobre el proyecto de reglamento¹⁷, tras la cual se iniciaron las conversaciones para consensuar el texto definitivo de la norma. Después de las arduas negociaciones celebradas en los primeros días de diciembre, Consejo y Parlamento llegaron a un acuerdo provisional que promete conducir a una aprobación definitiva de las normas armonizadas en materia de IA, el denominado «Reglamento

una propuesta de reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de estas tecnologías.

- 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), en la que se considera que estos ingenios inteligentes y autónomos, capaces de pensar y de discernir, pueden generar efectos perjudiciales personales y sociales.
- Carta Europea sobre el Uso Ético de la Inteligencia Artificial en los Sistemas Judiciales y su Entorno, (CEPEJ, 2018), que proporciona directrices éticas dirigidos a políticos, legisladores y operadores del derecho para elaboren y apliquen tecnologías avanzadas en el ejercicio de la función judicial con sujeción a cinco principios éticos básicos que deben aplicarse cuando exista un tratamiento automatizado de decisiones.

¹⁷ Con 499 votos a favor, 28 en contra y 93 abstenciones.

de Inteligencia Artificial». Hasta la adopción del visto bueno por el Comité de Representantes Permanentes (COREPER), el pasado 2 de febrero, el texto se perfeccionó técnicamente, y tras haber pasado por las comisiones de Mercado Interior y Libertades Civiles del Parlamento Europeo, el 13 de febrero de 2024, fue aprobado por la Eurocámara el 13 de marzo de 2024 y, finalmente, publicado el 12 de julio en el *Diario Oficial de la Unión Europea (DOUE)*¹⁸, con entrada en vigor el 2 de agosto, aunque con determinadas excepciones especificadas en el art. 113.

2. ASPECTOS MÁS IMPORTANTES DEL RECIENTE REGLAMENTO

Esta normativa tiene por objeto garantizar la seguridad, el respeto de los derechos y libertades, así como los valores de la Unión Europea, de todos los sistemas de IA que se utilicen en el territorio propio de la Unión o se introduzcan en el mercado europeo. Eso implica que este proyecto normativo debe conciliar ámbitos con intereses muy distintos, como son el funcionamiento de los servicios públicos, los derechos y libertades fundamentales, la confianza de los usuarios o el funcionamiento mismo del mercado. Su base jurídica son los arts. 16¹⁹ y 114²⁰ del Tratado de Funcionamiento de la Unión Europea, e, inicialmente, fijó como objetivos específicos los siguientes (COM (2021) 206 final):

- garantizar que los sistemas de IA introducidos y usados en el mercado de la UE sean seguros y respeten la legislación vigente en materia de derechos fundamentales y valores de la Unión;
- garantizar la seguridad jurídica para facilitar la inversión e innovación en IA;
- mejorar la gobernanza y la aplicación efectiva de la legislación vigente en materia de derechos fundamentales y los requisitos de seguridad aplicables a los sistemas de IA, y

¹⁸ Véase <https://tinyurl.com/3hskse29>.

¹⁹ En cuanto a la protección de las personas en relación con el tratamiento de los datos personales que restringen el uso de sistemas de IA para la identificación biométrica remota «en tiempo real» en espacios de acceso público con fines de aplicación de la ley.

²⁰ Porque trata la adopción de medidas para garantizar el establecimiento y el funcionamiento del mercado interior.

- facilitar el desarrollo de un mercado único para hacer un uso legal, seguro y fiable de las aplicaciones de IA y evitar la fragmentación del mercado.

Desde sus inicios se establecieron cuatro categorías de usos de IA en función del riesgo que generan para los ciudadanos.

En primer lugar, los sistemas que encierran un «riesgo inaceptable» (relacionados en el art. 5, están sujetos a una absoluta prohibición, porque llevan consigo un gran quebranto para los derechos y libertades de las personas, incluida su integridad física o su salud, aprovechando técnicas subliminales de manipulación, las vulnerabilidades de grupos específicos de personas, el empleo de identificación biométrica en determinados supuestos o las clasificaciones de las personas. En definitiva, que conducen (o pueden conducir) a tratos desfavorables.

