ACCESO A LA VISUALIZACIÓN ONLINE DE ESTE NÚMERO DE LA REVISTA

Para acceder a la visualización Online de este número de la revista envíe un correo electrónico a atencionalcliente@tirantonline.com con la siguiente información:

– Título y número de la revista

– Ticket de compra

A continuación recibirá un correo electrónico con las claves e instrucciones de acceso

TEORÍA & DERECHO TAMBIÉN SE VE Y SE ESCUCHA

Descargando este código puede acceder a los vídeos grabados por sus autores

No se admitirá la devolución de este libro si el código promocional ha sido manipulado

CONSULTE LOS NÚMEROS ANTERIORES DE LA REVISTA EN
teoriayderecho.tirant.com

teoría & derecho

REVISTA DE PENSAMIENTO JURÍDICO

REVISTA SEMESTRAL. DICIEMBRE 37/2024
teoriayderecho.tirant.com

PVP número suelto: 89 euros (IVA incluido)
PVP suscripción anual (2 números): 169 euros (IVA incluido)

Redacción, Administración y Suscripción
Artes Gráficas, 14, 46010 Valencia Tel.: (+34) 963610048 Fax: (+34) 963694151
E-mail: teoria.derecho@uv.es. Sitios web: http://teoriayderecho.tirant.com
Correspondencia: C/ Artes Gráficas, 14 - 46010 - Valencia

Edita: Tirant lo Blanch
C/ Artes Gráficas, 14 - 46010 - Valencia

Telfs.: 96/361 00 48 - 50 Fax: 96/369 41 51
Mail: tlb@tirant.com http://www.tirant.com
ISSN (impreso): 1888-3443 ISSN (online): 2695-6594
Depósito legal: V-2840-2007
Maqueta: Tink Factoría de Color

SUMARIO

DEBATE

ESTUDIOS

SUMMARY

DEBATE

ESTUDIES

DEBATE

PRESENTACIÓN: DERECHO E INTELIGENCIA ARTIFICIAL

PRESENTATION: LAW AND ARTIFICIAL INTELLIGENCE

Fernando H. Llano

Catedrático de Filosofía del Derecho
Universidad de Sevilla

DOI: https://doi.org/10.36151/TD.2024.103

PRESENTACIÓN: DERECHO E INTELIGENCIA ARTIFICIAL

Fernando H. Llano
Catedrático de Filosofía del Derecho
Universidad de Sevilla

Es un hecho constatable que la transfiguración de la justicia humana en el universo digital y la paulatina sustitución del paradigma humanista por el transhumanista afecta a todos los órdenes del mundo de las profesiones, y en esto el derecho no es una excepción.

En efecto, hoy son muchos los que consideran que nos encontramos cada vez más cerca de la realización del mito del juez-robot, hipótesis que, en caso de producirse, conllevaría la sustitución de los jueces humanos; en última instancia, la realización de esta hipótesis estará condicionada al modelo de justicia por el que se decanten las generaciones del futuro, y en buena parte esto dependerá de la defensa de la condición humanística del derecho, es decir, parafraseando a Ulpiano, de seguir concibiéndolo como arte de lo bueno y de lo justo

La hipótesis de los jueces-robot se plantea solo en el escenario de la segunda ola de la IA, es decir, una vez superada la fase original en la que los sistemas de IA eran utilizados como instrumentos de apoyo para resolver problemas judiciales y redactar documentos jurídicos. En esa primera ola de la IA, los sistemas expertos basados en reglas se desarrollaban mediante un proceso de aprendizaje en el que se trataba de extraer el conocimiento y los procesos de razonamiento de la mente de los expertos humanos en Derecho.

Este sistema era capaz de plantear preguntas a los usuarios, dar respuestas judiciales y redactar documentos jurídicos, con un nivel de calidad y precisión técnica que a menudo era superior al de los expertos humanos; aunque se trataba solo de un sistema experto de aprendizaje en el que se desplegaba un árbol de decisión gigante donde se representaban los patrones aparentes del razonamiento jurídico humano (Susskind, 2020, 308-310).

Mientras que la primera ola de IA fue programada por humanos, la segunda ola se caracteriza por su ilimitada capacidad de procesamiento de datos y por el uso de algoritmos que reconocen patrones de comportamiento a partir del análisis de ese ingente volumen de datos. Estos sistemas expertos son autónomos en la medida que ya no se basan en la ejecución de reglas predeterminadas (*code-driven*), sino que son capaces de generar sus propios modelos predictivos o de tomas de decisiones en función de los datos que se le suministren (*data-driven*) (Hildebrandt, 2018).

En resumidas cuentas, el aprendizaje automático de las máquinas inteligentes (*machine learning*) supone un cambio de paradigma computacional y constituye el rasgo distintivo de la actual IA (Solar Cayón, 2021, 335)[1].

Y esto es solo el comienzo de la revolución digital en el ámbito de la justicia y de la automatización de la práctica profesional del derecho. A este respecto, según el pronóstico de Richard Susskind, en el futuro habrá un "sinfín de avances" que irán más allá del aprendizaje automático, la técnica de IA dominante del momento. En cuestión de décadas, los sistemas expertos superarán a los jueces en su propio terreno, dictando sentencias razonadas con explicaciones parecidas a las que se ofrecen en las mejores sentencias dictadas por los jueces humanos, pero con la diferencia de que aquellas se habrán generado por *software* IA en vez de por *wetware* judicial (Susskind, 2020, 323).

En relación con el salto cualitativo operado en la evolución de IA jurídica y de la consiguiente aparición de start-ups que comercializan sistemas capaces de predecir el comportamiento de los tribunales (por ejemplo, *Lex Machina*) se ha especulado sobre el impacto que estas aplicaciones inteligentes tendrán tanto en la justicia como en el mercado de servicios jurídicos. En este sentido, algunos autores estiman que, conforme las máquinas vayan ampliando sus capacidades y asumiendo más competencias en la práctica cotidiana de las profesiones (entre ellas, también las profesiones vinculadas a la teoría y la praxis del derecho), la contribución de los humanos al mercado de trabajo irá reduciéndose proporcionalmente (Susskind, R. – Susskind, D., 2015; Susskind, D., 2020).

Sin embargo, para los más escépticos respecto a los supuestos efectos benéficos del gradual proceso de mecanización y/o robotización del trabajo tradicionalmente desempeñado por el hombre, siempre deberían quedar algunas áreas profesionales reservadas solo para humanos, especialmente en aquellos trabajos donde la interacción entre las personas es fundamental (Bessen, 2020, 83-88).

A propósito de la insistencia de una parte de la doctrina en la creencia de que hay tareas que nunca podrán realizar las máquinas de manera satisfactoria, pues en última instancia las máquinas serían incapaces de copiar y reproducir perfectamente la forma en la que funcionan los seres humanos, Richard Susskind ha rebatido recientemente esta opinión aduciendo que está basada en la "falacia de la IA", es decir, inspirada en la creencia de que la única forma de desarrollar máquinas que puedan funcionar al mismo nivel que los seres humanos consiste en "*copiar* la forma en que funcionan los seres humanos". Es más, en la actualidad, ya existen muchos sistemas de IA que no operan necesariamente copiando el comportamiento, ni dependiendo de los patrones de actuación humana,

sino de manera muy diversa (esto es, no humana); en este sentido, estos sistemas contemporáneos pueden "realizar predicciones, identificar documentos importantes, contestar preguntas y gestionar emociones a un nivel superior al de las personas" (Susskind, 2020, 316-317).

En suma, la creciente introducción de la IA en el ámbito profesional, en general, y en el mercado de los servicios jurídicos, en particular, se caracteriza porque está sometida a cambios que, en su primera fase, estará dominada por la *automatización* (lo cual supone que muchas labores cotidianas, arduas y/o reiterativas, se irán optimizando mediante la aplicación de la tecnología); posteriormente, en una segunda fase marcada por la *innovación tecnológica*, se introducirán herramientas y sistemas, cada vez más autónomos e inteligentes, que acabarán transformando el trabajo de los profesionales, "dando lugar a nuevos procedimientos para compartir el conocimiento práctico que aquellos poseen" (Solar Cayón, 2019, 49).

En esa segunda fase de la IA, que en el marco de la IA jurídica nos permite concebir la justicia digital de los tribunales on-line, se ahorrará tiempo y dinero (en la medida que las máquinas irán asumiendo cada vez más funciones de mediación hasta ahora reservadas a los abogados y los procuradores). A este respecto, cabe imaginar a corto/medio plazo un sistema de aprendizaje automático que, usando el *Big Data* y los algoritmos computacionales, ayude a las partes prediciendo el resultado del caso si finalmente decidieran presentarlo ante un juez humano.

En rigor, no parece viable que, desde un punto de vista técnico, las máquinas estén preparadas a corto y medio plazo para razonar, argumentar, pensar o sentir como lo haría un juez humano; por otra parte, tampoco cabría afirmar que las máquinas actuales, dotadas de sistemas de aprendizaje automático, pueden ofrecer los resultados del método judicial (es decir, que estén en disposición de dictar decisiones motivas); sin embargo, como demuestran aplicaciones inteligentes aplicadas al mercado de los servicios jurídicos, por ejemplo *Lex Machina*, existen ya sistemas inteligentes de justicia predictiva que permiten saber con antelación cuál será el comportamiento de los tribunales (la profecía del realismo jurídico representado por Oliver Wendell Holmes cumplida merced a los algoritmos, el *Big Data* y la estadística computacional).

Y esto es solo el comienzo de la revolución digital en el ámbito de la justicia y de la automatización de la práctica profesional del derecho. A este respecto, según el pronóstico de Richard Susskind, en el futuro habrá un "sinfín de avances" que irán más allá del aprendizaje automático, la técnica de IA dominante del momento. En cuestión de décadas, los sistemas expertos superarán a los jueces en su propio terreno, dictando sentencias razonadas con explicaciones parecidas a las que se ofrecen en las mejores sentencias dictadas por los jueces humanos, pero con la diferencia de que aquellas se habrán generado por *software* IA en vez de por *wetware* judicial (Susskind, 2020, 323).

Los artículos que conforman el presente número de la revista *Teoría & Derecho* se centran precisamente, desde una perspectiva interdisciplinar (iuspublicista e iusprivatista), en el estudio de las principales claves, retos y oportunidades que nos plantea la irrupción de

las tecnologías de frontera en la era de la revolución industrial 4.0, y muy especialmente la aplicación de la IA en la teoría y la praxis del derecho contemporáneo.

NOTAS

1. Señala Claudio Sarra que el abandono del paradigma antropocéntrico por el tecnoparadigma dataísta se debe, fundamentalmente, a la elevación a la máxima potencia del hombre frente a la naturaleza, es decir, a una especie de transfiguración por parte del individuo que supera la heroica oposición hombre-mundo de la modernidad (Sarra, 2019, 100-101).

BIBLIOGRAFÍA

Bessen, J. (2020), "Attitudes to Technology. Part 1", *Work in the Future. The Automation Revolution*, R. Skidelsky-N. Craig (eds.), Palgrave MacMillan, Cham (Switzerland), pp. 83-88.

Hildebrandt, M. (2018), "Algorithmic Regulation and the Rule of Law", *Philosophical Transactions Royal Society A* 376: 20170355, 2018. Disponible en: https://royalsocietypublishing.org/doi/full/10.1098/rsta.2017.0355 (última consulta, 12 de diciembre de 2024).

Solar Cayón, J. I., (2019), *La Inteligencia Artificial Jurídica. El impacto de la innovación tecnológica en la práctica del Derecho y el mercado de servicios jurídicos*, Thomson Reuters Aranzadi, Cizur Menor (Navarra).

— (2021), "IA en la abogacía: el futuro ya está aquí", en: *Inteligencia Artificial y Derecho. El jurista ante los retos de la era digital,* F. H. Llano Alonso-J. Garrido Martín (eds.), Thomson Reuters Aranzadi, Cizur Menor (Navarra).

Susskind, R. (1996), *The Future of Law. Facing the Challenges of Information Technology*, Clarendon Press, Oxford.

— (2013), *Tomorrow´s Lawyers. An Introduction to Your Future*, Oxford University Press, Oxford.

— (2020), *Tribunales online y la Justicia del futuro,* trad. esp., GEA Textos, La Ley - Wolters Kluwer, Madrid.

Susskind, R. – Susskind, D. (2015), *The Future of the Professions. How Technology will transform the Work of Human Experts*, Oxford University Press, Oxford.

Susskind, D. (2020), *A World Without World: Technology, Automation and How We Should Respond*, Allen Lane, London.

Susskind, R. (1996), *The Future of Law. Facing the Challenges of Information Technology*, Clarendon Press, Oxford.

Susskind, R. – Susskind, D. (2015), *The Future of the Professions*, Oxford University Press, Oxford.

CUESTIONES ACTUALES DEL DERECHO DE LOS ROBOTS*

CURRENT ISSUES IN ROBOT LAW

Moisés Barrio Andrés

Letrado del Consejo de Estado
Asesor en materia de regulación digital
Profesor de Derecho digital

RESUMEN

La entrada de la robótica y de la inteligencia artificial en el mundo jurídico es un hecho y plantea algunos retos que han impulsado una nueva rama jurídica: el Derecho de los robots. En este trabajo repasamos las cuestiones actuales —y aún irresueltas en términos jurídicos— referidas al estatuto legal general de los robots y el reconocimiento de una eventual persona robótica, los desafíos de la interacción persona-robot, los problemas de responsabilidad civil y el impacto más urgente derivado de los coches autónomos y la sustitución de trabajadores por robots.

PALABRAS CLAVE

Derecho de los robots, persona robótica, persona electrónica, coche autónomo, vehículo autónomo, responsabilidad civil robótica, responsabilidad por accidentes de circulación de vehículos de motor, trabajador robot.

ABSTRACT

The entrance of robotics and artificial intelligence into the legal world is a fact and poses some challenges that have led to the emergence of a new legal branch: Robot law. In this paper, we review the current —and still unresolved in legal terms— issues regarding the general legal status of robots and the recognition of an eventual robotic person, the challenges of human-robot interaction, civil liability issues and the most urgent impact arising from autonomous cars and the replacement of workers by robots.

KEYWORDS

Robot Law, robotic personhood, electronic person, autonomous car, autonomous vehicle, robotic civil liability, motor vehicle road accident liability, robot worker.

DOI: https://doi.org/10.36151/td.2024.104

* Estudio realizado en el marco del Proyecto I+D+i «Garantías institucionales y regulatorias. Autoridades electorales y de supervisión digital ante interferencias, narrativas hostiles, publicidad segmentada y polarización. (Dir-Politics)». Referencia PID2022-137245OB-I00. Financiado por la Agencia Estatal de Investigación.

CUESTIONES ACTUALES DEL DERECHO DE LOS ROBOTS

Moisés Barrio Andrés
Letrado del Consejo de Estado
Asesor en materia de regulación digital
Profesor de Derecho digital

1. INTRODUCCIÓN

Los robots son cada vez más omnipresentes. Como es sabido, ya pueden encontrarse robots en los hogares o las oficinas, las tiendas, las fábricas, los hospitales, las carreteras, en el aire, en el mar y hasta en el campo de batalla. Es, pues, previsible que, a medida que la robótica mejore, se incremente el número de robots capaces de llevar a cabo tareas cada vez más más complejas. De ello se deduce que la interacción entre los robots y las personas —y, naturalmente, el Derecho— se intensificará en casi todos los ámbitos imaginables, generando nuevos espacios de relación y de conflicto.

Es habitual definir a los robots como el producto de la mecanización del paradigma «OODA» (Observar, Orientar, Decidir, Actuar). A efectos de este trabajo, consideraré «robot» a cualquier dispositivo capaz de afectar al mundo exterior en respuesta a algún desencadenante detectado. Se trata de una definición muy amplia que abarca desde los dispositivos sencillos como un termostato de una habitación programable que cambia la calefacción o el aire acondicionado en respuesta a una elección de temperatura hasta los androides y otras formas de inteligencia artificial de perfil humanoide, pasando por los complejos algoritmos de negociación en programas informáticos que compran o venden acciones en respuesta a los movimientos del mercado o los sistemas punteros de defensa militar diseñados para derribar misiles (García-Prieto Cuesta, 2018).

Muchos robots sencillos no plantean problemas jurídicos inusuales y, por ello, no serán objeto de análisis en estas páginas. Un termostato está regulado por la legislación relativa a la seguridad general de los productos; asimismo, la actividad de un robot instalado en una cadena de montaje se rige por las normas de seguridad e higiene en el trabajo y las disposiciones sobre la responsabilidad civil. Ahora bien, las normas que regulan el funcionamiento de estos artefactos son las mismas que resultan aplicables a otros dispositivos más sofisticados emplazados en los hogares o en las empresas. Aunque incluso los robots sencillos pueden plantear ocasionalmente cuestiones de hecho complejas, normalmente no suscitan problemas teóricos que les afectan de manera exclusiva, en contraste con lo que sucede en el caso de otros productos y dispositivos conexos.

Por el contrario, como veremos más adelante, los robots más avanzados e interesantes —sobre todo, los que son capaces de comportarse de forma autónoma— plantean cuestiones jurídicas inéditas y a menudo no resueltas concernientes a la responsabilidad, la seguridad, la supervisión y la jurisdicción, así como una serie de problemas relativos a la identificación del responsable de los daños causados por los robots, la propia definición del daño y la casuística de la eventual restricción y aun la prohibición de determinados robots.

Para hacerse una idea aproximada de la complejidad del escenario, basta considerar que el hecho de que todos los robots, salvo los de *software* puro o *softbots*, sean corpóreos implica que las cuestiones jurídicas y la regulación relativa a los robots alcancen a todos los niveles de regulación, desde el local al internacional. Como ya advertimos en su momento (Barrio Andrés, 2018), especialmente en Estados Unidos y en Europa el Derecho de los robots está surgiendo como una nueva disciplina jurídica que simultáneamente «[…] adapta el Derecho general —en cuanto que previamente vigente— y, en la medida de lo necesario, genera un Derecho nuevo», tal y como sucede con los coches sin conductor, a los que me referiré más adelante.

Por si los desafíos apuntados hasta aquí fueran pocos, muchas de estas complejas cuestiones se solapan con la regulación jurídica de la inteligencia artificial (IA) o forman parte de ella, dado que los robots complejos e interesantes suelen estar controlados por una IA. Esta puede encastrarse en el seno del robot o funcionar a distancia en la plataforma del fabricante o proveedor (*cloud robotics*), dando origen a nuevos niveles de complejidad ligados al diseño robótico. Por ejemplo, los robots en el campo de batalla están concebidos para operar frente a los intentos del enemigo de interferir las comunicaciones; normalmente, la ubicación del puesto de control no es jurídicamente relevante, aunque puede tener un impacto jurídico en los casos en los que el piloto externo se encuentre en una jurisdicción distinta de la parte física del robot militar, o este sea controlado por diferentes personas, como sucede en determinadas misiones internacionales (Arellano Toledo, 2023).

Por tanto, una de las características del Derecho de los robots es la trasversalidad y la diversificación de los escenarios en los que opera. Aunque este trabajo tiene como objeto genérico los retos actuales aún no resueltos, se centrará en la problemática específica del empleo de vehículos autónomos, una realidad que ha puesto de manifiesto la necesidad de que Estados Unidos y la Unión Europea aborden una remodelación urgente del marco

jurídico que dé cabida a los avances técnicos en el campo de la robótica y la inteligencia artificial (IA) autónomos.

Ya anticipo que el Derecho estadounidense tendrá un mayor protagonismo, toda vez que en la Unión Europea ha prestado atención a la responsabilidad civil en numerosos estudios y resoluciones (Muñoz García, 2023; Llano Alonso, 2024), trabajos preparatorios que aún no se han concretado en una reforma legislativa en pro de una verdadera armonización, más allá de las previsiones normativas relativas la responsabilidad de los fabricantes por productos defectuosos. De hecho, la reciente aprobación del Reglamento Europeo de Inteligencia Artificial (RIA o AI Act)[1], que califica a los robots más inteligentes como sistemas de alto riesgo (art. 6 RIA), aborda cuestiones importantes que hemos examinado en otro lugar (Barrio Andrés, 2024*b*), pero no incorpora una regulación específica de la responsabilidad civil. Sea como fuere, parece lógico comenzar el trabajo con una referencia al estatuto jurídico de los robots.

2. EL ESTATUTO JURÍDICO GENERAL DE LOS ROBOTS

2.1. LA PERSONA ROBÓTICA

Con carácter general, el Derecho vigente trata a todos los tipos de robots como bienes muebles. Por el momento, la única excepción significativa es Arabia Saudí, que concedió la ciudadanía a un robot humanoide, Sophia, en 2017. En consecuencia, el punto de partida del análisis del Derecho de los robots debe ser la consideración instrumental del objeto regulado. Como observan Richards y Smart, «los robots son, y durante muchos años seguirán siendo, herramientas. Son herramientas sofisticadas que utilizan *software* complejo, sin duda, pero no diferentes en esencia de un martillo, un taladro eléctrico, un procesador de textos, un navegador web o el sistema de frenado de tu coche» (Richards y Smart, 2016: 18).

De esta consideración se infiere que la personificación jurídica de los robots —o de la inteligencia artificial— sigue siendo, por el momento, una posibilidad remota. A no ser que se produzca algún avance revolucionario en las redes neuronales o en otra tecnología disruptiva, parece poco probable que durante la próxima década las máquinas de cualquier clase alcancen el tipo de autoconciencia y sensibilidad que solemos asociar a una reivindicación legítima del conjunto de derechos y protección agrupados bajo la rúbrica de la personalidad. Como he sostenido en varias ocasiones (la última, en Barrio Andrés, 2024*a*), lejos de resultar tranquilizador, este impedimento puede tener algunos efectos indeseados, entre ellos la dificultad para resarcir a las víctimas de daños o la imposibilidad de establecer prohibiciones de uso de robots para determinados fines (por ejemplo, las prácticas pedófilas).