También están los sistemas de «alto riesgo», que son los que suponen el mayor peso del articulado; son sistemas que pueden provocar riesgos o daños para la salud, para la seguridad, para los valores y derechos fundamentales en la Unión o el medio ambiente. Un riesgo que, según el propio texto inicial, se aprecia en función de dos variables: la finalidad prevista del sistema conforme a la legislación vigente relativa a la seguridad de los productos y las modalidades de uso que efectivamente se le den.

Dada su potencialidad de riesgo, la salida al mercado de estos sistemas de riesgo alto debe cumplir los criterios regulados en el recién aprobado reglamento a través de una técnica armonizadora de nuevo enfoque que se basa en la existencia de unos requisitos esenciales obligatorios que deben ser respetados por el producto y que se fijan directamente por el acto legislativo que regula las especificaciones técnicas que facilitan la justificación de la conformidad del producto con dichos requisitos esenciales y, por último, por la existencia de controles que permiten acreditar que el producto cumple con dichos requisitos obligatorios (Álvarez y Tahiri, 2023). De esta forma, se ha incluido la necesidad de que estén sujetos a una evaluación de impacto en los derechos fundamentales de la persona.

Los sistemas con «riesgo limitado» se denominan así porque dicho riesgo se genera por falta de transparencia exigible, y, por ello, se exigen actuaciones tendentes a garantizar dicha transparencia y responsabilidad.

Finalmente, una última categoría residual, sistemas de «riesgo mínimo»; pueden ser sistemas de uso libre y, por tanto, exentos de las obligaciones contempladas en la propuesta de reglamento, que no quiere decir que se les exima de las obligaciones generales de seguridad de los productos ni que no se promueva la aplicación voluntaria de códigos de conducta.

La norma añade la creación de autorizades de control y, lógicamente, un régimen sancionador.

3. NOVEDADES INTRODUCIDAS DURANTE LA TRAMITACIÓN EN EL PARLAMENTO EUROPEO

Sobre la citada propuesta presentada por la Comisión, de 21 de abril de 2021, el Parlamento Europeo introdujo 771 enmiendas el 14 de junio de 2023, que no se pretenden abarcar en su totalidad, porque ello excedería con mucho el objeto de este trabajo.

Entre las modificaciones que presentan interés para este tema de estudio, destaca el art. 3, dedicado a las definiciones, que sufre importantes modificaciones cualitativas y cuantitativas. Una abundancia de definiciones —se añadieron una veintena a las cuarenta y cuatro existentes— que proporciona más seguridad, más claridad y más uniformidad. Así ocurre con la precisión de la definición de «sistema de inteligencia artificial», más acorde con la de la Organización para la Cooperación y el Desarrollo Económico (OCDE), como «un sistema basado en máquinas diseñado para funcionar con diversos niveles de autonomía y capaz, para objetivos explícitos o implícitos, de generar información de salida —como predicciones, recomendaciones o decisiones— que influya en entornos reales o virtuales». Sin perder de vista que, como ocurre en general con toda esta materia, estas definiciones pueden quedar sobrepasadas por los avances tecnológicos, por lo que exige un plus de atención y actualización (Barrio, 2023).

Además, tras el acuerdo provisional alcanzado en los primeros días de diciembre de 2023, los matices técnicos introducidos antes del referendo del Comité de Representantes Permanentes y su ratificación por las comisiones parlamentarias, es evidente que uno de los aspectos que acapara más interés regulador es el uso de la IA con finalidad de categorización biométrica, «en tiempo real» y «en diferido», conocido como identificación biométrica remota (RBI), utilizada en espacios de acceso público.

Se trata de sistemas que permitirían, mediante el análisis de datos físicos y emocionales de los sujetos, un control masivo de las personas en los ámbitos de la aplicación del derecho y la gestión de fronteras, en lugares de trabajo y en centros educativos (Barrio, 2023), o que agrupan a las personas que presentan características predefinidas de índole política, de orientación sexual, relativas a las creencias religiosas o incluso conexas con la raza. Pero, como aspectos positivos, la biometría permite excelentes aplicaciones beneficiosas para la sociedad y la seguridad de las personas, como es la identificación de

elementos terroristas o la prevención de accidentes de tráfico mediante la detección de conductores que generen riesgos a la seguridad vial.

Finalmente, parece que el empleo de estos sistemas en la identificación biométrica quedará autorizado en determinadas actuaciones de seguridad ciudadana, que, lógicamente, deberán gozar de las necesarias garantías mediante la debida autorización judicial o administrativa con carácter previo.

De igual manera, será necesario establecer medidas regulatorias sobre los sistemas de IA que, de una forma u otra, manipulan la voluntad de las personas, atrayéndolas a espacios virtuales en los que se las direcciona hacia decisiones políticas, de consumo o de posicionamiento social a través de un falseamiento de la realidad, la denominada «disonancia cognitiva», o «cámaras eco». En definitiva, algunos ejemplos de los denominados «riesgos inadmisibles».

También encontraremos que en el texto publicado se abordan con más rigor los denominados «sistemas de IA de uso general», y, específicamente, cuando estos sistemas se integran en sistemas de alto riesgo y se amplían los sistemas de alto riesgo, incluidos los utilizados por las grandes plataformas de redes sociales para recomendar contenidos a los usuarios, además de intensificar los requisitos técnicos y establecer la obligatoriedad de realizar una evaluación del impacto que un sistema de alto riesgo puede generar en los derechos fundamentales de las personas con unas exigencias específicas que se incorporan en el art. 27²¹.

En definitiva, un gran impulso a la necesidad reguladora que busca el punto de encuentro entre la arrolladora revolución tecnológica y los derechos y libertades de la persona; que desprende un incremento de la preocupación por las vulnerabilidades a las que estamos expuestos los seres humanos en el uso de estas tecnologías algorítmicas. Pero consciente de que es muy difícil regular al mismo ritmo que unos avances tecnológicos que no siempre tienen que ser positivos para el progreso del ser humano. Sabemos que, aunque toda mejora es un progreso, no todo progreso supone, *per se*, una mejora.

En este sentido, el legislador europeo ha ido actualizándose durante la tramitación de la norma con la incorporación de nuevos artículos, las modificaciones de muchos otros, o el establecimiento de los denominados «espacios controlados de pruebas», en los que los sistemas pueden ser ensayados antes de ser utilizados y comercializados.

Para que el cumplimiento de esta norma de IA sea efectivo en todos los Estados miembros, también se ha creado la Oficina Europea de Inteligencia Artificial, con las misiones, entre otras, de fiscalización de los modelos de IA de propósito general, investigar infracciones y elaborar códigos de prácticas.

²¹ Se introdujo a través de la enmienda 413.

Esta Oficina se sitúa dentro de la Comisión Europea y es dependiente de la Dirección General de Redes de Comunicación, Contenidos y Tecnología. La actividad de esta entidad se complementa con la del Comité Europeo de IA o la Autoridad Nacional de Supervisión de cada Estado miembro, para fiscalizar la evolución de estos modelos de IA más avanzados, actualizar las normas y garantizar el cumplimiento de estas²².

Sin embargo, nada garantiza que esta norma pueda hacer frente a la evolución de la tecnología, ni siquiera antes de su completa entrada en vigor, prevista, en algunos aspectos, para agosto de 2027.

Recordemos que esta andadura legislativa se inicia en abril de 2021 y, desde entonces, sistemas como ChatGPT, Sora y Dall-E, entre otros, eran ciencia ficción. Y, aunque todavía hoy encontramos fallos y errores en estas novedades, poco a poco se pulirán, y así la humanidad va asistiendo al fin de la realidad.

¿Será suficiente esta norma (junto con el resto de reglamentos europeos de protección de datos y de servicios digitales, entre otros) para conseguir la denominada centralidad de la persona, o son medidas huecas ante el imparable avance de la tecnología?

Nada impide pensar que los entrenamientos de los algoritmos se perfeccionarán y las denominadas «cajas negras» serán cada día menos inaccesibles, lo que implica que se irán asumiendo como buenos todos sus contenidos, determinando las decisiones de las personas, influyendo en sus vidas y acotando el libre albedrío.