Sea como fuere, el hecho de que el estado actual de desarrollo de esta tecnología resulte aún en parte incipiente no es óbice para identificar dos situaciones en las que la sociedad y

los legisladores podrían optar por reconocer algún tipo de derechos o protecciones jurídicas a los robots más allá de los que normalmente se atribuyen a los bienes muebles.

La primera está vinculada a la pertinencia de reconocer la utilidad social de la ficción jurídica de acuerdo con la cual un robot es una persona. Nadie, después de todo, cree seriamente que una empresa es una persona real, o incluso que una persona jurídica está viva o sea sensible. Sin embargo, aceptamos la ficción legal de la persona jurídica porque sirve a intereses que la sociedad reputa útiles, entre ellas la capacidad de realizar transacciones en su propio nombre y la limitación de la responsabilidad de sus accionistas. En esta línea, he defendido que el reconocimiento legal de la persona robótica puede reportar beneficios sociales similares (Barrio Andrés, 2024*a*). Otros aspectos de inscripción registral y fondo de aseguramiento han sido analizados con brillantez por Campuzano Gómez-Acebo y Sieira Gil (2023: 649 ss.). Incluso otros autores han llevado la cuestión al extremo y han estudiado la viabilidad de la hipótesis de que un robot o una IA puedan ser propietarios de sí mismos (así, Bayern, 2021; LoPuki, 2018).

Ahora bien, la posibilidad de que, a semejanza de las personas jurídicas, los robots sean titulares de determinados derechos es una cuestión que no ha concitado una postura teórica unánime. En nuestro entorno más cercano, las primeras sentencias dictadas sobre el tema por los tribunales españoles o de la Unión Europea descartan que pueda atribuirse a un robot o a una IA la titularidad de un derecho de autor o una patente incluso cuando dicha tecnología haya generado la obra o el invento en cuestión (Ortego Ruiz, 2022). Tales resoluciones toman como base la interpretación que de este derecho ha venido realizando el Tribunal de Justicia de la Unión Europea, conforme a la cual el derecho solo brindará protección a las obras originales, originalidad que debe reflejar una «creación intelectual propia del autor [humano]» (por todas, las sentencias *Infopaq*, de 16 de julio de 2009 (asunto C-5/08, caso *Infopaq International A/S c. Danske Dagblades Forening*) o *Cofemel*, de 12 de septiembre de 2019 (asunto C-683/17, caso *Cofemel - Sociedade de Vestuário, S.A c. G-Star Raw CV*). En contraste, un tribunal chino dictaminó que un texto producido por IA sí merecía la protección de los derechos de autor en virtud de la legislación china (Sawers, 2020). Se observa, pues, que la respuesta que se da a la cuestión de la equiparación de los robots a las personas jurídicas depende en buena medida de la cultura jurídica de referencia.

La segunda situación en la que podrían reconocerse algunos derechos a los robots es más plausible y parte de la siguiente premisa: los seres humanos tienden a antropomorfizar a los robots. Como observa de forma pionera Darling, «nuestra bien documentada inclinación a relacionarnos antropomórficamente con los animales se traslada extraordinariamente bien a los robots» (Darling, 2016: 223), y mucho más a los robots sociales de apariencia real diseñados para generar esa reacción. Como han puesto de manifiesto diversos autores, la gente no solo tiende a antropomorfizar a los robots, sino también a atribuirles características de género basadas en sus funciones, así como a asignarles patrones conductuales de carácter social como la cortesía y la reciprocidad (Rosenthal-von der Putten *et al.*, 2013; Nass y Moon, 2000). La cuestión tiene interés porque no debemos descartar que la atribución de derechos a los robots encuentre su fundamento en una exigencia social ligada a esta

empatía. A fin de cuestas, algo así ya viene ocurriendo con otras realidades; por ejemplo, los animales e incluso los ecosistemas, tradicionalmente considerados como objetos. La Ley 19/2022, de 30 de septiembre, para el reconocimiento de personalidad jurídica a la laguna del Mar Menor y su cuenca es elocuente en este punto.

Si los estudios sugieren que muchas personas están preparadas para percibir el daño a los robots humanoides como una conducta negativa (Darling, 2016:224), incluso cuando saben que se trata de una máquina (Nass y Moon, 2000), no es ilógico aventurar que algunos ordenamientos jurídicos puedan optar por instituir la prohibición del «abuso» a los robots (golpear, torturar, violar…). A a esta misma solución puede conducir la necesidad de brindar una suerte de protección indirecta de las personas si encontramos apoyo empírico a la hipótesis de que el comportamiento abusivo de los robots hacia las personas o animales será tanto más probable cuanto más parecidos sean a los humanos (Darling, 2016: 226-231). En cualquiera de estas líneas argumentales, los robots adquieren ciertos derechos frente a los humanos: al menos en términos hohfeldianos, si las personas tienen el deber de no dañar a un robot, entonces, correlativamente, el robot tiene derecho a no ser dañado por las personas (Coeckelbergh, 2010).

2.2. CUESTIONES JURÍDICAS DE LA INTERACCIÓN PERSONA-ROBOT

Desde el punto de vista del impacto sobre los individuos, nuestra tendencia a antropomorfizar a los robots puede, colateralmente, convertirlos en un medio para manipular a las personas que interactúan con ellos. Las posibilidades son innumerables. Si la gente se encariña con sus mascotas o con sus compañeros robóticos, es fácil que las empresas que fabrican o gestionan los robots caigan en la tentación de utilizar ese vínculo emocional para imponer pagos coactivos a los usuarios —por ejemplo, para mejorar sus prestaciones—. Los robots también podrían ser usados para introducir en los hogares anuncios publicitarios disfrazados de conversaciones normales, posibilidad que suscita preocupación, especialmente en el caso de los robots para niños. Tal vez formulen preguntas diseñadas para obtener y revelar información personal (Nyholm, 2020).

Un primer aspecto que debería inquietarnos en la interacción persona-robot es la excesiva confianza que pueden generar algunos robots creados para recopilar información sobre los usuarios, aunque no sea de naturaleza sensible en la terminología que comúnmente empleamos para referirnos a la protección de los datos personales (las «categorías especiales de datos personales» del art. 9 RGPD). Nadie discute la necesidad de recabar el consentimiento del usuario o de su representante legal para almacenar, por ejemplo, datos de un seguimiento médico con el fin de evitar su utilización desviada y el consiguiente daño a la privacidad del paciente. Los daños potenciales para el individuo no parecen menores en el caso de los robots; por ejemplo los robots de limpieza, que registran información sobre su entorno y la comparten con la empresa que fabrica los productos de limpieza (cfr. Kaminsky, 2015, que analiza los riesgos centrándose en el consentimiento y la divulgación de la información), o los dispositivos basados en el reconocimiento de voz, que pueden facilitar el acceso a las conversaciones de sus usuarios por terceros sin su consentimiento[2]. De hecho,

la falsificación parece una posibilidad tan presumible como indeseable en las interacciones persona-robot. Este tipo de conductas ilícitas adopta múltiples modalidades comisivas: desde los casos de robots que fingen ser personas en *chats* de pantalla hasta aquellos supuestos en que las personas simulan ser robots para que su producto parezca más sofisticado o sus consejos semejen tener una base científica (Brennan-Marquez, Levy y Susser, 2019).

Todos estos escenarios imponen la necesidad de establecer una regulación (entre otros, Hartzog, 2015), tarea que hasta ahora ha alcanzado un escaso grado de concreción. En algunos casos especiales como en el tratamiento de datos o las relaciones de representación, estos engaños son ilegales o constituyen una violación de la ética y los deberes profesionales. En otros contextos, sin embargo, aunque las acciones sean inmorales, únicamente pueden convertirse en ilícitos procesables cuando alguien puede demostrar un perjuicio. Por más que el Tribunal de Justicia de la Unión Europea haya establecido que el mero incumplimiento de la normativa europea de protección de datos da lugar a responsabilidad, lo cierto es que solo cabe indemnizar los daños probados (STJUE de 14 de diciembre de 2023, asunto C-456/22, caso *VX, AT c. Gemeinde Ummendorf*).

Independientemente de la cuestión su estatuto jurídico subyacente, un segundo aspecto que debería tomarse en consideración en la interacción persona-robot es que algunos robots son capaces de reaccionar mucho más rápidamente a los estímulos que las personas. Esta característica técnica puede ser muy valiosa para realizar determinadas tareas como ajustar los controles de vuelo de un avión, cambiar los parámetros de refrigeración de una central nuclear o apuntar a un blanco. No obstante, dado que los sensores y la programación nunca son perfectos, las reacciones rápidas también generan el riesgo de que se produzcan errores rápidos y a veces catastróficos. Esta eventualidad sugiere la conveniencia de que, cuando un robot tiene el poder de hacer un daño significativo, la ley exija la presencia de un «humano en el bucle» para reducir el riesgo de resultados no deseados (Llano Alonso, 2024: 135). Ahora bien, también existe el riesgo de que la acción conjunta del robot y el humano responda al peligro con lentitud, en algunos casos demasiado despacio como para garantizar la seguridad o la eficacia de la tarea.

El desempeño conjunto entre la persona y el robot nos acerca a la problemática de la responsabilidad. Elish describe una familia de riesgos que denomina «zonas de deformación moral», concretados en la asignación al ser humano de la responsabilidad —también la de índole legal— sin que realmente tenga la posibilidad de asumirla. Por ejemplo, si el sistema general está mal diseñado, cuando se produce una situación de crisis el ser humano en el bucle puede suceder que el ser humano nunca disponga del tiempo suficiente para evaluar los hechos y tomar una buena decisión. El fallo será más probable cuanto más rutinaria, repetitiva y tediosa sea la labor de supervisión, o cuando el humano sea responsable de tantas tareas simultáneas de robots que la supervisión real se haga prácticamente imposible. Tal sería el caso, por ejemplo, de un guardia de seguridad encargado de supervisar un gran número de robots de vigilancia (Elish, 2019).

En estos casos, la falta de personalidad jurídica de los robots supone que estos no puedan ser jurídicamente responsables de los daños o delitos que causen (Gómez-Riesco Tabernero de Paz, 2018). Por lo general, cabría esperar que los daños —y los delitos— causados por

un robot fueran imputados a la persona responsable de la acción del robot. Sin embargo, si no hay un humano en el bucle, ¿quién debería ser el responsable? Cuando se trata de daños civiles, por ejemplo, suele señalarse a la persona cuyas acciones fueron la causa próxima del daño o cuya negligencia propició su producción. A veces puede estar claro: alguien puso en marcha las acciones del robot, alguien no previó una contingencia previsible en la construcción o el funcionamiento del robot, o alguien que tenía un deber de diligencia mayor no supervisó adecuadamente las acciones del robot. Otras situaciones, sin embargo, serán mucho menos claras, y entonces puede ser necesario recurrir a conceptos y construcciones teóricas más elaborados de la negligencia, las presunciones sobre la culpa o el nexo causal o la objetivación de la responsabilidad (Barrio Andrés, 2022*a*:172 ss.).

En el caso más extremo, los delitos cometidos por —o a través de— robots presentan complejidades adicionales. No es difícil afirmar que, si una persona programara un robot para hurtar, esa persona sería penalmente responsable por el hurto del robot, al igual que cuando para cometer el ilícito utiliza cualquier otro instrumento (Domínguez Peco, 2018). Pero ¿qué ocurre si la conducta persona que posee o maneja el robot no es dolosa, como en el caso real de un robot programado para hacer pedidos aleatorios en la red oscura que terminó adquiriendo pastillas de éxtasis para su propietario? (Power, 2014) ¿Existe dolo cuando la persona que puso en marcha el robot no pretendía ni preveía cometer el delito de compra ilegal de estupefacientes? En ocasiones, como en el caso recién citado, puede ser suficiente sostener la tesis que la persona que maneja el robot debería haber previsto la acción y haberse protegido contra ella. A veces, sin embargo, ese juicio puede ser bastante problemático, especialmente si la posibilidad de perpetrar un delito es remota.

2.3. CUESTIONES JURÍDICAS DERIVADAS DEL COMPORTAMIENTO AUTÓNOMO

El problema de la atribución de la responsabilidad alcanza su punto álgido en el caso de los robots diseñados para aprender de la experiencia, particularmente cuando este proceso de aprendizaje tiene lugar con poca o nula supervisión humana. Por ejemplo, algunos robots convenientemente programados han aprendido a caminar solos (García-Prieto Cuesta, 2018). A medida que se incremente la interacción de los robots con las personas, el autoaprendizaje basado en la experiencia será más necesario, lo que puede generar una dinámica en la que los robots aprendan a hacer cosas de forma impredecible. De hecho, lo más realista es pensar que aprenderán a hacer cosas totalmente impredecibles.

Este comportamiento autónomo, es decir, «[…] el comportamiento que es útil pero que los operadores no pueden prever de antemano» (Calo, 2014: 5) es, en gran medida, una característica, no un error, pero genera oportunidades para que se produzcan lesiones desconocidas hasta la fecha. ¿Qué hacer cuando el comportamiento autónomo es el resultado de múltiples interacciones con distintas personas a lo largo de la cadena de fabricación del robot? ¿Es correcto culpar al diseñador por incluir una característica que normalmente no solo es positiva, sino que también puede ser necesaria para que el robot aprenda a realizar tareas complejas? Quizá deberíamos culpar al operador del robot por no supervisarlo

adecuadamente, pero la naturaleza de un comportamiento autónomo es que puede manifestarse sin previo aviso.

En la actualidad no existe consenso sobre el modo de atribuir la responsabilidad en caso de que la lesión sea causada por un comportamiento autónomo, inesperado, de un robot; de hecho, la doctrina jurisprudencial sobre esta cuestión es escasa. Sin embargo, el problema ya está sobre la mesa en el caso de los coches autónomos, por lo que parece adecuado centrar el análisis en los mismos.

3. VEHÍCULOS AUTÓNOMOS

Aunque contamos con una amplia experiencia sobre la asignación de la responsabilidad en los accidentes de tráfico y sobre la regulación del manejo de los coches en materia de seguridad, ello no significa que la tarea de asignar la responsabilidad de los accidentes en los que intervienen coches autónomos (también conocidos como «coches autoconducidos», «coches robot» o «vehículos autónomos») sea tan fácil como cabría esperar[3]. También surgirán nuevos problemas, e incluso problemas de carácter transnacional —por ejemplo, cuando los usuarios conduzcan un vehículo autoconducido más allá de las fronteras nacionales—.

A esta complejidad intrínseca se añade la existencia de un intrincada entramado regulador al que es preciso referirse con carácter previo.

Parte de la problemática de los vehículos autónomos estriba en que tanto en Estados Unidos como en la Unión Europea la competencia para regular los vehículos de motor está repartida entre las administraciones públicas nacionales, estatales y locales. La Administración Nacional de Seguridad del Tráfico en las Carreteras de Estados Unidos (la National Highway Traffic Safety Administration, o NHTSA) regula muchos aspectos de la seguridad de los vehículos de motor. Este organismo federal impone unas normas mínimas de seguridad, por ejemplo, en materia de resistencia a los choques, se encarga de realizar las llamadas a revisión cuando detecta defectos en el proceso de fabricación de los coches y regula o coordina una serie de normas de seguridad vial. Por su parte, los estados tienen a su cargo la expedición de los permisos de conducir, tarea sujeta a una serie de normas federales, entre ellas las relativas a las formas aceptables de identificación. Asimismo, la mayoría de las normas de tráfico, entre ellas las que establecen los límites de velocidad (sujetos a un tope nacional fijado por el Congreso) y regulan el aparcamiento, son competencia de los gobiernos estatales y, a menudo, de los municipios. Este reparto competencial supone que la aplicación de las leyes de tráfico varía en cada estado, aunque la policía local y los departamentos responsables de velar por el cumplimiento de las normas de tráfico suelen asumir la mayor parte de este cometido. Los coches robot y sus usuarios están sujetos a todos estos regímenes normativos.

En un escenario normativo que permite la diversidad reguladora, muchos estados de Norteamérica han promulgado leyes destinadas a fomentar las pruebas y el despliegue

final de los coches autoconducidos mediante la creación de excepciones a las normas que obligan a llevar conductor. Y varias empresas han puesto en marcha pequeñas flotas de coches, algunos con monitores en el asiento delantero, pero otros sin otro ocupante que el pasajero. Por ejemplo, la empresa Cruise introdujo un servicio de taxi sin conductor en algunas zonas de San Francisco, pero después de un accidente retiró todos sus robotaxis en Estados Unidos (Wessling, 2022).

La situación no es diferente en la Unión Europea, aunque en este ámbito disponemos del potencial armonizador del ordenamiento comunitario, así como del Foro Mundial para la Armonización de la Reglamentación sobre Vehículos, organismo intergubernamental auspiciado por la Comisión Económica para Europa de las Naciones Unidas (CEPE) que crea marcos normativos para garantizar el rendimiento y la seguridad de los vehículos e invita a los Estados miembros a adoptarlos. Por ejemplo, recientemente amplió la velocidad máxima recomendada para los vehículos terrestres con sistemas automatizados de mantenimiento de carril de 60 km/h a 130 km/h.

El principal cuerpo normativo de la regulación europea de los vehículos autónomos es el Reglamento (UE) 2019/2144 del Parlamento Europeo y del Consejo, de 27 de noviembre de 2019[4]. Esta norma introdujo las nociones fundamentales y las reglas básicas referidas a los procesos y métodos de homologación de los vehículos automatizados. En particular, el Reglamento establece que la incorporación de los sistemas de seguridad ha de constituir la base de la conducción automatizada, y que debe tenerse en cuenta el progreso tecnológico de esos sistemas en cada revisión de la legislación vigente a fin de que esta sea aplicable en un futuro con el fin de reducir o eliminar accidentes y, en consecuencia, lesiones en el transporte por carretera.

En virtud de la disposición de desarrollo prevista en el artículo 11 del Reglamento, relativa a los vehículos automatizados y los vehículos totalmente automatizados, la Comisión Europea adoptó el Reglamento de Ejecución (UE) 2022/1426 de la Comisión, de 5 de agosto de 2022, por el que se establecen normas para la aplicación del Reglamento (UE) 2019/2144 del Parlamento Europeo y del Consejo relativas a los procedimientos uniformes y las especificaciones técnicas para la homologación del sistema de conducción automatizada (ADS) de los vehículos totalmente automatizados.

El Reglamento de Ejecución (UE) 2022/1426 introduce una serie de requisitos y normas comunes y armonizadas cuya finalidad es evaluar, desde un punto de vista técnico, el sistema de conducción automatizada de los vehículos totalmente automatizados, entendiendo por tales aquellos que son capaces de desplazarse de manera autónoma sin que en ningún momento sean asistidos por la supervisión o la intervención de un conductor u operador. Es decir, el objeto del citado reglamento de ejecución es el establecimiento de los requisitos y las condiciones que deben cumplir los sistemas de conducción automatizada de los vehículos totalmente automatizados para su puesta en servicio y acceso al mercado común, sin perjuicio de la facultad de los Estados miembros para regular la circulación y la seguridad del funcionamiento de los vehículos totalmente automatizados cuando transiten en servicios de transporte local.

En España, la reciente modificación del texto refundido de la Ley sobre Tráfico, Circulación de Vehículos a Motor y Seguridad Vial, aprobado por Real Decreto Legislativo 6/2015, de 30 de octubre, que tuvo lugar por medio de la Ley 18/2021, de 20 de diciembre[5], introdujo en el texto legal la base jurídica para promulgar un nuevo conjunto normativo sobre conducción y vehículos automatizados destinado a regular la circulación de este tipo de vehículos, que, dada su naturaleza, trascienden la normativa actual, bajo la premisa contenida en el preámbulo de la Ley, de acuerdo con la cual «[…] el progreso tecnológico de la industria automotriz, está permitiendo el desarrollo de dispositivos y vehículos equipados con diversas tecnologías, que vienen a proponer distintos niveles de automatización, en su grado máximo, la conducción plenamente automatizada o autónoma. Por ello, se impone la necesidad de prever el diseño de un futuro marco normativo que regule la circulación de estos vehículos que por su naturaleza trascienden la regulación actual».

De esta manera, la meritada Ley 18/2021, de 20 de diciembre, ha positivizado una nueva competencia de la Administración General del Estado, la regulación del vehículo automatizado, atribuyendo, por un lado, al Ministerio del Interior el desarrollo de las normas en materia de tráfico y seguridad vial que deberán cumplir los vehículos dotados de un sistema de conducción automatizado, y, por otro, al Ministerio de Industria el desarrollo normativo referido a los requisitos técnicos para la homologación de los mismos. Asimismo, la modificación de 2021 habilita al Gobierno para «regular el procedimiento por el que se certifique que un vehículo dotado de un sistema de conducción automatizado cumple con las normas de circulación, así como la definición de las capacidades de automatización y de los entornos operacionales de uso que se harán constar tanto en el Registro de Vehículos como en los permisos de circulación».