Se trata, en resumen, de que las normas reguladoras de este fenómeno sean seguras, sostenibles en el tiempo, protectoras de la autonomía de la persona y de sus derechos fundamentales, para que los arrolladores vientos

²² Definida en el art. 3.1.42 de la propuesta (enmienda 197). El Gobierno español se ha anticipado a la aprobación y entrada en vigor del reglamento y ha dado los primeros pasos mediante la aprobación del Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial mediante el Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial (*BOE* n.º 210, de 2 de septiembre de 2023), así como el impulso de la creación de la Agencia Española de Supervisión de Inteligencia Artificial a través de la Ley 22/2021, de 28 de diciembre, de Presupuestos Generales del Estado para el año 2022, con misión de comprobar la puesta en marcha, uso o comercialización de sistemas que incluyan inteligencia artificial y, especialmente, aquellos que puedan suponer riesgos significativos para la salud, la seguridad, la igualdad de trato y no discriminación, en particular entre mujeres y hombres, y para los demás derechos fundamentales (art. 4.3.e del RD 729/2023).

tecnológicos no hagan zozobrar el lento y seguro (hasta hoy) progreso legislativo. Dicho de otra forma: es vital no olvidar que la persona, su paz, su tranquilidad y su bienestar estén protegidos frente a los usos equivocados o malintencionados de los sistemas de IA. Esa es la misión del legislador.

No hay duda de que la Unión Europea persigue una transformación digital basada en los valores europeos y sin apartar el foco del ser humano, de su dignidad y de los derechos y libertades inherentes a este. Proclama la necesidad de que la tecnología no se utilice para polarizar a las personas, sino para que sirva a las personas y a su bienestar; que los seres humanos puedan elegir los servicios digitales adecuados a sus necesidades y, sobre todo, que reciban la instrucción necesaria y continua para que la evolución tecnológica no les haga perder su condición de ciudadanos libres, marginándolos de nuevo a la posición social de súbditos, esta vez de una máquina.

Las tres instituciones europeas que participaron en este proceso legislativo, Parlamento, Consejo y Comisión, lo han proclamado solamente en una declaración común en el mes de enero de 2023 (2023/C 23/01, *DOUE*, C23/1).

Es posible que el legislador europeo tenga que mantener una vigilancia regulatoria incesante para conseguir un marco jurídico favorecedor de la innovación, a la vez que garante de la esfera de derechos y libertades de las personas. Porque, parafraseando a Daniel Innerarity, en su obra *Política para perplejos* (2018: 9), «el horizonte de lo posible se ha abierto tanto que nuestros cálculos acerca del futuro son especialmente inciertos».

V. A MODO DE CONCLUSIÓN

Razón y ciencia nos acompañan, desde la Ilustración, como herramientas necesarias para lograr el progreso y la plena libertad del individuo. Además, el principio de precaución nos permite detener cualquier avance tecnológico o científico, si no hemos constatado que no supone una vulnerabilidad para la humanidad. Así, la sociedad ha logrado armonizar lo humano con los ingenios que el humano fabrica.

Por otra parte, apelamos a la ética y al derecho para resolver las cuestiones que pudieran menoscabar el irrenunciable modelo de vida humana digna. Eso conlleva la necesaria salvaguarda de los derechos fundamentales de las personas. En definitiva, acompañar el progreso técnico (los avances tecnológicos) con el progreso ético (los intereses generales, la seguridad pública o las libertades democráticas).

La IA no ha venido de fuera a cambiar la sociedad, sino que es la propia sociedad la que ha sido capaz de crear y desarrollar estas herramientas para

satisfacer las nuevas necesidades que surgen del progreso de la humanidad y, a su vez, esa innovación tecnológica acaba transformando la propia sociedad.

Por eso utilizamos el término «revolución», y, por eso, ahora se habla de esta disruptiva Revolución 4.0 o Cuarta Revolución Industrial, caracterizada por el avance de unos sistemas algorítmicos que no solo son capaces de correlacionar cantidades estratosféricas de datos y de resolver con infinita rapidez complejísimas ecuaciones, sino que presentan la capacidad de «aprender a aprender», y, si nos detenemos a reflexionar mínimamente sobre esta afirmación tan simple, sentimos el vértigo de percibir la superación de la mente humana.