Así las cosas, en el momento de escribir estas líneas se está tramitando un proyecto de real decreto para introducir dos nuevas modificaciones normativas de rango infralegal. Por una parte, se modificará el Reglamento General de Circulación, aprobado por el Real Decreto1428/2003, de 21 de noviembre, para incorporar al texto dos nuevos artículos, a saber, el 3 *bis* y el 18 *bis*, que establecerán las condiciones y requisitos básicos para la circulación de estos nuevos vehículos, así como las obligaciones derivadas de su puesta en circulación. Adicionalmente, se añadirá un nuevo anexo V sobre las normas y condiciones particulares de circulación de los vehículos dotados de un sistema de conducción automatizada, que estará compuesto por dos secciones. La primera contendrá las condiciones generales de circulación que deben satisfacer los vehículos totalmente automatizados, y la segunda establecerá la creación del Manual de Circulación Segura, en el que, en términos generales, se fijarán las condiciones para la circulación segura de estos vehículos cuando transiten por las vías públicas. Por otra parte, se modificará el Reglamento General de Vehículos, aprobado por el Real Decreto 2822/1998, de 23 de diciembre, reforma que incidirá fundamentalmente en los apartados relativos a las autorizaciones administrativas necesarias para la circulación y los referidos a los datos que deben constar en el Registro de Vehículos.

El análisis del marco normativo realizado hasta aquí pretende poner de relieve que los retos que plantean los vehículos autoconducidos reclaman un tratamiento jurídico igualmente autónomo. En último término, la forma en que reaccionen la legislación y la polí-

tica dependerá de la seguridad técnica de este tipo de vehículos y, más aún, de su posible inseguridad y de las personas a las que esta pueda afectar. Estos vehículos solo serán beneficiosos para la sociedad si demuestran ser, en promedio, más seguros que los manejados por un humano; de hecho, muchas voces han sugerido que no debería permitirse su proliferación hasta que su seguridad no quede acreditada de forma fehaciente (así, de forma temprana, Marchant y Lindor, 2012).

Pero incluso una vez que los coches robot superen este listón, necesitaremos dar respuesta al problema de la responsabilidad civil. Para ello, previamente será preciso determinar si los daños que se producen en los accidentes en los que están involucrados estos vehículos reproducen los patrones existentes o si, por el contrario, tienden a causar accidentes diferentes y, por tanto, perjuicios distintos. Pensemos que aún queda por resolver la cuestión de cómo ajustar el equilibrio entre la seguridad entre los pasajeros y la de los demás conductores. Por ejemplo, los coches autoconducidos podrían ser más propensos que los tradicionales a causar desperfectos en otros vehículos o a atropellar a transeúntes que a dañar a los pasajeros. Si se concluyera que los coches autoconducidos son, en general, más seguros, pero tienen más probabilidades de atropellar a niños, ello debería hacer reflexionar al legislador (cfr. The Dawn Project, 2022).

Algunos autores han sugerido que los coches deberían ser capaces de preguntar a los pasajeros cuáles son sus preferencias de conducción; por ejemplo, si deben priorizar la seguridad de las personas dentro o fuera del coche, o si tienen que conducir rápido o con cuidado, con el fin de replicar el comportamiento de conducción humano (Matthias, 2004; Domínguez Álvarez, 2024). Parece poco plausible que esta propuesta prospere y probablemente sea mejor así.

La cuestión de las ventajas y desventajas de la seguridad cautivó la imaginación popular con el «dilema del tranvía» planteado la filósofa norteamericana Judith Jarvis Thomson, que explora las consecuencias morales de diversas acciones y omisiones y contempla accidentes con diferentes tipos de víctimas (Thompson, 1976). Algunos aspectos del dilema del tranvía se ludificaron recientemente en un sitio web llamado Moral Machine[6], que solicita a los participantes que manifiesten su opinión sobre «las decisiones morales tomadas por la inteligencia de las máquinas, como los coches autoconducidos». El sitio web ha atraído a millones de participantes de todo el mundo y la experiencia ha desembocado en una robusta crítica filosófica según la cual el «problema del tranvía es precisamente la herramienta equivocada» para pensar en el modo en que los vehículos automatizados deberían tomar decisiones de vida o muerte mientras se desplazan por diversos tipos de vías: «El problema del tranvía enmarca la cuestión como si todo lo que tuviéramos que hacer es averiguar lo que un individuo debería hacer mientras conduce, y luego hacer que esa sea la regla para los vehículos autónomos» (Suñé Llinás, 2020).

En este sentido, Jacques ha argumentado que las decisiones morales populares basadas en el *crowdsourcing* para generar reglas conducen a resultados repugnantes asociados a la formulación de preguntas equivocadas, el suministro de información errónea o enormemente incompleta para la toma de decisiones, sin dejar de mencionar que a menudo la información proporcionada en el sitio web (por ejemplo, el carácter o las ocupaciones de

las víctimas potenciales) no estaría disponible para un conductor, ya sea humano o robótico. Incluso si la información ofrecida es clara y visible (bebé en cochecito frente a víctima anciana), se confunde una elección individual con una política o algoritmo moralmente sensato. Si dictamináramos que debemos programar nuestros coches para que no den prioridad a los peatones imprudentes de edad avanzada frente a los más jóvenes, por ejemplo, eso equivaldría en la práctica a que el legislador asumiera que hay mayor probabilidad de «condenar» a muerte a los peatones imprudentes de edad avanzada (Jacques, 2019: 7-8). En conclusión, no hay una regla universal que permita dar una respuesta idónea a la cuestión del equilibrio de seguridad de los vehículos automatizados.

Esta reflexión conduce a analizar un segundo tipo de problemas ligados a la responsabilidad civil derivada del uso de estos vehículos: los que conciernen a la imputación. Por el momento, en el mercado están disponibles coches con asistencia al conductor y parcialmente autoconducidos, que devuelven el control al humano cuando el sistema de guía de a bordo se enfrenta a algo que no puede reconocer o con lo que no puede lidiar. Los automóviles con un «nivel 3» de automatización condicional del conductor, tal como son definidos en la clasificación internacional SAE[7], ya son notoriamente desatendidos cuando el sistema de guía está al mando del automóvil en cuestión. A medida que el nivel de automatización aumente hasta el «nivel 4», de alta automatización de la conducción, y las restituciones del control al ser humano sean menos frecuentes, podemos esperar razonablemente que los conductores estarán menos preparados para encarar aquellas contingencias. Llegado el caso, si los pasajeros nunca tienen el control de un vehículo que causa un accidente, ¿podemos exigirles la responsabilidad derivada del mismo? Parecería que no, dado que resultaría absurdo convertirlos de hecho en una suerte de asegurador de las eventualidades en las que el verdadero responsable es el coche robot.

En ausencia de un daño intencionado, lo habitual es imputar la responsabilidad de un accidente a la parte que pueda evitarlo con el menor coste social. Según la teoría económica, de este modo es más probable que se alineen los incentivos para evitar accidentes sin desincentivar demasiado las actividades que puedan causar daños. Entonces, ¿quién es el que puede evitar el daño de una manera más económica en un accidente causado por un vehículo totalmente autónomo? La respuesta es difusa y podría depender del contexto. Por ejemplo, el proveedor de un servicio de taxi podría convertirse en el titular de la responsabilidad si se recurre al argumento de que ejerció su libertad a la hora de elegir el tipo de taxi robot que adquirió. Pero lo más plausible es atribuir al fabricante la responsabilidad por los accidentes provocados por vehículos totalmente automatizados, dado que creó el *hardware* y diseño o eligió el *software* que ejecuta el vehículo autónomo y, por lo tanto, está mejor posicionado para determinar la inversión en pruebas orientadas a mejorar su seguridad y sus características.

Sin embargo, esta no es una respuesta clara y satisfactoria a la cuestión de la responsabilidad, al menos no en el Derecho norteamericano que regula en este campo de forma pionera, seguramente porque en este país los elevados costes médicos pueden incrementar significativamente la cantidad reclamada en cualquier litigio posterior relativo a la responsabilidad. Pues bien, a falta de legislación que modifique las normas de responsabilidad

civil, seguramente el modo en que Estados Unidos se determine la responsabilidad por los perjuicios causados por los coches autónomos —y, probablemente, por muchos otros tipos de robots— dependerá primeramente de si el robot se presenta como un producto o un servicio.

Si el robot se presenta como un servicio —por ejemplo, un taxi u otra prestación a la carta—, se aplicarán las normas ordinarias de responsabilidad civil, ya que el consumidor no es el comprador del producto y el proveedor del servicio puede estar sujeto a las regulaciones estatales de responsabilidad civil por producto (*products liability*), como se verá adelante. Dicho esto, incluso para los consumidores, abundan las complejidades en la determinación del causante de la negligencia en cuestión y del estándar de cuidado que debería adoptarse para medirla (Scherer, 2015). Por ejemplo, un coche robótico puede considerarse un producto cuando se vende. Pero si el *software* que lo hace funcionar se vendiera por separado y se actualizara continuamente, podría considerarse un servicio independiente. Por otra parte, si solo hay una opción de *software* para una determinada marca o modelo de coche autónomo, también podría argumentarse que ambos están inextricablemente unidos y que el *hardware* predomina en la transacción global, por lo que el *software* debería considerarse un producto.

Si los usuarios finales compran (o alquilan a largo plazo) su vehículo, incluido el *software* como hace actualmente la mayoría de los conductores estadounidenses, es de suponer que la legislación de Estados Unidos dispondrá que la ley de responsabilidad por productos defectuosos es aplicable a todo el vehículo. Si hay un defecto de fabricación, es decir, si debido a un error de fábrica el producto se desvía del diseño previsto, los estados siguen de modo uniforme el Restatement of Torts, Second § 402A (1965)[8] y atribuyen la responsabilidad solidaria al fabricante y a todas las partes que intervienen la cadena de venta del producto hasta que llega al usuario final. Ahora bien, ¿qué ocurre si el fabricante produjo el vehículo autónomo según lo previsto, pero existe un defecto de diseño subyacente? En este caso, los órganos de adjudicación de los estados no han establecido un criterio uniforme a la hora de determinar si se aplica una norma de culpa o de responsabilidad objetiva. Por tanto, tampoco hay uniformidad en relación con las pruebas que una víctima —ya sea el usuario final o un tercero— debe aportar en juicio para obtener una indemnización por los daños y perjuicios causados por el fabricante.

Con arreglo al Restatement of Torts, Second, modificado en los años sesenta y desarrollado posteriormente en la interpretación de la norma realizada por serie de resoluciones judiciales durante las décadas siguientes, la regla mayoritaria para los defectos de diseño pasó a ser la «prueba de las expectativas del consumidor», que atribuye la responsabilidad al vendedor de un producto si este se encuentra en un estado defectuoso irrazonablemente peligroso para el consumidor[9]. De acuerdo con este criterio, el juzgador de los hechos puede inferir la existencia de un defecto de diseño si el producto no satisface las expectativas razonables de los consumidores. Ello equivale a decir que el demandante puede obtener su derecho a una indemnización si prueba que el producto no funcionó de acuerdo con unas expectativas razonables como usuario, aunque no disponga de evidencias, directas o circunstanciales, que demuestren exactamente qué tipo de defecto de fabricación existía.

Como puede suponerse, la prueba de las expectativas del consumidor ha suscitado importantes críticas doctrinales por este y otros motivos (Owen y Davis, 2024, en § 5:16). Sin embargo, la principal ventaja del criterio probatorio de las expectativas es su aparente sencillez: cuando el producto falla en aquellos aspectos respecto a los que un consumidor medio podría tener expectativas bastante definidas, se aplica una regla de responsabilidad objetiva.

Sea como fuere, no parece que esta doctrina pueda ser aplicada a los vehículos autónomos. La prueba de las expectativas del consumidor no es aplicable cuando se alegan defectos técnicos y mecánicos que requieren la comprensión de comportamientos precisos de componentes opacos de productos en las complejas circunstancias que rodean a un accidente concreto[10]. La determinación exacta del momento en que los robots pasan de ser máquinas con piezas técnicas y opacas a productos tan comunes como los coches, los microondas o los frigoríficos es una cuestión en evolución y todavía en gran parte no resuelta. Las aspiradoras roomba son omnipresentes y, sin duda, entran en la categoría de producto de consumo ordinario, al igual que los drones y las mascotas robóticas, pero ¿lo haría el *software* de control de un coche de conducción autónoma, a diferencia de un Tesla con asistencia al conductor de nivel 3? ¿O la respuesta es que los consumidores no pueden alegar la doctrina de la falta de expectativas en los productos de Tesla porque Elon Musk proclama constantemente la excelencia del *software* de la compañía a través de sus tuits?

Además, hay que tener en cuenta que, posteriormente, el Third Restatement of Torts: Responsabilidad por productos defectuosos, § 2, ha modificado el modo en que deben construirse las demandas y las defensas en los casos de defectos de diseño. Durante muchos años, los críticos de la prueba de las «expectativas del consumidor» objetaron que los demandantes podían triunfar sin demostrar defectos en el diseño, y que asimismo no tenían obligación de presentar un diseño alternativo razonable. En cambio, si la defensa pretendía argumentar que no había un diseño práctico mejor, soportaba la carga de la prueba sobre esta cuestión una vez que el demandante hubiera demostrado *prima facie* lo que los consumidores esperarían del producto. Quizás por ello, la prueba de las «expectativas del consumidor» ha sido sustituida por la de la «utilidad del riesgo» (Twerski y Henderson, 2021). Con arreglo a esta doctrina, para que sea estimada la pretensión del demandante este debe demostrar que el riesgo del diseño supera su valor, lo que se traduce en la necesidad de presentar un diseño alternativo más seguro —en tanto que no hubiera causado el daño o, al menos, reducido en gran medida su probabilidad— y cuya producción no sea (sustancialmente) más costosa.

Eta reforma también ha suscitado una notable controversia y, de hecho, un número considerable de tribunales supremos estatales han decidido no adoptar una disposición del tercer Restatement[11], que, entienden, se aleja del criterio de responsabilidad objetiva, pues se centra en la previsibilidad del riesgo del daño e incluye un análisis de costes y beneficios: «En lugar de centrarse en el diseño del producto, se centra en la conducta del fabricante» (por todos, Owen y Davis, 2024: §§ 5.6 y 5.7). Dado que el demandado dispone de toda la información sobre el diseñó y la fabricación del robot, para la mayoría de los demandantes será difícil cumplir esta carga probatoria. En la mayoría de los casos, será necesario recurrir

a costosas pruebas periciales, lo que dificultará mucho la interposición de demandas y, eventualmente, obligará a los jurados a atender explicaciones probatorias excesivamente técnicas.

En resumen, en la actualidad la responsabilidad por lesiones causadas por robots en los Estados Unidos presenta hasta tres niveles de ambigüedad concretados en las siguientes cuestiones: *i)* si el robot se presenta como producto o servicio; *ii)* si el tipo de robot o la parte relevante del mismo es un producto de consumo ordinario sobre el que los consumidores tienen expectativas fundadas; *iii)* si el estado cuya ley se aplica al accidente utiliza la segunda edición del Restatement de responsabilidad objetiva por defectos de diseño o la tercera edición del Restatement, con reglas más parecidas a la culpa.

La regulación de la responsabilidad es diferente en de la Unión Europea. Por ejemplo, y como ha señalado Ebers respecto a la legislación alemana, a pesar de que la Directiva 85/374/CEE del Consejo, de 25 de julio de 1985[12] establece la responsabilidad objetiva de los fabricantes de productos defectuosos, en la práctica la responsabilidad principal recaerá en el poseedor del vehículo, debido a la dificultad de probar la culpa contra los fabricantes (Ebers, 2022). Esta conclusión puede extrapolarse a la mayoría de los Estados miembros, incluido nuestro país. Como ya sostuve en su momento (Barrio Andrés, 2019), en España la suficiencia en la responsabilidad basada en la culpa que regula nuestro Código Civil se volverá incierta en relación con los vehículos autónomos. Por ejemplo, no será fácil determinar quién es el «conductor» en un vehículo autónomo y si este puede ser considerado responsable de una violación de la ley cometida por el propio vehículo (por ejemplo, cruzar un semáforo en rojo).

Precisamente para evitar este tipo de incertidumbres, la regulación del nuevo Reglamento Europeo de Inteligencia Artificial iba a ser complementada por dos proyectos de Directivas cuya tramitación ha quedado interrumpida tras las elecciones al Parlamento Europeo celebradas en junio de 2024.

Por una parte, la Propuesta de Directiva del Parlamento Europeo y del Consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial (Directiva sobre responsabilidad en materia de IA) pretendía facilitar la prueba de la culpa y de la relación de causalidad en el caso de daños ocasionados por sistemas de inteligencia artificial que debían resolverse de acuerdo con las respectivas legislaciones nacionales sobre la responsabilidad por culpa.

Por otra, la Propuesta[13] de Directiva del Parlamento Europeo y del Consejo sobre responsabilidad por los daños causados por productos defectuosos tenía por objeto sustituir la aún vigente Directiva 85/374/CEE en materia de responsabilidad por los daños causados por productos defectuosos por una regulación adaptada a las nuevas necesidades de la IA y, también, de la economía circular.

En el momento de escribir estas líneas, el panorama de la regulación de la responsabilidad derivada de la utilización de vehículos autónomos es incierto. Desde el punto de vista de la instancia reguladora, es probable que las respuestas de los Estados no sean uniformes. La cuestión de la responsabilidad por accidentes de robots parece madura

para una solución nacional uniforme (tanto en Estados Unidos como en la propia Unión Europea). Idealmente, esta regulación debería reducir los costes sustanciales de litigación sobre estas cuestiones en cada estado o país. La consecución de este objetivo podría implicar la aprobación de una nueva legislación, o incluso la creación un regulador nacional de robots (Calo, 2014), si bien en algunos sectores, como el de los coches autoconducidos, también podría hallarse una solución parcial *de facto* si los fabricantes optan por incluir un seguro de coche para los pasajeros en cada venta con el fin de publicitar la seguridad de su producto.

Sea como fuere, considero que la regulación de la responsabilidad en asuntos relacionados con el tráfico transitará del procedimiento basado en el criterio la culpa a la adopción de modalidades de responsabilidad objetiva o de gestión de riesgos. A resultas de esta evolución, las víctimas tendrán que demandar a otros actores para reclamar la indemnización, entre ellos el constructor del vehículo o el fabricante del *software*. También surgirán nuevos modelos de aseguramiento, o incluso, como acertadamente apuntó Froomkin, «[…] un fondo general recaudado mediante impuestos (o mediante primas de seguros) pagados por los fabricantes, los usuarios o incluso todos nosotros» (Froomkin, 2018: 24).

En esta línea, Abraham y Ravin argumentan sólidamente que la previsible generalización del uso de los vehículos autónomos constituye una oportunidad para alejarse de los viejos paradigmas de responsabilidad civil por accidentes de tráfico y asumir un esquema de compensación administrativa (Abraham y Rabin, 2019). Sería interesante ver hasta qué punto en el futuro esta solución podría generalizarse a otros accidentes relacionados con los robots.

4. ROBOTS Y EMPLEO

Si los robots no nos matan, quizá solo nos quiten el trabajo. La gavilla de tareas que los robots (y las IA) parecen capaces de realizar —desde la construcción y el almacenaje hasta la conducción de camiones, pasando por trabajos altamente cualificados en los ámbitos financiero, jurídico y médico— suscita una inevitable preocupación por la posibilidad de que los robots desplacen a los trabajadores a gran escala. De hecho, en todo el mundo hay evidencia de que las empresas tratan de sustituir a los trabajadores —no solo a los empleados industriales y minoristas, sino también, cuando sea posible, a los profesionales— por robots, ya que estos pueden trabajar 24 horas al día y 7 días a la semana, a todos se les pueden enseñar nuevos comportamientos simultáneamente, son resistentes a las pandemias (aunque vulnerables a los virus informáticos) y no se declaran en huelga.

Acemoglu y Restropo calcularon que, en Estados Unidos, un robot industrial reprogramable más por cada mil trabajadores reduce la relación empleo-población agregada en aproximadamente 0.2 puntos porcentuales, y los salarios en aproximadamente un 0.42 % a nivel nacional. En la zona donde se despliega el robot, la relación empleo-población disminuye un 0.77 % y los salarios un 0.77 % (Acemoglu y Restropo, 2020). Por tanto, cada

robot añadido a una zona de desplazamiento (esto es, un área geográfica utilizada para el análisis económico) reduce el empleo en seis trabajadores en esa zona. Ambos autores conjeturan que los efectos futuros serán mayores a medida que la industria utilice más robots, y concluyen señalando que «[…] las empresas que adoptan tecnologías de automatización reducen sus costes y pueden expandirse a expensas de sus competidores». ¿Será preciso que en el proceso de sustitución de los trabajadores por robots tenga lugar una carrera a la baja para asegurar la pervivencia de las empresas?

En la actualidad, no hay consenso sobre si el creciente uso de robots (y de las IA) causará un desempleo permanente sustancial o si simplemente atravesaremos un periodo provisional de desajuste entre las cualificaciones y los puestos de trabajo. Pero incluso un efecto temporal podría afectar a una generación y ser grave no solo entre los empleados industriales, sino también entre otros trabajadores poco cualificados potencialmente reemplazables como el millón de camioneros que tienen pocas cualificaciones transferibles y los cerca de diez millones de cajeros, vendedores y supervisores minoristas de primera línea, cifras estimadas para el caso de los Estados Unidos (Anderson, 2020).

Además, y con carácter general, en la medida en que los robots exigen gastos de capital, la relación entre capital y trabajo se ve alterada. Con menos trabajadores, los índices de productividad de los que siguen empleados deberían aumentar, de manera que muchos de los que no son desplazados, incluidos los profesionales de cuello blanco, pueden verse sometidos a un control y una supervisión administrados por robots en los que cada acción sea computada en busca de la máxima productividad (Sánchez-Urán Azaña, López Sánchez y Grau Ruiz, 2019). Quizás por ello, la opinión pública ha empezado a cobrar conciencia de la amenaza que pueden representar los robots para el empleo, aunque hasta ahora los robots tienden a ser más impopulares en los Estados en los que hay mayores niveles de desigualdad (Macías García, 2024).

De momento, la estrategia para evitar el desplazamiento de trabajadores por robots parece centrarse en el ámbito fiscal. Si la utilización de de robots resulta atractiva desde la perspectiva del coste que supone la protección de trabajadores humanos, una reforma fiscal puede contribuir a equilibrar la tendencia.

La doctrina ha prestado mucha atención a las cuestiones fiscales, ya que afectan a la decisión de invertir en robots, y también a la posibilidad de utilizar la política fiscal para proteger algunos puestos de trabajo del desplazamiento de los humanos por robots (por ejemplo, Segura Alastrué, 2018). Los investigadores sugieren que el sistema fiscal vigente crea un incentivo para sustituir la mano de obra por capital, pues la inversión de capital en maquinaria es un coste deducible o crea un activo amortizable, en contraste con el carácter no deducible de los salarios. Igualmente, la obligación legal de pagar el seguro de desempleo y otros beneficios sociales a los trabajadores humanos, pero no a los robots, puede ser un incentivo tácito de la robotización.

Algunos miembros del Parlamento Europeo han propuesto la creación de un impuesto sobre los robots tanto para desincentivar esa sustitución como para crear un fondo público que ayude a los trabajadores desplazados, pero el Derecho europeo no ha asumido esta

propuesta. Hasta ahora, la idea ha fracasado por el temor a desincentivar la innovación y la dificultad de medir o definir qué número o fracción de puestos de trabajo han sido desplazados por la introducción de determinada máquina. Las propuestas de gravar los robots tropezarán también con otros problemas, entre ellos determinar cuáles serían los robots imponibles y cómo decidir el nivel de imposición por robot más adecuado. En el fondo, este debate académico se ve ensombrecido por la conciencia de que las previsiones sobre los efectos destructores del empleo generados por la automatización que fueron frecuentes en los años sesenta resultaron totalmente infundadas (Sánchez Castro Díaz Guerra, 2023). ¿Será diferente esta vez?

5. CONCLUSIONES

Como es sabido, cada nueva tecnología importante —ya sea la electricidad, el ferrocarril, Internet, la IA o los robots— crea nuevos problemas éticos, jurídicos y sociales. Actualmente, nos encontramos en las primeras etapas de la robotización generalizada: la calidad, la cantidad y la variedad de los robots evidencian un rápido crecimiento del fenómeno. Este avance tecnológico genera un trabajo importante y necesario para los legisladores, las autoridades reguladoras, los jueces y los académicos, si lo que se pretende es maximizar los beneficios de la robotización y minimizar sus daños. El presente estudio, que no es más que una instantánea en el tiempo, deja muchas cosas fuera, pero aun así pone de manifiesto la trascendencia de las tareas que tenemos ante nosotros.

En ocasiones, las tecnologías disruptivas también crean oportunidades de cambio en los acuerdos sociales fundamentales. La robotización del trabajo, por ejemplo, ofrece la esperanza de liberarse de los riesgos laborales y la monotonía, pero en algunos escenarios podría abrir la puerta al desempleo masivo y la miseria. Las armas autónomas seducen porque su utilización puede reducir las bajas militares humanas, pero también presentan claros peligros, entre ellos la atenuación de la responsabilidad de la cadena de mando, el aumento de las matanzas civiles o la tentación que para algunos dirigentes sin escrúpulos pueda suponer el hecho de que el ejército agresor en una guerra sufra pocas bajas.

Las probabilidades de que los robots se conviertan en nuestros amos en un futuro próximo parecen insignificantes, pero es cualquier cosa menos inevitable que todos los cambios generados por los robots sean beneficiosos y que utilicemos los robots en beneficio de la humanidad. Los retos para los juristas, los responsables políticos, los expertos en ética y para la ciudadanía en general están ahí para ser resueltos, entre otras herramientas, con los principios generales del Derecho de los Robots que propuse en la 2.ª edición de *Derecho de los Robots*, publicada en 2019 y las reformas legislativas que sean necesarias.

NOTAS

1. 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).

2. Vid., por ejemplo: https://www.xatakahome.com/a-fondo/amazon-mantiene-todas-tus-grabaciones-voz-alexa-nube-asi-puedes-eliminarlas-saber-cuando-te-escucha-tu-altavoz

3. En Europa, uno de los Estados miembros pioneros en el abordaje de esta cuestión ha sido Alemania. Ya en julio de 2016, el entonces ministro federal alemán de Transporte e Infraestructura Digital, Alexander Dobrindt, nombró un comité nacional de ética para la conducción automatizada y conectada, que comenzó su actividad en septiembre de 2016. En junio de 2017, este comité presentó un código ético que fue asumido por el Gobierno federal. Disponible en: https://www.bmdv.bund.de/SharedDocs/DE/Publikationen/DG/bericht-der-ethik-kommission.pdf?__blob=publicationFile

4. Reglamento (UE) 2019/2144 del Parlamento Europeo y del Consejo, de 27 de noviembre de 2019, relativo a los requisitos de homologación de tipo de los vehículos de motor y de sus remolques, así como de los sistemas, componentes y unidades técnicas independientes destinados a esos vehículos, en lo que respecta a su seguridad general y a la protección de los ocupantes de los vehículos y de los usuarios vulnerables de la vía pública, por el que se modifica el Reglamento (UE) 2018/858 del Parlamento Europeo y del Consejo y se derogan los Reglamentos (CE) n.º 78/2009, (CE) n.º 79/2009 y (CE) n.º 661/2009 del Parlamento Europeo y del Consejo y los Reglamentos (CE) n.º 631/2009, (UE) n.º 406/2010, (UE) n.º 672/2010, (UE) n.º 1003/2010, (UE) n.º 1005/2010, (UE) n.º 1008/2010, (UE) n.º 1009/2010, (UE) n.º 19/2011, (UE) n.º 109/2011, (UE) n.º 458/2011, (UE) n.º 65/2012, (UE) n.º 130/2012, (UE) n.º 347/2012, (UE) n.º 351/2012, (UE) n.º 1230/2012 y (UE) 2015/166 de la Comisión.

5. Ley 18/2021, de 20 de diciembre, por la que se modifica el texto refundido de la Ley sobre Tráfico, Circulación de Vehículos a Motor y Seguridad Vial, aprobado por el Real Decreto Legislativo 6/2015, de 30 de octubre, en materia del permiso y licencia de conducción por puntos.

6. Disponible en: https://www.moralmachine.net/

7. Vid. SAE, Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles J3016_202104, de 30 de abril de 2021. Disponible en: https://www.sae.org/standards/content/j3016_202104/

8. Vid. Restatement of Torts, Second § 402A (1965).

9. Vid. Sección § 402A.

10. Vid. Soule vs. General Motors Corp., 882 P.2d 298, 305 (Cal. 1994).

11. Por todos, vid. Aubin vs. Union Carbide Corp., 177 So. 3d 489, 506 (Fla. 2015).

12. Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados miembros en materia de responsabilidad por los daños causados por productos defectuosos.

13. Estando este trabajo a punto de imprimirse, ha sido publicada la nueva Directiva (UE) 2024/2853 del Parlamento Europeo y del Consejo, de 23 de octubre de 2024, sobre responsabilidad por los daños causados por productos defectuosos y por la que se deroga la Directiva 85/374/CEE del Consejo.

BIBLIOGRAFÍA

ABRAHAM, Kenneth y Robert RABIN (2019): «Automated Vehicles and Manufacturer Responsibility for Accidents: A New Legal Regime for a New Era», *Virginia Law Review*, 105, 127-171.

ACEMOGLU, Daron y Pascual RESTREPO (2020): *Robots and Jobs: Evidence from US Labor Markets*. Chicago: The University of Chicago Press.

AMIDEI, Andrea (2019): «Intelligenza Artificiale e "product liability": sviluppi del diritto dell'Unione Europea», *Giurisprudenza Italiana*, 7.

ANDERSON, D. Augustus (2020): «Retail Jobs Among the Most Common Occupations» [en línea] *U.S. Census* <https://www.census.gov/library/stories/2020/09/profile-of-the-retail-workforce.html>. [Consulta: 12/07/2023.]

ARELLANO TOLEDO, Wilma (dir.) (2023): *Derecho, Ética e Inteligencia Artificial*. Valencia: Tirant lo Blanch.

BARONA VILAR, Silvia (2021): *Algoritmización del Derecho y de la Justicia. De la inteligencia artificial a la Smart Justice*. Valencia: Tirant lo Blanch.

— (2021): *Justicia algorítmica y neuroderecho. Una mirada multidisciplinar*. Valencia: Tirant lo Blanch.

BARRIO ANDRÉS, Moisés (dir.) (2018): *Derecho de los Robots*. Madrid: Wolters Kluwer.

— (2019): «Consideraciones jurídicas acerca del coche autónomo», *Actualidad jurídica Uría Menéndez*, 52, 101-108.

— (2022*a*): *Internet de las Cosas*. Madrid: Editorial Reus, 3.ª edición.

— (2022*b*): «Inteligencia artificial: origen, concepto, mito y realidad», *El Cronista del Estado Social y Democrático de Derecho*,100.

— (2023): *Los derechos digitales y su regulación en España, la Unión Europea e Iberoamérica*. A Coruña: Colex.

— (2024*a*): «De nuevo sobre la persona robótica», *Inteligencia Artificial*, 73, 65-79.

— (dir.) (2024*b*): *El Reglamento Europeo de Inteligencia Artificial*. Valencia: Tirant lo Blanch.

— (2024*c*): *Manual de Derecho digital*. Valencia: Tirant lo Blanch, 3.ª edición.

— (dir.) (2024*d*): *Comentarios al Reglamento Europeo de Inteligencia Artificial*. Madrid: La Ley.

BAYERN, Shawn (2021): *Autonomous Organizations*. Cambridge: Cambridge University Press.

BRENNAN-MARQUEZ, Kiel, Karen LEVY y Daniel SUSSER (2019): «Strange Loops: Apparent versus Actual Human Involvement in Automated Decision Making», *Berkely Tehnolgoy Law Journal*, 34, 745-771.

CALO, Ryan (2014): «The Case for a Federal Robotics Commission», *Brookings Institution Center for Technology Innovation* [en línea] <https://www.brookings.edu/articles/the-case-for-a-federal-robotics-commission/>. [Consulta: 15/12/2024.]

CAMPUZANO GÓMEZ-ACEBO, Jimena y Jesús SIEIRA GIL (2023): «Legal Tech y función registral», en Barrio Andrés, Moisés (dir.): *Legal Tech. La transformación digital de la abogacía*. Madrid: Editorial La Ley, 2.ª edición.

CASTILLA BAREA, Margarita (2023): «¿Qué queda hoy de la resolución del parlamento europeo de 20 de octubre de 2020 sobre un régimen de responsabilidad civil en materia de inteligencia artificial?», en M.D. CERVILLA GARZÓN, y A.M. BALLESTEROS BARROS (dirs.): *Temas actuales de Derecho Privado II*, Cizur menor: Aranzadi.

CERDEIRA BRAVO DE MANSILLA, Guillermo (2024): «¿Humanizar o personificar? Inteligencia artificial y fundaciones robóticas», *Actualidad civil*, 3.

COECKELBERGH, Mark (2010): «Robot rights? Towards a social-relational justification of moral consideration», *Ethics and Information Technology*, 12, 209-221.

DARLING, Kate (2016): «Extending Legal Protection to Social Robots», en R. CALO, A. FROOMKINE I. KERR (eds.): *Robot Law*, Cheltenham: Editorial Edward Elgar.

DOMÍNGUEZ ÁLVAREZ, José Luis (2024): *Iusalgoritmia: las reglas de derecho y la inteligencia artificial*. Valencia: Tirant lo Blanch.

DOMÍNGUEZ PECO, E. (2018): «Los robots en el Derecho penal», en M. BARRIO ANDRÉS, (dir.) (2018): *Derecho de los Robots*. Madrid: Wolters Kluwer.
EBERS, Martin (2022): «Civil Liability for Autonomous Vehicles in Germany», *Humboldt University of Berlin* [en línea] < https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4027594>. [Consulta: 15/12/2024.]
Elish, Madeleine (2019): «Moral Crumple Zones: Cautionary Tales in Human Robot Interaction», *Engaging Science, Technology, and Society*, 5.
FAGGIANI, Valentina y Gabrielle SALES SARLET (dirs.) (2024): *Retos del Derecho ante la IA. Apuntes desde una perspectiva interdisciplinar*. Barcelona: J.M. Bosch Editor.
FROOMKIN, A. Michael (2018): «Prólogo», en M. BARRIO ANDRÉS (dir.) (2018): *Derecho de los Robots*. Madrid: Wolters Kluwer.
GARCÍA-PRIETO CUESTA, Juan (2018): «¿Qué es un robot?», en M. BARRIO ANDRÉS (dir.) (2018): *Derecho de los Robots*. Madrid: Wolters Kluwer.
GÓMEZ-RIESCO TABERNERO DE PAZ, José María (2018): «Los robots y la responsabilidad civil extracontractual», en M. BARRIO ANDRÉS (dir.) (2018): *Derecho de los Robots*. Madrid: Wolters Kluwer.
HARTZOG, Woodrow (2015): «Unfair and Deceptive Robots», en *Maryland Law Review*, 74, 785-829.
JACQUES, Abby Everett (2019): «Why the Moral Machine is a Monster», *University of Miami Law School: We Robot Conference* [en línea] <https://robots.law.miami.edu/2019/wp-content/uploads/2019/03/MoralMachineMonster.pdf>. [Consulta: 15/12/2024.]
KAMINSKY, Margot (2015): «Robots in the Home What Will We Have Agreed To?», *Idaho Law Review*, 51, 661-677.
LLANO ALONSO, Fernando (2018): *Homo Excelsior. Los límites ético-jurídicos del transhumanismo*. Valencia: Tirant lo Blanch.
— (dir.) (2021): *Inteligencia artificial y derecho. El jurista ante los retos de la era digital*. Pamplona: Aranzadi.
— (dir.) (2022): *Inteligencia Artificial y Filosofía del Derecho*. Murcia: Ediciones Laborum.
— (2024): *Homo Ex Machina. Ética de la inteligencia artificial y Derecho digital ante el horizonte de la singularidad tecnológica*. Valencia: Tirant lo Blanch.
LOPUKI, Lynn (2018): «Algorithmic Entities», *Washington University Law Review*, 4, 95, 887-953.
MACÍAS GARCÍA, Marina (2024): *El impacto de la robótica y la Inteligencia Artificial en la Seguridad y Salud de las personas trabajadoras*. Murcia: Ediciones Laborum.
MARCHANT Gary y Rachel LINDOR (2012): «The coming collision between autonomous vehicles and the liability system», *Santa Clara Law Review*, 52, 1321-1340.
MARCHENA GÓMEZ, Manuel (2022): *Inteligencia artificial y jurisdicción penal*. Discurso de ingreso en la Real Academia de Doctores.
MATTHIAS, Andreas (2004): «The responsibility gap: Ascribing responsibility for the actions of learning automata», *Ethics and information technology*, 6(3), 175-183.
MUÑOZ GARCÍA, Carmen (2023): *Regulación de la inteligencia artificial en Europa. Incidencia en los regímenes jurídicos de protección de datos*. Valencia: Tirant lo Blanch.
NASS Clifford y Youngme MOON (2000): «Machines and Mindlessness: Social Responses to Computers», *Journal of Social Issues*, 56, 81-103.
NYHOLM Sven (2020): *Humans and robots: ethics, agency, and anthropomorphism*. Lanham: Editorial Rowman & Littlefield Publishers.
ORTEGO RUIZ, Miguel (2022): *La personalidad jurídica internacional de los robots*. Valencia: Tirant lo Blanch.
OWEN, David y Mary DAVIS (2024): *Products Liability*. Eagan: Thompson Reuters.
POWER, M. (2014): «What happens when a software bot goes on a darknet shopping spree?», *The Guardian* [en línea] <https://www.theguardian.com/technology/2014/dec/05/software-bot-darknet-shopping-spree-random-shopper>. [Consulta: 15/12/2024.]

PUZIO, Anna (2023): «Robot, let us pray! Can and should robots have religious functions? An ethical exploration of religious robots», *AI & Society*, 1-17.
RICHARDS, Neil y William SMART (2016): «How Should the Law Think About Robots?», en R. Calo, A. Froomkin e I. Kerr (eds.): *Robot Law*, Cheltenham: Edward Elgar.
ROSENTHAL-VON DER PUTTEN, Astrid, Nicole KRAMER, Laura HOFFMANN, Sabrina SOBIERAJ y Sabrina EIMLER (2013): «An experimental study on emotional reactions towards a robot», *International Journal of Social Robotics*, 5(1), 17-34.
ROSINO CALLE, Roberto Carlos (2023): «Hacia una protección efectiva de la propiedad intelectual desde el Convenio Europeo de Derechos Humanos», *Revista LA LEY Derecho Digital e Innovación*, 16.
SÁNCHEZ CASTRO DÍAZ GUERRA, Consuelo (2023): *Estatus Jurídico Constitucional de los Robots*. Valencia: Tirant lo Blanch.
SÁNCHEZ-URÁN AZAÑA, Yolanda, José Ignacio LÓPEZ SÁNCHEZ y Maria Amparo GRAU RUIZ (2019): «The impact of robotics and computerization on the labour market: inclusive insight from a Law and Economics perspective», *Revista LA LEY de Derecho Digital e Innovación*, 3.
SAWERS, Paul (2020): «Chinese court rules AI-written article is protected by copyright», *VentureBeat* [en línea] <https://venturebeat.com/ai/chinese-court-rules-ai-written-article-is-protected-by-copyright/>. [Consulta: 15/12/2024.]
SCHERER, Matthew (2015): «Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies», *Harvard Journal of Law & Technology*, 29, 354-398.
SEGURA ALASTRUÉ, Mario (2018): «Los robots en el derecho financiero y tributario», en M. BARRIO ANDRÉS (dir.) (2018): *Derecho de los Robots*, Madrid: Wolters Kluwer.
SUÑÉ LLINÁS, Emilio (2020): *Derecho e Inteligencia Artificial. De la robótica a lo posthumano*. México: Tirant lo Blanch.
THOMSON, Judith (1976): «Killing, letting die, and the trolley problem», *The Monist*, 59, 204-217.
TWERSKI, Aaron y James HENDERSON (2021): *Products Liability: Problems and Process*, Nueva York: Aspen Opco.
WESSLING, Bernhard (2022): «Cruise hits milestone by charging for robotaxis rides in SF», *The Robot Report* [en línea] <https://www.therobotreport.com/cruise-begins-charging-the-public-for-robotaxis-rides/>. [Consulta: 15/12/2024.]
ZHANG, Zhuo *et al.* (2023): «Dataset and Baselines for IID and OOD Image Classification Considering Data Quality and Evolving Environments», *International Journal of Interactive Multimedia and Artificial Intelligence*, 8(1) 6-12.

Fecha de recepción: 1 de julio de 2024.
Fecha de aceptación: 20 de octubre de 2024.

INTELIGENCIA ARTIFICIAL Y ADMINISTRACIONES PÚBLICAS: PARA QUÉ Y CÓMO*

ARTIFICIAL INTELLIGENCE AND PUBLIC ADMINISTRATIONS: FOR WHAT AND HOW

Alejandro Huergo Lora

Catedrático de Derecho Administrativo
Universidad de Oviedo

RESUMEN

Existen distintos grados y formas de utilizar la tecnología digital en la actividad administrativa: desde la comunicación electrónica con los ciudadanos («administración electrónica») hasta la inteligencia artificial, pasando por la introducción de aplicaciones que roboticen actividades administrativas regladas. La inteligencia artificial, basada en predicciones algorítmicas (incluidos instrumentos de inteligencia artificial generativa y modelos fundacionales), puede servir, entre otras cosas, para concretar conceptos probabilísticos o predictivos (como el de «riesgo») o en fases preparatorias de la actividad de la Administración, por ejemplo a la hora de distribuir los recursos disponibles o de dirigir en otra dirección inspecciones o medidas de apoyo. Utilizarla para la adopción de decisiones finales es muy cuestionable, incluso por razones estructurales. En cuanto al control y regulación, es necesario distinguir la utilización de inteligencia artificial para elaborar y para justificar actuaciones administrativas.

PALABRAS CLAVE

Inteligencia artificial – algoritmos – Administración Pública – procedimiento administrativo.

ABSTRACT

There are different degrees and ways of using digital technology in administrative activity: from electronic communication with citizens («e-government») to artificial intelligence, including the introduction of applications that robotize administrative activities. Artificial intelligence, based on algorithmic predictions (including generative artificial intelligence instruments and foundation models), can serve, among other things, to specify probabilistic or predictive concepts (such as «risk») or as a preparation for administrative action, i.e. when distributing available resources or targeting inspections or support measures. Using it for final decision-making is highly questionable, even for structural reasons. Regarding control and regulation, it is necessary to distinguish the use of artificial intelligence to prepare and to justify administrative actions.

KEY WORDS

Artificial intelligence – algorithms – Public administration – Administrative procedure.

DOI: https://doi.org/10.36151/TD.2024.105

---

* Este trabajo se enmarca en el proyecto de investigación PID2021-126881OB-I00, financiado por el Ministerio de Ciencia e Innovación y titulado «Herramientas algorítmicas para ciudadanos y administraciones públicas». Agradezco a Gustavo Manuel Díaz González la revisión y valiosas sugerencias al borrador.