La novedad de esta revolución y de los debates éticos y jurídicos que nos plantea radica en que, hasta ahora, la humanidad había creado herramientas que, con más o menos perfección, servían al ser humano para hacer o facilitar la realización de una tarea dirigida y decidida por quien maneja la herramienta. El utensilio era una prolongación del cuerpo humano. Sin embargo, a través de la IA, es la propia herramienta la que puede analizar la situación, valorar, decidir y ejecutar la tarea, sustituyendo la principal misión de quien manejaba dicha herramienta, que es la de decidir qué, cómo, cuándo y dónde hacer de manera autónoma.

La decisión es una cualidad eminentemente humana que centraliza cualquier actuación, porque la persona no puede dejar de ser dueña de esa autonomía. Si la persona no puede dejar de ser dueña de sus decisiones, toda norma reguladora de la IA debe tener como centro de su objeto a la persona. Porque, en el momento en el que el ser humano deje libre el hueco central, algo lo ocupará. Y entonces veremos que la máquina sustituye a la persona.

No se trata de impedir el avance de la tecnología ni desarrollar algo parecido a la «tecnofobia». Se trata de que este cambio paradigmático y trascendental en el uso de las herramientas que fabricamos active todas las disciplinas del saber, para que los beneficios que generan dichos avances sean canalizados hacia el disfrute de toda la humanidad. Tenemos el deber de lograr una IA subordinada a la inteligencia humana y no al contrario. Es decir, nuestra responsabilidad es mantener el control humano total «de» la máquina y no «por» la máquina.

La Unión Europea trata de conciliar todos los aspectos que concurren en este fenómeno y ponerse a la vanguardia de estas novedades, pero es consciente de que ciertos usos de las tecnologías suponen una clara amenaza para la dignidad y los derechos y libertades fundamentales que pueden comprometer el Estado de derecho y el derecho mismo tal y como lo entendemos hoy. Pilares del constitucionalismo europeo (de Giovanni, 2022: 23).

La persona se muestra como el núcleo de la revolución digital en la actividad de las distintas instituciones de la Unión y, por supuesto, en la creación de las normas dirigidas a regular la IA. Pero piénsese que cuando la Comisión presentó la iniciativa de Reglamento de la IA aún no conocíamos la existencia del famoso ChatGPT, que lleva ya más de un año generando consecuencias dignas de todo tipo de valoraciones. Una lección aprendida sobre tempos desacompasados: el de evolución de la técnica y el de la elaboración de las normas que la regulan.

Que la IA debe estar centrada en la persona es unánime. Pero ¿cómo se logra que en este ambiente arrollador pueda mantenerse la promoción del ser humano en todas sus dimensiones, transformando y liberando a la persona? ¿Cómo se garantiza que la norma no se desacople de la evolución de la IA? Y, sobre todo, ¿qué mecanismos permitirán el aseguramiento de su cumplimiento? ¿Pueden evitarse los «usos atípicos»? Mejor aún, ¿pueden detectarse esos usos atípicos?

Entre los siglos I y XVIII, si hubiéramos podido dormir a una persona y despertarla cien o doscientos años después, apenas hubiera tenido problemas de adaptación al nuevo escenario. Es decir, que los acontecimientos novedosos que pudieran haberse producido durante su letargo podían ser absorbidos fácilmente debido a la lentitud del progreso humano. De la misma forma, esa lentitud ha permitido desarrollar unos procedimientos legislativos mesurados, reposados y, por ende, eficaces.

En definitiva, subirse al tren de la materia desregulada y hacerse con los mandos de este hasta conseguir controlarlo no solo no era imposible, sino que era el procedimiento normal. No obstante, la vertiginosa rapidez con la que se desarrolla esta revolución digital dificulta, cuando no frustra, cualquier absorción de novedades para quien pierde, por un momento, el ritmo del progreso. El tren tecnológico de hoy ha mejorado el diseño de tal forma y ha alcanzado tal velocidad que exige igualmente tácticas novedosas para poder subirse a él en marcha.