INTELIGENCIA ARTIFICIAL Y ADMINISTRACIONES PÚBLICAS: PARA QUÉ Y CÓMO

Alejandro Huergo Lora
Catedrático de Derecho Administrativo
Universidad de Oviedo

1. INTRODUCCIÓN

Desde aproximadamente 2020 se han multiplicado las noticias sobre la inteligencia artificial. No es sorprendente que se haya hablado de «ansiedad algorítmica» (Kearns/Roth). Con independencia de los avances técnicos concretos que se han ido conociendo (gran parte de los cuales se produjeron en la década entre 2010 y 2020, antes de esta popularidad de la inteligencia artificial, con la excepción de los llamados «modelos fundacionales»,

«modelos de uso general» o grandes modelos de lenguaje, difundidos a partir de 2022), la inteligencia artificial es un tema de moda, utilizado incluso para llamar la atención y captar audiencia de manera forzada. En su presentación se ha pasado de un enfoque «optimista», en el que se veía en ella la solución para los grandes retos sociales y una gran promesa de desarrollo económico, a otro «pesimista» en el que ha pasado a convertirse —supuestamente— en una amenaza existencial para la humanidad.

Por otro lado, es notorio que sobre esta materia ha habido una gran profusión doctrinal (que me sorprende y me abruma, a decir verdad), a la que yo mismo he contribuido, y que no me sorprendería que produjera en gran parte de la doctrina y el público cierto rechazo.

Creo que la línea dominante e incluso repetitiva es la que, reconociendo las posibilidades de la inteligencia artificial, se centra sobre todo en sus riesgos, identificados principalmente con los «sesgos». La solución que se propone generalmente es establecer muchos requisitos, sobre todo formales y procedimentales (publicidad, evaluaciones de riesgo), además de otros de carácter genérico cuya definición concreta debe establecerse caso por caso (línea seguida por el Reglamento de Inteligencia Artificial). En mi opinión, ese enfoque debe ser complementado para explicar bien los distintos usos y los diferentes riesgos, en lugar caer en una especie de nebulosa que lleve a prohibir algo sin conocer realmente de qué se trata, dando a todos los usos de la inteligencia artificial el mismo tratamiento. No debemos renunciar a una técnica útil e incluso imprescindible, sobre todo para organizaciones que, como las Administraciones Públicas, viven instaladas en la demora en el cumplimiento de sus funciones —con graves perjuicios a los ciudadanos— y en la carencia crónica de personal —dedicado en gran medida a tareas rutinarias—, y están sometidas, además, a un control muy poco efectivo —pese a las apariencias— sobre su actividad, debido a la falta de transparencia. La inteligencia artificial puede contribuir a paliar estos problemas.

Es preciso, además, evitar caer en el mismo marasmo que corroe al planeamiento urbanístico (donde la acumulación de trámites procedimentales ha dado lugar a frecuentes anulaciones y a retrasos que han acabado por inutilizar el instrumento) o en una regulación para iniciados en la que, como sucede en cierto modo con la protección de datos, la acumulación de exigencias genéricas parcialmente contradictorias haga que sólo al final de un procedimiento sancionador se pueda saber qué quería realmente decir (exigir, prohibir) la ley.

Por otro lado, hay que partir de que la inteligencia artificial es sólo un instrumento para ejercer derechos, poderes y facultades que son anteriores a ella y están sometidos a un marco jurídico que no cambia porque se utilice este instrumento: nada está prohibido ni permitido sólo porque se haga a través de la inteligencia artificial. Sólo desde esta constatación se puede decidir en qué medida es necesario, además de hacer cumplir ese marco jurídico general, aprobar, además, normas específicas para la utilización de inteligencia artificial. Por ello, voy a centrarme en examinar qué hace y qué puede hacer la inteligencia artificial, distinguiendo usos y viendo hasta qué punto suponen un cambio cualitativo en el Derecho administrativo.

2. DE LA DIGITALIZACIÓN A LA INTELIGENCIA ARTIFICIAL: UNA CLASIFICACIÓN

2.1. INTRODUCCIÓN

Como he venido haciendo desde 2020, parto de una clasificación conceptual, una taxonomía, de los distintos usos de la tecnología en el ámbito administrativo.

La utilización de las TIC (de la informática) no ha esperado a la regulación. Se ha ido introduciendo porque se entendía inicialmente que no tenía trascendencia jurídica, aunque no siempre sea así. Esa introducción sin una previa habilitación normativa es algo común, que se ha aplicado incluso a la inteligencia artificial (que se está regulando ahora, años después de su introducción en actividades públicas y privadas). Es cierto que no existe una reserva de ley general ni una «vinculación positiva» en materia tecnológica (aunque se establecía una exigencia de aprobación previa de los programas en el artículo 45.3 de la Ley 30/1992, seguramente no siempre cumplida). A la Administración no le está prohibido utilizar cualquier tecnología que no esté expresamente permitida. Entre otras cosas, porque tampoco la tecnología «analógica» estaba expresamente regulada. (Por ejemplo, la legislación no establecía una forma específica para realizar las notificaciones —artículo 59.1 de la Ley 30/1992 y 41.3 de la Ley 39/2015—, y los detalles de la regulación de la notificación en papel quedaban encomendados en parte a reglamentos de correos, a pesar de que ése sea sólo un medio de hacerla). Y ello a pesar de que los medios técnicos que se utilizan, con su particular estructura y forma, siempre ejercen alguna influencia sobre la actividad. Nadie se plantea, por ejemplo, si la Administración puede utilizar el buscador de Google para acceder a información, a pesar de que es notorio que puede influir en el tipo de información que se obtiene. La regulación se necesita y se obtiene cuando la utilización de tecnología afecta a la norma o cuando se pretende imponer obligaciones nuevas (presentación de escritos por vía electrónica, obligación de atender a las notificaciones electrónicas), como vamos a ver a continuación.

2.2. LA «ADMINISTRACIÓN ELECTRÓNICA»

La «administración electrónica» afecta especialmente a la comunicación entre los ciudadanos y la Administración, en ambas direcciones: presentación electrónica de escritos y documentos, notificaciones electrónicas.

Inicialmente, se presenta como un plus, una extensión al tráfico administrativo de las facilidades de la comunicación digital (¿por qué tener que comunicarse a través de una notificación en papel en la era del correo electrónico?), pero más tarde ese canal digital (más cómodo, pero que exige ciertas habilidades y medios que no todos los afectados tienen, y que está sometido a una característica fragilidad que requiere que exista un canal de refuerzo) tiende a convertirse en único y a sustituir (por ahorro de costes) al otro, con los consiguientes riesgos y perjuicios (brechas digitales, daños cuando se produce una avería o supuesto de denegación del servicio).

En la medida en que afecta a procesos que están regulados y que producen efectos jurídicos importantes (registros de entrada, notificaciones), es imprescindible una base jurídica, una regulación, máxime cuando el uso de estas técnicas se convierte en obligatorio (obligación de relacionarse electrónicamente con la Administración).

Es una digitalización que afecta, además de a las comunicaciones a/con los ciudadanos, al funcionamiento interno de la Administración, al *back office*: archivos electrónicos, comunicaciones electrónicas internas, actas electrónicas, etc.

Esa digitalización no afecta *directamente* al contenido de la actuación administrativa. Indirectamente, sí puede condicionarla en la medida en que genera más datos y puede facilitar el conocimiento de dicha actividad (útil para extraer conclusiones), la elaboración matemática y algorítmica y también el control (aunque este segundo aspecto se encuentra casi inédito).

2.3. LA PROGRAMACIÓN, ROBOTIZACIÓN O «AUTOMATIZACIÓN» DE LA ACTIVIDAD ADMINISTRATIVA

Cuando la tecnología interviene en la decisión administrativa, y no sólo en la forma de comunicarla, nos encontramos en primer lugar con lo que podemos denominar programación o robotización, donde la tecnología elabora el producto (a veces —no siempre— automatizando también el último tramo, la toma de la decisión), *pero sin influir sobre su contenido*. Hablamos (necesariamente) de actividad administrativa reglada y de aplicaciones informáticas en las que se vuelca esa programación normativa para incrementar la capacidad de resolver expedientes. Es una intervención informática que muchas veces resulta *invisible* y (jurídicamente) *irrelevante*.

La clave es que, cuando se trata de actuaciones individuales, para su control jurídico no se necesita conocer si se ha utilizado una aplicación informática y si la misma ha funcionado bien o mal, porque la legalidad de la decisión se deduce de su comparación con el marco normativo. Eso no significa que la aplicación informática no pueda producir sus propios problemas (básicamente, un error de programación conduce a la producción de decisiones ilegales). Y, por supuesto, si se automatiza el último paso, y la aplicación informática produce directamente el acto que se notifica al ciudadano, serán necesarias todas las garantías establecidas para la actividad administrativa automatizada. Pero no existe «autonomía» del sistema, que es un concepto utilizado por el Reglamento de inteligencia artificial.

La mera programación o robotización de la actividad administrativa opera en entornos de *actividad reglada*, en los que se puede crear un programa que recoge los parámetros normativos que guían a la Administración. El robot o la aplicación informática reproducen el hilo lógico que también debería seguir un operador humano y llegan al mismo resultado al que éste tendría que llegar. El programador no introduce ninguna innovación (salvo que cometa un error, cuya consecuencia será que se dicten decisiones sistemáticamente ilegales por no ajustarse al marco normativo). Esta categoría se ha discutido en —y a propósito de— las sentencias sobre el algoritmo «Bosco» (sobre las que hay referencias en la bibliografía). Con frecuencia, el verdadero problema está en la calificación de los hechos y, en

la medida en que aquí hay factores que requieren una decisión o una apreciación, vuelve a quedar en manos de un operador humano (así sucede, por ejemplo, en el caso de las decisiones de seguridad social). La pregunta en estos casos es si tiene sentido que, estando disponibles estas técnicas, un operador humano emplee su tiempo en preparar decisiones tan sencillas. *Parece casi obligatorio hacer uso de la tecnología para acelerar este tipo de decisiones, de forma que el operador humano se ocupe de vigilar el funcionamiento de la aplicación, facilitar el acceso de la misma a los ciudadanos o facilitar la comprensión de sus resultados*.

Otra cosa son las actuaciones complejas, en las que no podemos reproducir la aplicación de la norma (y controlar, por tanto, la decisión administrativa) sin comprobar el funcionamiento de la aplicación (el «algoritmo»). Allí donde la aplicación informática es uno de los elementos que contribuye a determinar el producto jurídico-administrativo, es necesario un acceso suficiente al mismo, para comprobar si se ajusta al marco normativo o no, si ha funcionado bien. Es algo que se desprende de la normativa vigente, porque forma parte del expediente administrativo. Para que ese control pueda producirse es imprescindible, naturalmente, saber que se ha utilizado un sistema de este tipo.

El test es sencillo: en la informatización o automatización «simple», la decisión administrativa no necesita citar la aplicación informática, ni siquiera como antecedente. En cambio, en los otros casos sí.

2.4. INTELIGENCIA ARTIFICIAL: MODELOS PREDICTIVOS (INCLUIDOS LOS FUNDACIONALES O GRANDES MODELOS DE LENGUAJE). UTILIZACIÓN PARA CONCRETAR CONCEPTOS INDETERMINADOS O PROBABILÍSTICOS (COMO EL RIESGO)

Sin necesidad de entrar en debates terminológicos (nos los ahorra el Reglamento), la inteligencia artificial supone un paso más frente a la mera programación o robotización de la actividad administrativa, que opera en entornos de *actividad reglada*, en los que se puede crear un programa que recoge los parámetros normativos que guían a la Administración. Las aplicaciones de inteligencia artificial influyen en el contenido de la decisión administrativa, que se mueve en un *entorno no reglado*, en el que existe un margen de apreciación.

El primer campo de uso la inteligencia artificial es el de las predicciones algorítmicas, muchas veces relacionadas con el perfilado. En definitiva, se trata de intentar responder a una pregunta analizando los datos del pasado (cómo sucedieron las mismas cosas de las que ahora intentamos saber cómo van a ocurrir), para encontrar en esos datos correlaciones (rasgos que se registraron con más frecuencia en aquellos casos en que sucedió lo que estamos intentando prever) y fabricar, de esa forma, una especie de retrato robot que podemos tomar como base para futuras decisiones administrativas. Es muy importante darse cuenta de que la utilización de estos modelos supone, a veces, sustituir una actuación administrativa *uniforme* por una actuación *personalizada* (así sucede, por ejemplo, con las «carpetas ciudadanas» que sirven para suministrar informaciones personalizadas a los ciudadanos, en función de lo que se supone que les puede interesar) y, en otros casos, sustituye a la toma de decisiones basadas en criterios subjetivos (como se hace cuando se inspecciona a quienes

parece que tienen mayor perfil de riesgo, en lugar de elegir los objetivos al azar o según la iniciativa del órgano competente). La inteligencia artificial promete eficiencia (al sector privado ya se la ha dado), en el sentido de que, en la actuación tradicional, se lanzan al viento muchas semillas y sólo germinan unas pocas, mientras que la inteligencia artificial permitiría identificar cuáles son las que van a germinar, de forma que sólo se planten esas. Por otro lado, también es importante retener que la inteligencia artificial suele ser aproximativa (ayuda a reducir la incertidumbre inicial, a pasar de 100 a 10), no dirigida a tomar decisiones definitivas y cerradas, para las que suele ser necesaria la intervención humana y la confrontación con datos reales, no sólo con predicciones.

Otro campo, relacionado con el anterior, porque también se basa en predicciones, es el de los «grandes modelos de lenguaje» (GPT, Gemini, etc.). Pueden tener (a qué negarlo) un uso instrumental en el acceso y organización de la información, como el de los buscadores. Pero no parece que se vaya a justificar jurídicamente una decisión administrativa diciendo que procede de uno de estos modelos.

2.5. LAS LIMITACIONES TECNOLÓGICAS DE LA INTELIGENCIA ARTIFICIAL («CAJAS NEGRAS») Y DE OTRAS TÉCNICAS Y SABERES UTILIZADOS POR LA ADMINISTRACIÓN

Es necesario ver qué papel tiene la inteligencia artificial en la aplicación del programa normativo que en cada caso vincula a la Administración. Por ejemplo, si uno de los requisitos necesarios para que la Administración pueda o deba tomar una determinada decisión es que se produzca una situación de «riesgo», que es un concepto predictivo y probabilístico, no parece haber ningún problema en que, en lugar o junto a otras técnicas que permitan verificar ese riesgo, se utilice un sistema de inteligencia artificial, un modelo predictivo basado en el análisis de datos pasados.

Salvo que la norma diga otra cosa, no existe una lista tasada de tecnologías que permitan concretar los conceptos legales. De hecho, la regla de libertad de medios de prueba que se establece en el artículo 77.1 LPAC («Los hechos relevantes para la decisión de un procedimiento podrán acreditarse por cualquier medio de prueba admisible en Derecho, cuya valoración se realizará de acuerdo con los criterios establecidos en la Ley 1/2000, de 7 de enero, de Enjuiciamiento Civil») cubre perfectamente la utilización de esta técnica, como de cualquier otra de las muchas que se usan en la práctica.

Cuando se enfatiza que los modelos predictivos son «cajas negras», en el sentido de que la complejidad de las «redes neuronales» o estructuras matemáticas que les sirven de base (el número de capas que contienen y de parámetros que se tienen en cuenta) impide recorrer de forma exhaustiva e integral el recorrido que lleva desde los datos de entrada hasta la respuesta proporcionada por el modelo, tal vez no tenemos en cuenta cuál es la realidad de otros medios técnicos que se utilizan en la aplicación del Derecho administrativo o en la actividad administrativa en general. El fenómeno de la «caja negra» hace referencia a una limitación tecnológica, a un grado de incertidumbre residual que no puede evitarse. Pero también (por poner varios ejemplos) los modelos de predicción meteorológica o demos-

cópica tienen márgenes de error. Del mismo modo, los sistemas de geolocalización tienen limitaciones en su grado de exactitud, a veces buscados deliberadamente por razones de privacidad o seguridad. Igualmente, los modelos físicos que predicen la trayectoria de proyectiles tienen márgenes de incertidumbre. De eso estamos hablando, no de algo que sea exclusivo de los sistemas de inteligencia artificial.

La consecuencia jurídica que, en mi opinión, debe extraerse de ese fenómeno de la «caja negra» es no admitir el uso de ese medio de prueba cuando ese margen de error pueda ser relevante para la decisión. Si utilizo un modelo predictivo para decidir si llevo un paraguas al salir de casa o al hacer el equipaje (como hacemos con frecuencia al mirar la predicción meteorológica en el teléfono móvil), el hecho de que exista un margen de error del 10% es irrelevante, porque lo peor que puede pasar es que me moje o que cargue con un paraguas en vano, y en la mayoría de los casos el sistema me da un consejo acertado. Del mismo modo, si utilizo un sistema similar para decidir a quiénes inspecciono, un margen de error similar puede ser asumible, porque lo peor que puede pasar es que se inspeccione a alguien que no ha cometido ninguna infracción (inspeccionar no es sancionar; sencillamente, no se descubrirá ninguna infracción y no se iniciará ningún procedimiento sancionador). En cambio, si se utilizara un sistema de inteligencia artificial para dictar resoluciones con efectos jurídicos sobre terceros, y el asunto pudiera ser dudoso y depender de ese margen de error, no se podría utilizar ese medio porque llevaría a conclusiones no fiables.

En la aplicación de las normas (y, concretamente, en la concreción de conceptos indeterminados que éstas utilizan) ya se emplean tecnologías que plantean problemas similares. Por ejemplo, en asuntos relativos a obras (ejecución de contratos de obras, responsabilidad patrimonial por accidentes, etc.) o en responsabilidad patrimonial en el ámbito sanitario, se utilizan con frecuencia, tanto en vía administrativa (y en la posterior contencioso-administrativa) como en la civil, técnicas y saberes que tienen parecidos márgenes de incertidumbre o cuyas conclusiones están sujetas a discusión en términos similares. Normalmente, esos conocimientos son incorporados a los procedimientos a través de informes periciales, frecuentemente varios que se oponen entre sí. Al final decide un lego (el juez, el titular del órgano administrativo) a partir de las explicaciones de los técnicos. En la práctica se produce una «tormenta perfecta» que hace que la decisión final (administrativa o judicial) se aleje bastante de una aproximación científica a la verdad. A lo que más se parece es al resultado de una batalla o de un enfrentamiento deportivo, que dependen de múltiples factores: la estrategia, la motivación de los contendientes o incluso el azar, y no sólo la fuerza o habilidad que teóricamente tienen los participantes. Con frecuencia, la convicción final de quien decide (órgano administrativo o judicial) está condicionada o determinada por las explicaciones que ofrecen los distintos peritos, enfrentados en un interrogatorio siempre limitado en el tiempo, lo que arroja varios problemas: desde la desigual elocuencia de los peritos (que a veces tiene más importancia que la solidez de sus tesis) al hecho de que pueden hacer afirmaciones carentes de fundamento, pero que pueden calar si no son adecuadamente desenmascaradas (lo que a veces es sencillamente imposible por la propia ordenación procesal de los tiempos y los turnos de intervención). Esta elemental constatación de la realidad hace que las habituales lamentaciones sobre los límites de los

modelos predictivos y de su explicabilidad pierdan, una vez puestas en contexto, buena parte de su fuerza.

2.6. LA UTILIZACIÓN DE TECNOLOGÍA PARA «ELABORAR» O PARA «JUSTIFICAR» UNA DECISIÓN PÚBLICA: CONSECUENCIAS JURÍDICAS

En todo caso, lo que me parece evidente es que, en la medida en que se utilice un modelo predictivo o cualquier otro sistema de inteligencia artificial para *justificar* una decisión que produzca efectos jurídicos, el control de dicha decisión (control que es imprescindible en un Estado de Derecho, en virtud, entre otras cosas, del derecho fundamental a la tutela judicial efectiva), incluirá un control acerca de si el sistema ha funcionado correctamente (es decir, si ha funcionado de acuerdo con sus objetivos) y de si su diseño es correcto, es decir, si los criterios que se le ha pedido que busque suponen una traslación o traducción adecuada del concepto jurídico que se está aplicando.

En definitiva: una cosa es utilizar un modelo informático para *elaborar* una decisión y otra cosa es utilizarlo para *justificarla*. El grado de visibilidad, relevancia jurídica y exigencias no puede ser el mismo en un caso y en otro. Sobre ello voy a insistir en los apartados siguientes.

2.7. INTELIGENCIA ARTIFICIAL Y ACTUACIONES ADMINISTRATIVAS PRELIMINARES O APROXIMATIVAS

Los sistemas de inteligencia artificial pueden utilizarse, como hemos visto, para concretar conceptos probabilísticos, por ejemplo, para demostrar que existe una situación de «riesgo» (cuando éste sea el presupuesto que permite adoptar una determinada decisión) o que se va a producir una reducción de la competencia en un sector (hecho que es relevante para autorizar, o no, una fusión). También puede utilizarse, por ejemplo, en la elaboración y justificación de un plan urbanístico o de una norma reglamentaria.

Pero puede suceder, y es lo más frecuente, que esas técnicas se utilicen para tomar otra clase de decisiones o para realizar otras actuaciones, y estoy pensando sobre todo en dos posibilidades. En primer lugar, actuaciones materiales, típicas en la actividad prestacional: asignar prioridades en la atención, incluir a personas en un concreto procedimiento, realizar un diagnóstico, etc. En segundo lugar, y ya en el campo de la actividad jurídica, actuaciones de carácter preparatorio o incluso provisional: distribuir los recursos materiales o humanos que van a dedicarse a esa actuación (refuerzos en uno u otro lugar) o identificar objetivos para la inspección.

En el campo de la actuación material o prestacional, y sin entrar ahora ni siquiera a resumir todo un sector del Derecho administrativo, hablamos de actuaciones que normalmente no están formalizadas en un procedimiento administrativo dirigido a la producción de un acto administrativo. Eso no quiere decir que carezcan de relevancia jurídica: pueden tenerla por la vía de la responsabilidad patrimonial e incluso pueden formalizarse o convertirse en contenciosas cuando el ciudadano presenta una reclamación. Algunas de estas

actuaciones en las que se utiliza la inteligencia artificial ni siquiera se refieren directamente a ciudadanos concretos, como, por ejemplo, cuando se utiliza para la asignación de recursos (en qué oficinas se va a situar más personal, por ejemplo, a partir de las necesidades predichas por un modelo informático). En el campo de los procedimientos administrativos, hablamos normalmente de actos de trámite, que, debido a que no producen efectos jurídicos directos sobre los ciudadanos y que son preparatorios, están sometidos a un control o programación normativos muy ligeros (no es una afirmación subjetiva, es lo que se deriva de que no se admita la impugnación autónoma de los actos de trámite no cualificados). Decisiones en las que la Administración tiene un margen de maniobra amplio, calificado o no de discrecionalidad (puesto que esta categoría suele reservarse para las resoluciones o actos finalizadores del procedimiento).

Es fácil estar de acuerdo en que la utilización de medios tecnológicos no puede suponer en absoluto un paso atrás en el control y en las garantías de la actuación administrativa. Pero, ¿deben establecerse más garantías o más controles, más desconfianza, en suma, sólo por el hecho de que se haya actuado «digitalmente»? Cuando la tecnología se utiliza para «justificar» la decisión, desde luego que es necesario someter a control esa tecnología (como ya se hace actualmente con cualquier otra), pero, ¿tiene sentido que una tecnología esté prohibida o se le exijan unos elevados requisitos sólo por el hecho de que puedan producirse en su aplicación errores que no tienen consecuencias cuando los comete un operador humano? Sería necesario, en todo caso, repensar los niveles de exigencia para ese tipo de decisiones en general, no sólo cuando en su elaboración se utiliza la IA.

2.8. ¿ES POSIBLE IR MÁS ALLÁ EN LA UTILIZACIÓN DE SISTEMAS DE INTELIGENCIA ARTIFICIAL?

¿Tiene sentido utilizar sistemas de inteligencia artificial para tomar decisiones públicas, más allá de estos dos supuestos analizados, es decir, la concreción de conceptos de tipo probabilístico y la preparación de decisiones preparatorias o aproximativas? En mi opinión, es muy discutible, porque, fuera de esos dos casos concretos, las decisiones públicas deben basarse, porque así lo impone su correspondiente programación normativa, en un determinado hilo argumental que no puede ser replicado por la inteligencia artificial salvo que admitamos la sustitución de ese hilo argumental por un equivalente o alternativo algorítmico que, en realidad, supone incumplir la programación normativa.

De nuevo acudo, para apoyar esta afirmación, a las normas atributivas de potestades (como se ha hecho previamente para admitir la utilización de programas que ayuden a elaborar decisiones regladas y para admitir el uso de inteligencia artificial para elaborar y justificar decisiones en las que se concreten conceptos normativos de tipo probabilístico, como el de riesgo, o se hará para exigir base normativa para la actuación automatizada).

Las normas atributivas de potestades (que forman esa programación normativa de la Administración) suelen establecer un supuesto de hecho y también la decisión o decisiones que la Administración puede tomar cuando el mismo se verifique. Partimos de la base de que el supuesto de hecho no es probabilístico (pues en ese caso nos encontraríamos en el

supuesto ya estudiado previamente), sino que se trata de un hecho que es necesario verificar. Resulta difícil aplicar aquí un sistema de inteligencia artificial, porque éste acredita probabilidades, no hechos. Todo lo más, puede aplicarse para probar o acreditar indicios, que, como es sabido, pueden servir como prueba en algunos casos. Probar hechos no es lo mismo que acreditar probabilidades. Por eso, lo normal es que la inteligencia artificial se utilice para aproximarse, para saber dónde y cuándo iniciar un procedimiento, pero que la decisión final se tome posteriormente sin inteligencia artificial, en función de hechos probados.

Si pasamos a la segunda parte de la norma habilitante de la potestad, es decir, la que especifica qué decisión puede o debe tomar la Administración ante la constatación del presupuesto de hecho, la Administración tiene que razonar por qué elige, de entre todos los posibles, un determinado contenido para la decisión. Si no hay margen decisional, se tratará de decisiones regladas, en las que tendrá cabida otro tipo de apoyo tecnológico ya analizado previamente. Pues bien, el razonamiento «valorativo», que es el que se lleva a cabo para justificar la elección, es totalmente ajeno al ámbito de la inteligencia artificial. Por ejemplo, una vez constatado el riesgo de contagio de COVID, una comunidad autónoma decide imponer restricciones a la actividad hostelera para reducir los contagios y otra opta, en cambio, por lo contrario (libertad de horarios y aforos) porque entiende que el riesgo es asumible, que también se van a producir contagios en otros ámbitos —aunque se cierre la hostelería— y que un parón de la actividad económica es igualmente indeseable y dañino. Son ejemplos reales, como todos sabemos. ¿Qué papel puede jugar aquí la inteligencia artificial? A mi juicio, ninguno, salvo el de verificador de hipótesis: cuantificar el riesgo de contagio, cuantificar los daños que sufriría la economía con una reducción de la actividad hostelera, etc. Pero, en la medida en que haya juicios de valor o preferencias entre intereses enfrentados entre sí, no veo cómo podrá sustituirse el juicio humano (la preferencia humana) por un juicio informático. Una vez que alguien decida (valorativamente) una opción (por ejemplo, minimizar el riesgo de contagio, al precio que sea), se podrá convertir esa decisión en una función matemática y utilizar tecnología (incluidos modelos de inteligencia artificial) para calcularla, pero hay un paso previo (valorativo, decisional) que es necesariamente humano, no ya por razones legales, sino puramente estructurales.

2.9. UTILIZACIÓN DE SISTEMAS DE INTELIGENCIA ARTIFICIAL GENERATIVA (GRANDES MODELOS DE LENGUAJE) PARA ELABORAR DECISIONES PÚBLICAS

A pesar de lo que se ha sostenido hace un momento, alguien podría decir que también se puede redactar una resolución judicial o administrativa con un modelo de lenguaje (LLM o *large language models*, tipo Gemini o GPT). Sin embargo, a mi modo de ver tampoco aquí se produciría ningún cambio, porque la decisión valdría en la medida en que detrás de ella hay un operador humano (que es el que ha decidido utilizar el modelo de inteligencia artificial) que hace suyo ese texto, y no porque éste tenga esa procedencia tecnológica.

Es importante distinguir entre el uso de la inteligencia artificial para «justificar» o motivar decisiones jurídicas, y su uso para elaborar decisiones. Aunque sea banal, hemos visto con frecuencia noticias que hablan de «demandas escritas con Chat-GPT» y lo mismo podemos decir de sentencias o incluso, recientemente, de leyes. ¿Qué decir en estos casos, además de subrayar un evidente uso propagandístico, por un lado, en favor de una determinada aplicación informática que está en el mercado y compite con otras, y, por otro lado, como forma de que un determinado operador jurídico (despacho de abogados) obtenga notoriedad, algo que en estos momentos se consigue, con frecuencia, incluyendo la expresión «inteligencia artificial» en cualquier noticia? Para empezar, conviene destacar las dudas sobre la fiabilidad de estos sistemas, que los convierten en una buena herramienta para acceder a información que ya se conoce, pero no tan buena como forma de investigar sobre un tema que no se conoce. Del mismo modo que antes hemos visto que los modelos de inteligencia artificial pueden ser útiles como aproximación, más que para tomar una decisión definitiva que afecte a personas concretas (o, por poner otro ejemplo, que un equipo deportivo muy económico puede ser más que suficiente para una excursión de senderismo, pero tal vez no para una expedición más exigente), cuando hablamos de tomar decisiones importantes, la información que proporcionan estos sistemas no resulta fiable (y, si es necesario comprobarla sistemáticamente, puede ser poco eficiente utilizarlos).

En todo caso, imaginemos que el titular de un órgano administrativo utiliza un «modelo de lenguaje» (recordemos que no existe sólo uno) para elaborar el texto de una resolución administrativa. Como las decisiones suelen basarse en informes técnicos, que constituyen su motivación, lo más probable sería que un técnico utilizara el modelo fundacional para preparar el informe. Para ello, haría la correspondiente «pregunta» o *prompt* al modelo, suministrándole unos parámetros o datos para que éste cree un texto. ¿Qué supone eso desde el punto de vista jurídico? En mi opinión, prácticamente nada.

En la situación que estoy planteando, el técnico o el titular del órgano administrativo «enfocan» el asunto, seleccionando un determinado marco (y unas fuentes) para su solución, y piden a un modelo de inteligencia artificial que redacte el texto. O bien se inclinan, en virtud de un juicio personal, por una u otra de las soluciones en presencia (legalizar una determinada construcción o bien considerarla ilegalizable), y encargan la redacción del texto a un sistema de inteligencia artificial, del mismo modo que, antes de la aparición de estos modelos, podía encargarse la redacción a una persona del equipo. ¿Podemos decir que «la decisión la ha tomado un modelo de inteligencia artificial»? Evidentemente, no, aunque de él proceda el texto.

Profundizando en el análisis, diríamos que hay tres elementos o fases. El primero es la toma de la decisión, que es un proceso puramente personal del titular del órgano (o, en muchos casos, del técnico que elabora el informe, cuyo criterio va a ser seguido por el órgano que toma la decisión). Ese proceso decisional tiene un margen de maniobra, de holgura, en el sentido de que se puede optar por la decisión A o por la decisión B, siendo ambas jurídicamente admisibles siempre que se justifiquen adecuadamente (algo que no es imposible ni difícil). En unos casos hablamos de potestad discrecional, otras veces de margen de apreciación y en otros casos ni siquiera se formaliza la calificación con tanta

precisión porque no se trata de resoluciones en sentido estricto, sino de actos de trámite o incluso de decisiones ajenas a un procedimiento administrativo, pero el hecho cierto es que existe ese margen y que la opción la toma el operador humano en un proceso subjetivo.

¿Qué factores son los que «pesan más» en ese proceso subjetivo? Es imposible saberlo: el intelecto humano es una auténtica «caja negra» mucho más difícil de descifrar que cualquier sistema de inteligencia artificial. Pueden ser decisivas consideraciones jurídicas (que después aparecerán en la motivación de la decisión): el operador puede sentirse vinculado (siguiendo el ejemplo de las restricciones por el COVID) por la protección de la salud, o, en sentido contrario, por la necesidad de limitar al máximo las restricciones a la actividad económica, pero también pueden entrar en juego muchos otros factores, incluidos los de tipo político (favorecer a unos determinados responsables políticos o erosionarlos), ideológico, de carrera, etc. Por supuesto que en algunos casos podríamos hablar de auténtica desviación de poder, pero es casi imposible demostrarla ni siquiera indiciariamente. Y también es cierto que algunas técnicas organizativas (de selección de personal) pueden, al incidir sobre el sistema de incentivos, favorecer o perjudicar que las decisiones se tomen por razones más «alineadas» con los intereses públicos, pero lo cierto es que, en mayor o menor intensidad, la decisión se toma al margen de su justificación, que es un elemento posterior. En todo caso, lo que me parece clave es que se pueden justificar jurídicamente distintas decisiones (no sólo una), de modo que la toma de la decisión no se explica por razones jurídicas (que las hay para justificar distintas decisiones), sino por otras razones personales previas. En el mejor de los casos, esas razones coinciden con las que se utilizarán en la motivación (eso indica que el operador humano «se cree» esas razones jurídicas y no incurre en desviación de poder), pero hay algo previo que le ha llevado a optar por esas razones y no por otras (que le habrían permitido llegar a una decisión diferente, igualmente susceptible de ser motivada), tal como explicó, en nuestra doctrina de Derecho administrativo, el recordado Alejandro Nieto.

La segunda fase es la redacción y justificación de la decisión, que es un proceso condicionado o determinado por la decisión misma, en el sentido de que cada una de las posibles decisiones se justificará en función de un argumento distinto. Si se opta (por continuar con la analogía de las medidas de prevención de contagios COVID) por el cierre de establecimientos, se hará hincapié en la protección de la salud y el riesgo de contagio, mientras que, si se opta por reducir al máximo las restricciones, se insistirá en las dudas sobre el riesgo real de contagio, en la posibilidad de que también se produzcan contagios por otras vías y en la importancia de la responsabilidad individual y la libertad de empresa. Es, como digo, una labor condicionada por la anterior. Aunque, en un análisis puramente lógico, pueda parecer lo contrario (aparentemente, las premisas jurídicas conducen a una decisión concreta), la realidad es que es la decisión previamente tomada la que obliga a seguir un determinado hilo argumental para su justificación. Esta segunda fase no es creativa, puesto que lo relevante es la toma de la decisión, no tanto su motivación. Por lo tanto, el hecho de que la redacción se realice se con la ayuda de un modelo de IA no supondría un cambio cualitativo con consecuencias jurídicas importantes: no más, en todo caso, que el hecho de

encomendar a una persona del equipo la redacción del texto de la decisión, cosa que es frecuente en la forma de trabajo en la Administración e incluso en algunos órganos judiciales.

El tercer elemento es el del control, desarrollado en muy distintas instancias, incluidas (pero no sólo) las de tipo judicial. Ese control consiste, fundamentalmente, en verificar si la decisión se ajusta al marco jurídico a que se encuentra sometida, a la programación normativa de la misma, a fin de llegar a una conclusión positiva (que lleva a admitir su validez) o negativa (cuya consecuencia puede ser, en algunos casos, la anulación o revocación de la decisión). En cierto modo, lo que se verifica es la motivación de la decisión, que es, a su vez, el conjunto de argumentos utilizados por el operador humano que dicta la decisión para demostrar que la misma se ajusta a su marco jurídico. El control supone una especie de revisión de la motivación (una motivación a la inversa), o una forma de «remontar» la motivación para comprobar si la decisión se ajusta al marco jurídico.

En ese esfuerzo argumentativo, que va de la decisión al marco jurídico para ver si la primera se ajusta al segundo, el hecho de que la decisión se haya elaborado con un modelo de IA es irrelevante, no se tiene en cuenta. Esa utilización de un modelo fundacional no opera en ninguna de dos dimensiones posibles en que podría ser relevante para la justificación de la decisión. No se ha utilizado el modelo para «traducir» un concepto normativo de tipo probabilístico (es lo que se estudió en el epígrafe anterior, así como sus consecuencias). Y, por otro lado, el hecho de que el texto se haya elaborado con la ayuda de un modelo informático es indiferente, porque lo relevante para la verificación de la legalidad de la decisión (que es el núcleo de la función de control) es lo que ella dice y cómo se justifica jurídicamente, pero no cómo se haya elaborado la misma.

Cabría objetar y defender una tesis alternativa. Para la validez de una decisión administrativa, no importa sólo el «qué» (la parte dispositiva) y el por qué (la justificación o motivación), sino también el cómo. Esto se manifiesta jurídicamente en instituciones como las causas de abstención y recusación, las normas sobre formación de voluntad de los órganos colegiados, etc. Precisamente porque existe un elemento decisional, es decir, porque la decisión no es una «función» del ordenamiento jurídico (algo que se derive «matemáticamente» de la norma y de los hechos), y porque se pueden justificar jurídicamente distintas decisiones, es importante que ese proceso decisional, que es anterior al Derecho, no esté viciado y se realice de la forma legalmente prevista. La inserción en ese proceso de un modelo de inteligencia artificial lo viciaría y afectaría a la legalidad y validez de la decisión adoptada.

Sin embargo, creo que este análisis no sería acertado, porque el sistema de inteligencia artificial opera *después* de que se haya tomado la decisión, y, por tanto, para labores instrumentales. Por decirlo en términos de análisis lingüístico, lo que importa de una decisión administrativa no son sus palabras (que pueden ser perfectamente predecibles), sino el «metadato» de que el órgano competente haya optado por esa decisión y no por ninguna otra alternativa.

Un tribunal puede tener, para un tipo de pleito más o menos predecible (pensemos en una acción reivindicatoria, o en una petición de reequilibrio de una concesión por ries-

go imprevisible, o, por mencionar un ejemplo absolutamente real, conceder o denegar la suspensión cautelar del acto administrativo impugnado en un proceso contencioso-administrativo), dos plantillas, una para la decisión estimatoria y otra para la desestimatoria. El contenido de las sentencias es totalmente previsible (muchas veces, salvo el párrafo en que se justifica la subsunción del caso concreto en el marco elegido) y la labor de redacción no existe, puesto que se apoya en la correspondiente plantilla. Pero ello no resta un ápice de importancia a la sentencia como decisión, puesto que lo importante no son las palabras de la resolución, sino la decisión de optar por una u otra plantilla.

Por tanto, el uso de modelos de lenguaje (IA generativa) como elemento auxiliar para la redacción del texto de decisiones que producen efectos jurídicos (incluidos, por ejemplo, informes que indirectamente producen efectos jurídicos por su influencia sobre la resolución) carece de relevancia jurídica.

Es cierto que el instrumento utilizado puede tener su importancia, en el sentido de que por esta vía se produzca un texto que no sea igual al que saldría «de la pluma» de un operador humano, pero son efectos que se producen en cualquier contexto y que se consideran despreciables, de modo que no me parece justificado que se les dé importancia sólo porque en este caso el instrumento sea un modelo fundacional. También influye, lógicamente, que sea una u otra persona la que redacte materialmente el texto, y no por eso se exige, como condición de validez de las decisiones administrativas y judiciales, que sólo hayan intervenido en su redacción el o los titulares del órgano. Por no hablar de otros instrumentos que se utilizan sin ningún problema o prevención y que también influyen en el contenido de la decisión, e incluso de forma más incisiva. Por ejemplo, el uso de buscadores para el acceso a la información (siendo el de Google el más difundido, con una enorme diferencia sobre los demás) puede influir en la información a la que se accede, y no por ello se prohíbe su uso ni se exige su exhaustiva documentación. Incluso el hecho de que *no se utilice un buscador*, y se limite la información a la que figura en una determinada biblioteca, también influye, pero se considera algo despreciable o irrelevante (aunque pueda no serlo, como saben, por ejemplo, las empresas que invierten en realizar actos de difusión de información favorable a sus tesis, puesto que saben que esa labor de «concienciación» es relevante).

Sería posible, teóricamente, que el operador no utilizase un modelo fundacional para *redactar* la decisión que ha tomado (elaborando lingüísticamente el material que se le suministra), sino para *tomar* la decisión. El sistema arrojaría, en principio, la decisión más previsible a la vista de los datos con los que se ha entrenado, es decir, en principio, la decisión más parecida a las que suelen tomarse en esos casos. De nuevo se trata de algo que puede ser comprendido con categorías conocidas o que se parece a fenómenos que ya conocemos. Por ir a un ejemplo extremo, en el sistema actual es posible (y no siempre descartable) que el operador humano, cuando puede —de hecho— elegir entre dictar una resolución estimatoria o desestimatoria de la solicitud (o de la demanda) y justificar cualquiera de ellas sin incurrir en prevaricación, tome su decisión en función de circunstancias aleatorias, desde lanzar una moneda al aire a hipótesis mucho más probables, como la mayor o menor simpatía o arrogancia mostradas por el abogado o representante del interesado, la respuesta mediática que una u otra decisión pueden provocar o sus efectos sociales o económicos

(por ejemplo, la mayor o menor simpatía por la instalación de parques eólicos o fotovoltaicos en el medio rural puede llevar a interpretar en un determinado sentido la normativa aplicable, aunque no sea un factor relevante en el marco jurídico de la decisión que en ese caso deba tomarse). ¿Invalidan la decisión adoptada esas consideraciones que no se exteriorizan y que no influyen en su justificación jurídica? Sólo en aquellos casos extremos en que puedan acreditarse y nos encontremos en desviación de poder, lo que no sucede casi nunca. Lo mismo deberíamos decir en estos casos.

También cabría una objeción más fundamental, que tal vez exigiría un epígrafe independiente. Los modelos de IA generativa no «piensan», sino que arrojan un resultado (por ejemplo, un texto) que es el que el modelo «sugiere» a la vista de la pregunta formulada y a partir de los datos que se le hayan suministrado para su entrenamiento. Algo parecido (aunque con un nivel muy superior de complejidad, obviamente) a lo que hace un teléfono móvil cuando, mientras el usuario está escribiendo un mensaje, le «sugiere» cuál podría ser la próxima palabra. El operador humano no actúa de la misma manera, puesto que antes de escribir «decide» qué quiere hacer y después lo escribe, sin limitarse a reproducir patrones. En este sentido, surge la objeción de que, en realidad, también los operadores humanos, cuando aparentemente deciden, no hacen otra cosa que arrojar el resultado previsible a partir de una serie de condicionantes, como su ideología, formación académica, origen social, expectativas e intereses, etc. Es la objeción determinista, permanentemente actualizada. Es una tesis defendible en términos filosóficos, pero creo que se puede afirmar que el ordenamiento jurídico no la comparte. Si realmente los comportamientos humanos estuvieran tan influidos por factores previos, habría que cambiar reglas básicas del ordenamiento, que distinguen entre comportamientos «libres» y comportamientos obligados (por ejemplo, el consentimiento prestado a una cláusula abusiva es irrelevante, la cláusula no obliga, mientras que sí obligan otras cláusulas a las que se entiende que se ha prestado un consentimiento libre) y que hacen a las personas responsables de aquellos de sus comportamientos considerados libres (por ejemplo, se obliga a cumplir los contratos).