De ahí que extraigamos cuatro reflexiones:

- i) La tecnología no espera al legislador y, en este caso, la dignidad de la persona y la protección de sus derechos inviolables pueden verse afectados.
- ii) La norma que regule estos sistemas tiene que nacer con una cultura preventiva, con disposición permanente al cambio y con adaptación continua a las circunstancias, de forma proactiva, pragmática y realista, generando confianza en el derecho y en la tecnología.
- iii) Aunque suene utópico, deberíamos aspirar a una reglamentación vinculante global, pero, al menos, deben garantizarse y protegerse

la dignidad humana y sus derechos inviolables. Un tesoro irrenunciable que no ha sido fácil conquistar para el ser humano. Si bien se ha dado un paso importante hace unos días con la firma del Convenio Marco del Consejo de Europa sobre Inteligencia Artificial y derechos humanos, democracia y Estado de derecho, como primer instrumento internacional jurídicamente vinculante sobre los sistemas de IA.

- iv) Finalmente, la ley no debe ceder ante las intrigas tecnológicas, debe eliminar cualquier laguna o «gatera» que socave la propia regulación. Lo contrario sería dinamitar nuestros pilares éticos y jurídicos a medida que esta revolución digital, deshumanizada, se utilice en la vida diaria y las decisiones robóticas sean una realidad, relegándonos a la condición de víctima.

No olvidemos que estos ingenios han sido creados por el ser humano y que utilizan datos de seres humanos o generados por los seres humanos. Es fácil pensar que estos sistemas, que han alcanzado capacidades hasta ahora reservadas a los seres humanos, nunca pueden ser neutros ni neutrales; porque contienen, de forma más o menos camuflada, las virtudes consustanciales al ser humano, como son la bondad, la solidaridad, el servicio a los demás o el progreso; consecuentemente, tampoco se han podido eliminar los vicios y demás imperfecciones de la raza humana, como son la maldad, los complejos, las discriminaciones, las ambiciones o las frustraciones.

A medida que se eleve el nivel de complejidad de los algoritmos, será más difícil detectar esos vicios. Cuando los algoritmos sean los encargados de crear y fiscalizar los propios algoritmos, la capacidad de comprobación por el ser humano será más difícil, cuando no imposible. En ese momento, la persona habrá sido superada por la máquina y, por tanto, sustituida. Estará obligada a aceptar decisiones situadas extramuros del proceso de racionalidad humana y la persona ya no estaría en el centro.

En definitiva, regular la IA centrada en la persona se asemeja al vértigo de tener que cambiar la rueda del coche en marcha. A no ser que dejemos que sea la IA la que regule a la IA, dado que solo ella se conoce y se comprende a sí misma.

Bibliografía

- Aldama, Z. (2018). Videovigilancia: China se queda con tu cara. *El País*, 27-4-2018. Disponible en: <https://tinyurl.com/4ct872hb>.

- Álvarez García, V. y Tahiri Moreno, J. (2023). *La regulación de la inteligencia artificial en Europa a través de la técnica armonizadora del nuevo enfoque*. Madrid: INAP. Disponible en: <https://tinyurl.com/3xrws4rb>.
- Ansele, A. (2022). Tener un sensor en la cabeza será de rigor en 10 años, igual que ahora todo el mundo tiene un teléfono inteligente. *El País*, 5-1-2022. Disponible en: <https://tinyurl.com/3u9rwp4y>.
- Barrio Andrés, M. (2023). *Novedades en la tramitación del próximo Reglamento europeo de inteligencia artificial*. Madrid: Real Instituto Elcano. Disponible en: <https://tinyurl.com/3nujxeb4>.
- Byung-Chul, H. (2021). *No cosas*. Madrid: Taurus.
- Byung-Chul, H. (2022). *Infocracia*. Madrid: Taurus.
- Collingridge, D. (1980). *The Social control of Technology*. New York: St. Martin's Press.
- Cortina, A. (2019). Ética de la inteligencia artificial. *Anales de la Real Academia de Ciencias Morales y Políticas*, 96, 379-394.
- Deutsche Welle. (2021). China—¿Estado policial o laboratorio del futuro? *DW Documental*, 17-10-2021. Disponible en: <https://tinyurl.com/3v9ja8ae>.
- García, G. (2021). Innovación y tecnología en Japón: lecciones por aprender. *Documento de Opinión IEEE*, 39, 247-263. Disponible en: <https://tinyurl.com/4d8wyenh>.
- Giovanni de, G. (2022). *Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorithmic Society*. Cambridge: Cambridge University Press.
- Innerarity, D. (2018). *Política para perplejos*. Barcelona: Galaxia Gutenberg.
- Jonas, H. (1995). *El principio de responsabilidad. Ensayo de una ética para la civilización tecnológica*. Barcelona: Herder.
- Pedreño, A. y Moreno, L. (2020). *Europa frente a EE. UU. y China, prevenir el declive en la era de la inteligencia artificial*. Alicante: Autoedición.
- Pérez Royo, J. (2016). *Curso de Derecho Constitucional*. Madrid: Marcial Pons.
- Schneider, S. (2021). *Inteligencia artificial*. Badalona: Koan.
- Seibt, S. (2021). Clearview AI y la aplicación de reconocimiento facial que destruye la privacidad. *France 24*, 8-3-2021. Disponible en: <https://tinyurl.com/2p8ypvn4>.
- Skolnikoff, E. B. (1994). *The elusive transformation: science, technology, and the evolution of international politics*. Princeton: Princeton University Press. Disponible en: <https://doi.org/10.1515/9781400820924>.
- Tenorio Cueto, G. A. (2021). El derecho a una vida libre de algoritmos. *Revista del Instituto de Ciencias Jurídicas de Puebla*, 15 (48), 115-135. Disponible en: <https://doi.org/10.35487/rius.v15i48.2021.708>.
- Velarde, G. (2021). *Era artificial*. Düsseldorf: Prica.