La clave, en mi opinión, se encuentra en que, aunque la decisión que tome un operador humano pueda ser la misma que se tome si se le hace la pregunta a un modelo de IA (a fin de cuentas, una sentencia estimatoria tiene el mismo texto tanto si es el resultado de una sesuda y honrada reflexión como si se ha escrito con una pistola apuntando a la cabeza de quien la redacta y la firma), su significado es diferente. El operador humano introduce un elemento decisional o volitivo que está ausente del modelo fundacional y también puede evolucionar o cambiar o de opinión, o reaccionar ante nuevas circunstancias, mientras que el modelo fundacional depende de los datos con los que se le haya entrenado, que pueden estar actualizados o no.

2.10. EL USO «REACTIVO» DE LA INTELIGENCIA ARTIFICIAL PARA CONTROLAR, Y NO PARA TOMAR, DECISIONES PÚBLICAS

Normalmente, se piensa en la inteligencia artificial (o sistemas tecnológicos más sencillos, como los que soportan la automatización) como una ayuda —o incluso una susti-

tución— en la toma o adopción de decisiones. En esa clave se analizan los riesgos de que por esa vía se produzcan decisiones que discriminen a determinadas personas o les causen perjuicios por otra vía.

Con ser importante esta posibilidad, también lo es, en mi opinión, otro uso de la inteligencia artificial, probablemente menos peligroso y de gran potencialidad. Se trata de la posibilidad de utilizar el análisis de datos masivos para saber realmente cómo está funcionando la Administración (o la jurisdicción), cómo se está cumpliendo (o incumpliendo) el ordenamiento jurídico y qué efectos produce una determinada norma.

El tipo de control jurídico de las decisiones administrativas al que estamos acostumbrados, y que se tiende a considerar suficiente, resulta, con frecuencia, extremadamente defectuoso y debería ser superado con la ayuda de técnicas de datos masivos. Cuando hablamos de decisiones en las que existe un margen de apreciación (y vuelvo a repetir, como lo he hecho varias veces en este trabajo, que eso sucede casi siempre y que, cuando no sucede, ya estamos tardando demasiado en implantar sistemas informáticos que eviten retrasos y errores y permitan dedicar recursos humanos a tareas más importantes), los procedimientos de recurso consisten, casi exclusivamente, en comparar la decisión adoptada con el marco jurídico y ver si se ha motivado correctamente que esa decisión es compatible con el mismo. Difícilmente se puede comprobar en el seno de esos procedimientos de control cómo se ha aplicado el marco jurídico a otros casos similares. Pues bien, en ausencia de ese análisis comparativo, el procedimiento de control es muy insuficiente y se queda por debajo de lo que, en mi opinión, debería garantizar el derecho fundamental a la tutela judicial efectiva. De poco sirve, por ejemplo, decir que un determinado acto es conforme a derecho porque en él se aplica correctamente el baremo legalmente aprobado (para la resolución de un procedimiento selectivo, o para el reconocimiento de un título, por ejemplo, o para la concesión de una ayuda a un proyecto de investigación, o a un contrato predoctoral), si en otros casos ese mismo baremo se aplica de una forma diferente (más generosa, o más estricta), pero también considerada como una interpretación admisible del mismo. El procedimiento de control es ciego ante esas desviaciones, lo que supone una deficiencia muy importante. No estoy diciendo que un órgano (en una determinada composición) esté vinculado por lo que otros miembros anteriores de la misma comisión hayan decidido, o que sea imperativo que se aplique siempre la misma interpretación, pero al menos se debería ser consciente de esas desviaciones y valorar hasta qué punto son admisibles. Seguramente sería necesario diseñar mejor estos procedimientos para establecer mecanismos de unificación de criterios, también de forma preventiva y no sólo a través de mecanismos individuales de impugnación.

En esta labor, que me parece imprescindible, es necesaria la *transparencia* y seguramente también el apoyo tecnológico. A través de la anonimización se pueden publicar las decisiones de manera generalizada y hacer posible su análisis, que normalmente requerirá la utilización de procedimientos de inteligencia artificial que permitan, por ejemplo, detectar patrones de comportamiento y sesgos. Dicho de otra forma, detectar las razones ocultas que llevan a que el margen de apreciación de que dispone quien toma la decisión se utilice en unos casos de una forma y en otros de otra, las razones realmente relevantes, que son

previas al despliegue de la motivación exigida en cada caso, que es la única que normalmente analizan los procedimientos de recurso.

Al margen de los procesos de impugnación o revisión, la tecnología puede ayudar a comprobar cómo se está aplicando una norma. Por ejemplo, en el ejercicio de la potestad sancionadora es importante saber cómo se está aplicando la horquilla legalmente prevista para las sanciones impuestas a un determinado tipo de infracciones graves, una horquilla que muchas veces es amplísima (puesto que su extremo superior es, con frecuencia, 10 veces mayor que su extremo inferior, y a veces el margen de divergencia es aún más alto, como ocurre con las infracciones muy graves) y para la que se establecen criterios de graduación que con frecuencia son muy genéricos y se pueden concretar de maneras muy diversas (basta remitirse al artículo 29.3 LRJSP). Puede ocurrir que, en entornos descentralizados, esos márgenes se apliquen de manera muy diferente, sin que ello aflore en los procedimientos de revisión o impugnación porque todas esas decisiones sancionadoras son, tomadas individualmente, conformes a Derecho (es decir, se encuentran dentro de la horquilla normativa y son justificables con arreglo a los criterios de graduación) y porque a los interesados que interpongan esos recursos les falta información sobre la aplicación general de las normas, más allá de su expediente concreto. Como digo, para conseguir este objetivo —que me parece claramente deseable— hacen falta obligaciones intensas de transparencia y también, seguramente, la aplicación de tecnología, que tendría también como derivada la posibilidad de realizar una auténtica justicia predictiva.

Por otro lado, el artículo 130.1 LRJSP ha convertido en obligación legal la deseable «evaluación normativa», estableciendo, en su apartado 1, que «[l]as Administraciones Públicas revisarán periódicamente su normativa vigente para adaptarla a los principios de buena regulación y para comprobar la medida en que las normas en vigor han conseguido los objetivos previstos y si estaba justificado y correctamente cuantificado el coste y las cargas impuestas en ellas». A mi juicio, esta evaluación es una auténtica asignatura pendiente, tanto más importante en la medida en que la aprobación de normas se ha convertido en una finalidad en sí misma y en una forma relativamente fácil de justificar la gestión política. El cambio normativo, no pocas veces puramente lingüístico o cosmético (problema acentuado por la distribución de competencias, que lleva a las comunidades autónomas, pero también al Estado, a querer legislar a cualquier precio incluso en materias en las que apenas tienen competencias), conduce a la aprobación de normas que nunca se evalúan. Para ello debe servir también la tecnología, analizando cantidades ingentes de datos. De hecho, las normas jurídicas —y los propios sistemas de inteligencia artificial— son (o deben ser) objeto de procesos de *auditoría inversa*, en los que se verifique si cumplen determinados principios o requerimientos (en este caso, los explicitados en el artículo 129 LRJSP), no mediante un análisis racional de su contenido, sino mediante un estudio empírico de su funcionamiento, buscando correlaciones más que causas (o, mejor dicho, estableciendo las causas a partir de las correlaciones observadas, y no de forma teórica o apriorística).

Este análisis, como todos, está sujeto a discusión y a posibles errores y debe llevarse a cabo cumpliendo premisas metodológicas bien establecidas. Ejemplos recientes (como el análisis de los efectos económicos de la legislación que limita la renta en el arrendamiento

de viviendas, o de las normas que elevan el salario mínimo interprofesional) muestran que también aquí las conclusiones son discutidas y con frecuencia sesgadas ideológicamente, pero no parece que estemos ante una técnica «de riesgo» cuya utilización deba limitarse, sino ante una forma difícilmente sustituible de evaluar las políticas públicas.

3. CONSECUENCIAS JURÍDICAS: ¿EVOLUCIÓN O REVOLUCIÓN?

3.1. LA ACTUACIÓN ADMINISTRATIVA AUTOMATIZADA COMO CONCEPTO JURÍDICO: ¿EL CONTENEDOR ADECUADO PARA LA INTELIGENCIA ARTIFICIAL?

Antes de que se regulase la inteligencia artificial, se reguló la actividad administrativa automatizada, entendida como aquella que se produce sin intervención humana (o, al menos, sin intervención humana en la aplicación del programa, aunque sí en su creación). Tanto la legislación nacional (en España, el artículo 41 LRJSP; en Alemania, el parágrafo 35a de la Ley Federal de Procedimiento Administrativo; existe también regulación autonómica) como —para las decisiones automatizadas en general— el RGPD (artículo 22) establecen un régimen para este caso. El artículo 22 del RGPD se ha convertido, de hecho, en la norma imperativa más relevante en el abordaje de la actividad automatizada y, por extensión, de la utilización práctica de la inteligencia artificial. En la práctica, viene a condicionar la adopción de decisiones que estén basadas únicamente en el tratamiento automatizado de datos y que produzcan efectos jurídicos sobre el interesado o le afecten significativamente de modo similar, a que estén autorizadas por Derecho de la Unión o de los Estados miembros y a que esa habilitación normativa contenga medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado. Se trata de una regulación transversal, que afecta tanto a decisiones públicas como privadas. El artículo 22 (con el precedente del artículo 15 de la Directiva 95/46) viene a considerar que el hecho de que el tratamiento automatizado de los datos dé lugar por sí solo (sin intervención humana, o sin intervención humana significativa) a una decisión, constituye en sí mismo una inmisión en la libertad del ciudadano, adicional a la que deriva del propio tratamiento, por lo que está sometido a requisitos específicos, incluida la exigencia de una habilitación normativa, que deberá establecer garantías, entre las que deberá incluirse «como mínimo el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión».

No hay espacio para una exposición, ni siquiera resumida, de los problemas que plantean las actuaciones (en este caso, administrativas) automatizadas o para un comentario al artículo 22. Sí puede decirse, en mi opinión, que, en términos generales y al margen de una regulación concreta, no se puede discutir que la automatización exige una regulación y una habilitación. La atribución a un determinado órgano de la competencia para dictar un acto (que es algo que existe en todos los casos como habilitación normativa a la Administración) supone, salvo que se indique lo contrario, que la competencia se ejercerá

de forma personal o humana por ese órgano. Para que eso cambie, o para que se entienda que un acto que en realidad deriva de un sistema informático es imputable o procede de un órgano administrativo, es necesario que una norma lo autorice y también es necesario que se establezcan garantías, incluida la cadena de legitimación y también algún tipo de garantía frente a posibles errores.

En la práctica, se está ensanchando el ámbito de aplicación del artículo 22, a través de una interpretación amplia tanto de lo que se entiende por decisión que produce efectos jurídicos, como de cuándo se entiende que la decisión está «basada únicamente en el tratamiento automatizado». Se tiende a aplicar este artículo a todo supuesto de empleo de inteligencia artificial (aunque no suponga automatización) y también a los casos de influencia «casi decisiva» del sistema sobre el operador humano (sesgo de automatización), aunque tampoco haya, en sentido estricto, automatización. Un hito relevante es la sentencia del TJUE en el asunto SCHUFA (sentencia de 7 de diciembre de 2023, asunto C-634/21), que, en el fallo, establece que el artículo 22 del RGPD «debe interpretarse en el sentido de que la generación automatizada, por una agencia de información comercial, de un valor de probabilidad a partir de datos personales relativos a una persona y acerca de la capacidad de esta para hacer frente a compromisos de pago en el futuro constituye una «decisión individual automatizada», en el sentido de la mencionada disposición, cuando de ese valor de probabilidad dependa de manera determinante que un tercero, al que se comunica dicho valor, establezca, ejecute o ponga fin a una relación contractual con esa persona». En los párrafos 42-68 se analiza muy detenidamente la interpretación de los requisitos del artículo 22, destacando que «existiría un riesgo de elusión del artículo 22 del RGPD y, en consecuencia, una laguna en la protección jurídica, si se optara por una interpretación restrictiva de esta disposición según la cual la generación del valor de probabilidad debiera considerarse un mero acto preparatorio y solo el acto adoptado por el tercero pudiera, en su caso, calificarse de «decisión», en el sentido del artículo 22, apartado 1, de dicho Reglamento» (párrafo 61). El Tribunal entiende que no es necesario que la predicción automatizada sea la decisión final del banco (es decir, no sería necesario que estemos ante una «decisión automatizada» en sentido estricto), siendo suficiente que ejerza una influencia determinante sobre la decisión. Ésta fue, asimismo, la clave de la sentencia de 5 de febrero de 2020 del Tribunal de Distrito de La Haya (ECLI:NL:RBDHA:2020:1878) en el conocido «caso Syri», puesto que el Tribunal entendió que un «informe de riesgo» sobre una persona, que indica que existen factores que hacen sospechar que puede haber cometido un fraude, y que puede ser utilizado para iniciar una inspección, es análoga a una decisión que produce efectos jurídicos sobre esa persona.

Posiblemente, la automatización, con ser importante, no lo sea tanto como otras circunstancias (por ejemplo, que se utilice, o no, inteligencia artificial para determinar el contenido de la decisión administrativa). La extensión del artículo 22 puede estar llevando a que se traten de forma igual fenómenos que son diferentes. Habrá que ver en qué medida la aplicación del Reglamento de Inteligencia Artificial cubre necesidades que ahora se abordan, a falta de otra norma, por el artículo 22.

3.2. ¿CUÁLES SON REALMENTE LOS RIESGOS DE LA INTELIGENCIA ARTIFICIAL?

La literatura jurídica (y no jurídica) sobre inteligencia artificial se refiere, de manera predominante, a sus riesgos, que también son el eje del Reglamento de Inteligencia Artificial, cuyo enfoque regulatorio es precisamente éste, es decir, que se trata de una tecnología que produce riesgos y por ello debe estar sometida a una regulación preventiva. Por esta abundancia de análisis, sólo voy a hacer algunas precisiones.

A diferencia de lo que sucede con otras tecnologías de riesgo, en las que están muy claros los daños que pueden producir (pensemos en el uso de automóviles, maquinaria en general, energía nuclear, etc.), que muchas veces son de tipo corporal, en el caso de la inteligencia artificial no son tan visibles las lesiones jurídicas que puede producir, salvo la discriminación (que es un riesgo evidente). Por ello, creo que es necesario teorizar el ámbito de los derechos fundamentales y las lesiones concretas que puede provocar la inteligencia artificial, más allá de la discriminación. Un ejemplo de estos daños específicos que pueden producirse con la inteligencia artificial lo tenemos en el sector de la competencia, en el que se puede utilizar la inteligencia artificial para realizar conductas colusorias de forma novedosa, sin necesidad de un acuerdo entre las partes.

No niego esos peligros, pero me parece necesario profundizar en ellos más allá de las etiquetas. No siempre queda claro si el riesgo son las predicciones, las «malas» predicciones o su utilización por operadores humanos (que no es lo mismo). Por ejemplo, en el caso de la discriminación algorítmica, una de las cuestiones clave es si, una vez detectado por un sistema de inteligencia artificial que existen correlaciones basadas en factores que están prohibidos por el artículo 14 CE (por ejemplo, que personas de una determinada procedencia geográfica han tenido un desempeño peor en determinada actividad, lo que podría llevar, «estadísticamente», a colocarlos en peor posición para acceder a un empleo en ese sector), se debe imponer una especie de «ceguera» ante dichos factores, impidiendo, directamente, que se tengan en cuenta en este tipo de predicciones. Parece claro que sí.

Al margen de esa discriminación «estructural», que debe evitarse incluso cuando pueda estar estadísticamente «fundamentada», el peligro de la inteligencia artificial es que suministre predicciones equivocadas y que ello provoque consecuencias negativas para las personas a las que se «castiga» como si hubieran hecho algo que no han realizado, pero que un sistema de inteligencia artificial les imputa. De hecho, el denominador común de la regulación del Reglamento de Inteligencia Artificial es evitar esas decisiones erróneas (de ahí la exigencia de que los datos sean adecuados, que se garantice la ciberseguridad, que se investiguen desde el principio las posibles fuentes de errores, etc.). Todo esto es importante, pero no deja de ser un enfoque parcial, puesto que las predicciones equivocadas sólo serán problemáticas *si admitimos que a las personas se las puede tratar de un modo u otro (favorable o desfavorable) en función de «predicciones» y no en función de hechos probados*. Y esto, en la mayoría de los casos, sencillamente no es posible, ni con inteligencia artificial ni sin ella (me remito a las consideraciones previas sobre los usos de la inteligencia artificial).

Otro de los peligros de la inteligencia artificial (colocado en un plano muy diferente) es la sustitución de algunos trabajadores, lo que da lugar a lo que podemos llamar objeción «ludita». Parece evidente que la inteligencia artificial, como cualquier otra innovación tecnológica, produce ganadores y perdedores, desde luego a corto plazo, pero no sólo. Al igual que la globalización, tiene consecuencias económicas muy serias, aunque creo que a corto plazo seguramente serán menos intensas que las que ya ha producido en las últimas décadas la globalización. Estas consecuencias operan en varias direcciones y suscitan distintas reflexiones.

El hecho de que la inteligencia artificial pueda encargarse de tareas rutinarias, descargando de ellas a ciertos trabajadores (o, en el peor de los casos, haciendo redundantes a los trabajadores que actualmente se encargan sólo de ellas), tiene más ventajas que inconvenientes, en mi opinión, puesto que muchas organizaciones (sobre todo, las públicas) tienen una endémica carencia de personal y, por ello, carece de sentido que se insista en acumular retrasos y demoras en lugar de aprovechar avances que podrían aumentar la eficiencia y mejorar el servicio. Con todo, esa sustitución es más aparente que real y, en todo caso, es sólo parcial. Lo esperable es la convivencia de operadores humanos y sistemas de inteligencia artificial, no una sustitución total, sobre todo porque estos sistemas pueden cometer errores y, tanto por ello como por las dificultades que muchas personas pueden tener para interactuar con ellos, sigue siendo necesaria una asistencia humana a las personas sobre las que se ejerce esa gestión asistida por inteligencia artificial. Además, también hay un problema de seguridad: la inteligencia artificial no tiene la misma fiabilidad que el operador humano en ciertas tareas, por lo que requiere, al menos, una revisión (pensemos en los errores que cometen los modelos fundacionales cuando se les pide que elaboren un informe o un texto sobre un tema concreto).

En todo caso, la resistencia «ludita» no ha frenado hasta la fecha ninguna tecnología (sí lo han hecho, en algunos casos, las objeciones éticas, como ha pasado hasta cierto punto con la biotecnología). En caso de que la sustitución de mano de obra fuera realmente masiva, de modo que una parte importante de la población no pudiera trabajar, entraríamos de verdad en un terreno casi desconocido en la historia, en el que sería necesario realizar reformas y cambios muy intensos, pero no inimaginables, si se quieren evitar tensiones insuperables.

No muy lejos de esa sustitución se encuentran los problemas y conflictos que genera la inteligencia artificial, no ya para los trabajadores, sino para otras formas de capital, como es la propiedad intelectual o los medios de comunicación. No es casual que en este momento se encuentren planteadas varias demandas colectivas que actúan como grandes escenarios de tales conflictos: la demanda de un grupo de escritores (como Jonathan Franzen) contra las grandes tecnológicas por violación de sus derechos de autor (en la misma dirección que la huelga de guionistas de Hollywood, que reaccionaba contra los perjuicios derivados del uso de inteligencia artificial), la demanda (en España) de un grupo de medios de comunicación escritos contra Meta (matriz de Facebook, Instagram y Whatsapp) por el daño que supuestamente les causa la publicidad personalizada (que no es un asunto lateral o secundario, sino una de las aplicaciones prototípicas y más desarrolladas de la inteligencia

artificial), la demanda del New York Times contra OpenAI (por haber, supuestamente, utilizado textos protegidos por derecho de autor para entrenar su modelo GPT) o la más reciente demanda de grandes periódicos europeos contra Google. No es este el momento para estudiar estas cuestiones, pero sí para destacar que aquí se marcarán, seguramente, algunas líneas muy relevantes para el futuro, estableciendo el punto de equilibrio entre intereses enfrentados, como se hizo en el pasado, por ejemplo, en algunas acciones judiciales muy relevantes de derecho de daños, como las que enfrentaron a las autoridades sanitarias y a las tabaqueras. En todo caso, se pueden apuntar algunos de esos peligros de la inteligencia artificial que necesitan una construcción jurídica.

Por otro lado, y conectando la publicidad personalizada con la desinformación, hasta qué punto debe considerarse práctica prohibida el sometimiento de las personas a unos impactos o mensajes (no sólo de publicidad comercial o política, también de información) que estén pensados específicamente para ese perfil de personas sin que ellas tengan una conciencia clara de ello. ¿Son suficientes las actuales peticiones de consentimiento que se hacen al abrir una página web, para aceptar que el ciudadano ha consentido libremente someterse a la experiencia de verse e informarse en un entorno digital que está pensado específicamente para él (no es el mismo que ven otros ciudadanos) y que, por ello, puede «confundirle» con cierta facilidad?