Documentación normativa e institucional

- Blueprint for an AI Bil of rights making automated systems work for the American people, October 2022. Publicado por la Oficina de Política Científica y Tecnológica de la Casa Blanca. Disponible en: <https://tinyurl.com/4rz5fff5>.

- Comunicación de la Comisión al Parlamento Europeo, al Consejo Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. Inteligencia Artificial para Europa. Bruselas, 25-4-2018. COM (2018) 237 final. Disponible en: <https://tinyurl.com/bdh8yc9b>.
- II (Comunicaciones) Declaraciones Comunes. Parlamento Europeo, Consejo, Comisión Europea. Declaración Europea sobre los Derechos y Principios Digitales para la Década Digital (2023/C 23/01). *DOUE*, C23/1. Disponible en: <https://tinyurl.com/mr4yn74r>.
- Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law. Disponible en: <https://tinyurl.com/4445ayt8>
- Decisión del Parlamento Europeo, de 18 de junio de 2020, sobre la creación de una comisión especial sobre Inteligencia Artificial en la era digital y la definición de sus responsabilidades, fuerza numérica y mandato (2020/2684 (RSO)). Disponible en: <https://tinyurl.com/24crrewd>.
- Estrategia Nacional de Inteligencia Artificial de España. Disponible en: <https://tinyurl.com/5n8zruwu>.
- Libro Blanco sobre la inteligencia artificial—un enfoque europeo orientado a la excelencia y la confianza. Bruselas, 19-2-2020. COM (2020) 65 final. Disponible en: <https://tinyurl.com/yttf5uac>.
- Propuesta de Reglamento del Parlamento Europeo y del Consejo 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) 206 final. Disponible en: <https://tinyurl.com/mrxss5vx>.
- Puesta en marcha la Alianza Digital Unión Europea-América Latina y el Caribe. (2023). *España Digital 2026*, 15-3-2023. Disponible en: <https://tinyurl.com/37trzcbk>
- Recomendación sobre la ética de la inteligencia artificial, adoptada el 23 de noviembre de 2021. Disponible en: <https://tinyurl.com/bde6vevs>.
- Recomendación sobre la ética de la inteligencia artificial, adoptada por la UNESCO el 23 de noviembre de 2021. Disponible en: <https://is.gd/0tC4EJ>.
- Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo de 13 de junio de 2024 por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial). Disponible en: <https://tinyurl.com/3hskse29>.
- Resolución del Parlamento Europeo, de 20 de enero de 2021, sobre inteligencia artificial: cuestiones de interpretación y de aplicación del derecho internacional en la medida en que la UE se ve afectada en los ámbitos de los usos civil y militar, así como de la autoridad del Estado fuera del ámbito de la justicia penal (2020/2013[INI]). Disponible en: <https://tinyurl.com/mrxrhx4>.