3.3. LA INTELIGENCIA ARTIFICIAL COMO INSTRUMENTO: CONSECUENCIAS JURÍDICAS

Una de las formas erróneas de presentar la inteligencia artificial y su impacto es el símil antropomórfico. Hablar de «inteligencia artificial» ya es un antropomorfismo, porque, aunque estos sistemas puedan producir un objeto (un texto, una imagen) que es similar o idéntico a otro que podría producir un humano, el procedimiento es diferente, en la medida en que el sistema de inteligencia artificial no hace otra cosa que buscar la respuesta que parezca más adecuada o parecida a la que se deduzca de un conjunto de datos con el que ha sido previamente entrenado. Pero, con frecuencia, sobre ese antropomorfismo cabalga otro, cuando se la presenta como un ser omnisciente que podría dominar a los humanos.

Esta presentación es, a mi juicio, equivocada, porque la inteligencia artificial es, hoy por hoy, un *instrumento* en manos de los humanos (de unos más que de otros), a los que permite actuar de manera más eficiente, utilizando los datos para estar un poco menos a ciegas. El empresario utiliza la inteligencia artificial para ejercer el poder que ostenta en las relaciones laborales, a la hora de llevar a cabo acciones como seleccionar a su personal, organizar su trabajo, concretar sus retribuciones o aplicar sanciones y despidos. Pero la inteligencia artificial no le da ni le quita ningún poder, simplemente le suministra criterios (no vinculantes para él) a la hora de ejercerlo.

Lo mismo sucede con cualquier otro ámbito en el que se utiliza la inteligencia artificial, incluida la Administración. Al tratarse de un *instrumento*, la inteligencia artificial se encuentra sometida a las mismas normas, al mismo marco jurídico, que regula la actividad o la potestad en la que se utiliza. Por tanto, es rotundamente falso que la inteligencia artificial

carezca de regulación. Discriminar con inteligencia artificial es igual de inconstitucional que hacerlo sin inteligencia artificial. No es necesario retocar el artículo del Código Penal que castiga el homicidio cada vez que se inventa una nueva arma o un nuevo procedimiento para matar. Otra cosa es, por supuesto, que algunas armas o algunos procedimientos generen tantos riesgos que sea necesario establecer una regulación preventiva que reduzca al máximo la posibilidad de que se utilicen para causar daños.

No hay nada que, estando prohibido en el mundo analógico, pase a estar permitido sólo porque se actúe en un entorno digital o con inteligencia artificial. Otra cosa es que pueda ser más difícil aplicar determinadas normas o detectar su incumplimiento, pero conviene no poner el carro delante de los bueyes y no dar impresiones engañosas.

Cuando una decisión (de cualquier tipo) se toma con inteligencia artificial, se encuentra sometida al mismo marco jurídico, que será lo primero que haya que aplicar para ver si es admisible la utilización del instrumento, en la línea que hemos seguido al principio en relación con las decisiones administrativas.

Cuando los riesgos son demasiado altos es cuando tiene sentido una regulación preventiva, aquí y en cualquier otro campo. Ése es el sentido del Reglamento de Inteligencia Artificial, que prohíbe algunos usos de la inteligencia artificial y que, para aquellos que, estando permitidos, son de «alto riesgo» por el contexto en que se utilizan (es decir, porque las decisiones que ayudan a adoptar afectan más directamente a humanos en asuntos sensibles o pueden producir accidentes o daños materiales), establece unos requisitos de precaución y calidad que deberían minimizar el riesgo de que se produzcan predicciones o resultados erróneos y, por tanto, perjudiciales.

Sin embargo, esa regulación preventiva general no impide la aplicación de la normativa sectorial, que creo que acabará siendo tanto o más importante. El cumplimiento del Reglamento será algo fundamental, que sucederá «aguas arriba». Ningún producto de inteligencia artificial podrá ponerse en funcionamiento sin ese control, pero es algo previo a la solución de los problemas concretos que se planteen en su uso. La aplicación de inteligencia artificial, por ejemplo, a la actividad administrativa, exigirá el cumplimiento del Derecho administrativo, tanto de sus normas generales como de las que se establezcan (o incluso deban establecerse previamente) para el uso de la inteligencia artificial.

4. CONCLUSIONES

Los sistemas de inteligencia artificial se construyen para suministrar información (por ejemplo, predicciones) que sirva para responder a una pregunta, todo ello a partir del análisis y el hallazgo de correlaciones en datos del pasado. El sistema tendrá más probabilidades de acertar en la medida en que esos datos se parezcan a la situación actual a la que se aplica. La inteligencia artificial pretende reducir la incertidumbre y, por tanto, tendría cabida —técnicamente— en actividades (también de tipo jurídico o administrativo) en las que exista un margen de actuación, en las que sean posibles varias alternativas.

Los modelos de inteligencia artificial pueden servir para concretar conceptos normativos que son probabilísticos o de pronóstico, como los relacionados con el riesgo. Es un instrumento técnico cuyo encaje jurídico no sería muy diferente del de otros métodos técnicos que se utilizan en la actividad pública.

Otro campo habitual sería el de ayudar a centrar la actividad administrativa, como una aproximación que le indique campos en que puede ser más oportuna, necesaria o efectiva (parámetros definidos previamente por operadores humanos), pero sin tomar decisiones o resoluciones.

Ir más allá resultaría estructuralmente muy difícil (con independencia de que se admita legalmente o no) porque, o bien incumpliría el marco normativo (sustituyendo la constatación de hechos por una predicción o una probabilidad), o en realidad la intervención del sistema de inteligencia artificial sería instrumental de una decisión humana previa, que es la que elige las preguntas que va a responder el sistema o fija los parámetros con los que éste realizará una actividad auxiliar.

Una cosa es utilizar un modelo informático para *elaborar* una decisión a partir de un marco jurídico que la vincula y determina, y otra cosa es utilizarlo para *justificarla*, es decir, para argumentar por qué se elige una concreta opción. El grado de visibilidad, relevancia jurídica y exigencias no puede ser el mismo en un caso y en otro.

Cuando una decisión jurídicamente relevante se *justifica*, al menos en parte, en que es el resultado que ha producido un sistema de inteligencia artificial, en los procedimientos de control de dicha decisión se podrá y deberá verificar si ese sistema de inteligencia artificial ha funcionado de forma adecuada, lo que supone, al menos, dos cosas: que su diseño sea adecuado al marco jurídico que se está aplicando (es decir, que se le haya programado para buscar justamente aquello que es jurídicamente relevante, y no otra cosa distinta) y que su funcionamiento en concreto haya sido coherente con ese diseño.

El hecho de que en la *elaboración* de una decisión jurídicamente relevante se haya utilizado algún producto o procedimiento tecnológico (por ejemplo, un sistema de inteligencia artificial, incluidos los de tipo generativo) no necesariamente afecta a la validez de dicha decisión ni supone que en su control sea necesario analizar a fondo el funcionamiento de ese sistema de inteligencia artificial, porque el mismo permanece en un segundo plano y resulta irrelevante para verificar la legalidad de esa decisión.

La inteligencia artificial no es un ente que se enfrente a los humanos o amenace con dominarlos, sino un instrumento utilizado por algunos humanos para ejercer sus poderes, derechos y potestades (que son los mismos con independencia de que se utilice, o no, ese instrumento). La utilización de inteligencia artificial se encuentra automáticamente sometida al marco jurídico de esas potestades. Lo que está prohibido cuando se realiza de forma humana también lo está cuando se hace con inteligencia artificial. El marco jurídico o programación normativa de la actividad administrativa, que incluye normas sobre su contenido y sobre el procedimiento necesario para su producción, así como vías de impugnación y recurso, se aplica con independencia de que en su elaboración se hayan utilizado modelos de inteligencia artificial, que son un simple instrumento.

La inteligencia artificial es una actividad tecnológicamente compleja en la que pueden producirse resultados erróneos por múltiples razones. Si esos resultados son tomados como base para la toma de decisiones que afecten a personas, las decisiones podrán ser inadecuadas y contrarias al ordenamiento jurídico de muchas formas, empezando por la discriminación, pero no exclusivamente. Esto justifica que se haya elaborado una normativa específica para la inteligencia artificial, que intenta prevenir esos errores. Pero, con independencia de ella, el problema fundamental es que esas posibles decisiones inadecuadas están prohibidas por el marco jurídico que regule cada decisión en la que se haya utilizado inteligencia artificial, marco que debe aplicarse en todo caso, también cuando se utilice la IA.

BIBLIOGRAFÍA

ÁLVAREZ GONZÁLEZ, Elsa Marina (2022): *La función normativa y la técnica legislativa en España. Una nueva herramienta: la inteligencia artificial*, Valencia: Tirant lo Blanch.

ARIAS, Joaquín, Mar MORENO-REBATO, José A. RODRÍGUEZ-GARCÍA, José A y Sascha OSSOWSKI (2023): «Automated legal reasoning with discretion to act using s(LAW)», *Artificial Intelligence and Law*.

BARONA VILAR, Silvia (2022): «La seductora algoritmización de la justicia. Hacia una justicia poshumanista (Justicia+) ¿utópica o distópica?», *El Cronista*, 100, 36-47.

BOIX PALOP, Andrés (2020): «Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la Administración para la adopción de decisiones», *Revista de Derecho Púbico: Teoría y Método*, 1, 223-270.

BURKOV, Andriy (2019): *The hundred-page machine learning book*.

CAMPUZANO FERNÁNDEZ, Susana y Pilar VEGA VICENTE (2021): «Riesgos y oportunidades de la inteligencia artificial desde la perspectiva de la competencia. Un análisis desde la CNMC», *Boletín Económico de ICE (Información Comercial Española),* 3131, 43-55.

CASTETS-RENARD, Céline y Jessica EYNARD (dirs.) (2023): *Un Droit de l'intelligence artificielle. Entre règles sectorielles et régime général. Perspectives comparées*, Bruxelles: Bruylant.

CERRILLO I MARTÍNEZ, Agustí (2019): «El impacto de la inteligencia artificial en el Derecho administrativo. ¿Nuevos conceptos para nuevas realidades técnicas?», *RGDA*, 50.

COTINO HUESO, Lorenzo, (2023): «Discriminación, sesgos e igualdad de la inteligencia artificial en el sector público», en E. Gamero Casado (dir.) y F. Pérez Guerrero, (coord.), *Inteligencia artificial y sector público: retos, límites y medios*, Valencia: Tirant lo Blanch, 275-351.

COTINO HUESO, Lorenzo y Adrián TODOLÍ SIGNES (coords.) (2022): *Explotación y regulación del uso del big data e inteligencia artificial para los servicios públicos y la ciudad inteligente*, Valencia: Tirant lo Blanch.

CUBERO MARCOS, José Ignacio (2022): «Potestad administrativa de inspección e inteligencia aritificial: cuestiones controvertidas», en G. Vestri (dir.), *La disrupción tecnológica en la Administración Pública: retos y desafíos de la inteligencia artificial*, Cizur Menor: Aranzadi, 139-156.

DEGLI-ESPOSTI, Sara (2023): *La ética de la inteligencia artificial*, Madrid: CSIC-Catarata.

DE LA SIERRA, Susana (2020): «Inteligencia artificial y justicia administrativa: una aproximación desde la teoría del control de la Administración Pública», *Revista General de Derecho Administrativo*, 53.

DÍAZ GONZÁLEZ, Gustavo Manuel (2020), «Algoritmos y actuación policial: la policía predictiva», en A. Huergo Lora (dir.) y G.M. Díaz González (coord.), *La regulación de los algoritmos*, Cizur Menor: Aranzadi, 181-234.

DÍEZ SASTRE, Silvia (2008): *El precedente administrativo. Fundamentos y eficacia vinculante*, Madrid: Marcial Pons.

DIMATTEO, Larry A., Cristina PONCIBÒ, Cristina y Michel CANNARSA, Michel (eds.) (2022): *The Cambridge Handbook of Artificial Intelligence. Global Perspectives on Law and Ethics*, Cambridge: Cambridge University Press.

EBERS, Martin y Marta CANTERO GAMITO (eds.) (2021): *Algorithmic Governance and Governance of Algorithms. Legal and ethical Challenges*, Berlín: Springer.

FERNÁNDEZ FARRERES, Germán (2022): «El control judicial de la administración ante algunas cuestiones problemáticas», *REDA*, 219, 13-24.

FERNÁNDEZ SALMERÓN, Manuel (2019): «De la reutilización de sentencias al *big data* judicial. Aproximación a la metamorfosis experimentada por los modelos de uso de la información en el marco de la actividad jurisdiccional», en F. Gómez Manresa, M. Fernández Salmerón y M. Arenas Ramiro (dirs.), *Modernización digital e innovación en la administración de justicia*, Cizur Menor: Aranzadi, 63-102.

FROSINI, Vittorio (2019): *Cibernética, Derecho, internet y sociedad*, Santiago de Chile: Ediciones Olejnik.

GALLONE, Giovanni (2023): *Riserva di umanità e funzioni amministrative. Indagine sui limiti dell'automazione decisionale tra procedimento e processo*, Milano: Wolter Kluwer-CEDAM.

GAMERO CASADO, Eduardo, (2021) (dir.): *La potestad administrativa. Concepto y alcance práctico de un criterio clave para la aplicación del Derecho administrativo*, Valencia: Tirant lo Blanch.

— (2023): «Las garantías de régimen jurídico del sector público y del procedimiento administrativo común frente a la actividad automatizada y la inteligencia artificial», en E. Gamero Casado (dir.) y F. Pérez Guerrero, (coord.), *Inteligencia artificial y sector público: retos, límites y medios*, Valencia: Tirant lo Blanch, 397-461.

GAMERO CASADO, Eduardo (dir.) y Francisco PÉREZ GUERRERO, (coord.) (2023): *Inteligencia artificial y sector público: retos, límites y medios*, Valencia: Tirant lo Blanch.

GARCÍA MEXÍA, Pablo (2022): «Contra el "ciberfatalismo". Beneficio y riesgo en la sociedad digital», *Revista de las Cortes Generales*, 114, 285-354.

GARCÍA FERNÁNDEZ, Manuel (2022): «Collusion detection in public procurement auctions with machine learning algorithms», *Automation in Construction*, 133.

— (2023): «European public procurement (2012-2022): indicators and competition & SME maps for crossborder analysis based on TED data», *Research project for the Publications Office of the EU.*

GUCKELBERGER, Annette, (2019): Öffentliche Verwaltung im Zeitalter der Digitalisierung. Analysen und Strategien zur Verbesserung des E-Governments aus rechtlicher Sicht, Baden-Baden: Nomos.

HERNÁNDEZ PEÑA, Juan Carlos (2022): *El marco jurídico de la inteligencia artificial. Principios, procedimientos y estructuras de gobernanza*, Cizur Menor: Aranzadi.

HUERGO LORA, Alejandro (dir.) y Gustavo Manuel DÍAZ GONZÁLEZ, (coord.), (2020) *La regulación de los algoritmos*, Cizur Menor: Aranzadi.

HUERGO LORA, Alejandro, (1995): «La desigualdad en la aplicación de potestades de gravamen: remedios jurídicos», *RAP*, 137, 189-238.

— (2019): «Regulation versus tort. La función y los límites del Derecho administrativo», *Almacén de Derecho* [en línea], <https://almacendederecho.org/regulation-versus-tort-la-funcion-y-los-limites-del-derecho-administrativo>. [Consulta: 12/04/2024].

— (2020): «Una aproximación a los algoritmos desde el Derecho administrativo», en A. Huergo Lora (dir.) y G.M. Díaz González (coord.), *La regulación de los algoritmos*, Cizur Menor: Aranzadi.

— (2022) «Inteligencia artificial y sanciones administrativas», en M. Rebollo Puig, A. Huergo Lora, J. Guillén Camarés y T. Cano Campos (dirs.), *Anuario del Derecho Administrativo Sancionador*, 275-309.

— (2024a) «Por qué aciertan las sentencias sobre el «algoritmo» del bono social eléctrico», *Almacén de Derecho* [en línea], <https://almacendederecho.org/por-que-aciertan-las-sentencias-sobre-el-algoritmo-del-bono-social-electrico>. Consulta [12/05/2024].

— (2024b) «Distopía y algoritmo», *Almacén de Derecho* [en línea], <https://almacendederecho.org/distopia-y-algoritmos>. [Consulta: 17/05/2024].
IZQUIERDO CARRASCO, Manuel, (2022): «Actuaciones automatizadas en la sanción de irregularidades en el orden social», en L. Cotino Hueso y A. Todolí Signes (coords.), *Explotación y regulación del uso del big data e inteligencia artificial para los servicios públicos y la ciudad inteligente*, Valencia: Tirant lo Blanch, 261-288.
KEARNS, Michael y Aaron ROTH (2020): *The ethical algorithm. The science of socially aware algorithm design*, Oxford: Oxford University Press.
KOHL, Uta y Jacob EISLER (eds.) (2021): *Data-Driven Personalisation in Markets, Politics and Law*, Cambridge: Cambridge University Press.
LLANO ALONSO, Fernando y Joaquín GARRIDO MARTÍN (eds.) (2021): *Inteligencia artificial y Derecho. El jurista ante los retos de la era digital*, Cizur Menor: Aranzadi.
MARTÍN DELGADO, Isaac (2009): «Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada», *RAP*, 180, 353-386.
MARTÍN LÓPEZ, Jorge (2023): *Inteligencia artificial y comprobación tributaria: transparencia y no discriminación*, Cizur Menor: Aranzadi.
MARTINI, Mario (2019): *Blackbox Algorithmus -Grundfragen einer Regulierung Künstlicher Intelligenz*, Berliín: Springer.
MENÉNDEZ SEBASTIÁN, Eva (2023): *From bureaucracy to artificial intelligence*, Milano: Wolters Kluwer-CEDAM.
MIRANZO DÍAZ, Javier (2023): *Inteligencia artificial y Derecho administrativo. Especial referencia a los instrumentos de colaboración público-privada y contratación pública*, Madrid: ICA-Tecnos.
MORENO REBATO, Mar, (2021): *Inteligencia artificial (umbrales* éticos, Derecho y Administraciones Públicas), Cizur Menor: Aranzadi.
NIETO, Alejandro, (2000): *El arbitrio judicial*, Barcelona: Ariel.
— (2019): *Una introducción al derecho*, Valencia: Tirant lo Blanch.
NIKOLANIKOS, Nikos Th., (2023): *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies -The AI Act*, Berlin: Springer.
OROFINO, Angelo G. (2020): *La trasparenza oltre la crisi. Accesso, informatizzazione e controllo civico*, Bari: Circolo Editore.
PALMA ORTIGOSA, Adrián (2022): *Decisiones automatizadas y protección de datos. Especial atención a los sistemas de inteligencia artificial*, Madrid: Dykinson.
PONCÉ SOLÉ, Juli (2023): «Seres humanos e inteligencia artificial: discrecionalidad artificial, reserva de humanidad y supervisión humana», en E. GAMERO CASADO (dir.), *Inteligencia artificial y sector público: retos, límites y medios*, Valencia: Tirant lo Blanch, páginas 195-225.
— (2024a): «Por qué se equivocan las sentencias sobre el algoritmo del bono social eléctrico», *Almacén de Derecho* [en línea], <https://almacendederecho.org/por-que-se-equivocan-las-sentencias-sobre-el-algoritmo-del-bono-social-electrico>. [Consulta:13/06/2024].
— (2024b): «Utopía y algoritmos», *Almacén de Derecho* [en línea] < https://almacendederecho.org/utopia-y-algoritmos>. [Consulta: 12/06/2024].
PRESNO LINERA, Miguel Ángel (2022), *Derechos fundamentales e inteligencia artificial*, Madrid: Marcial Pons.
PUENTES COCIÑA, Beltrán y Andrei QUINTIÁ PASTRANA (2019): *El derecho ante la transformación digital. Oportunidades, riesgos y garantías*, Barcelona: Atelier.
RIBES, Aurora (2021): «La posición del contribuyente ante los Sistemas de Inteligencia Artificial utilizados por la Administración», *Quincena Fiscal*, 18.
RIVERO ORTEGA, Ricardo (2023): *Derecho e inteligencia artificial. Cuatro estudios*, Santiago de Chile: Olejnik.

SANCHEZ-GRAELLS, Albert, (2021): «Procurement Corruption and Artificial Intelligence: Between the Potential of Enabling Data Architectures and the Constraints of Due Process Requirements», en S. Williams y J. Tillipman (eds), *Routledge Handbook of Public Procurement Corruption*, Oxford: Routledge.

SAPOLSKY, Robert (2018): *Compórtate. La biología que hay detrás de nuestros mejores y peores comportamientos*, Madrid: Captain Swing.

SCHMIDT-ASSMANN, Eberhard (2003): *La Teoría General del Derecho administrativo como sistema. Objeto y fundamentos de la construcción sistemática*, Madrid: Marcial Pons.

SOLAR CAYÓN, José Ignacio (2019): *La inteligencia artificial jurídica. El impacto de la innovación tecnológica en la práctica del Derecho y el mercado de servicios jurídicos*, Cizur Menor: Aranzadi.

SORIANO ARNANZ, Alba (2021): *Data protection for the prevention of algorithmic discrimination. Protecting from discrimination and other harms caused by algorithms through privacy in the EU: possibilities, shortcomings and proposals*, Cizur Menor: Aranzadi.

— (2022): «El uso de sistemas automatizados por los poderes públicos. El caso del sistema empleado por el servicio público de empleo austriaco», en L. Cotino Hueso y A. Todolí Signes (coords.), *Explotación y regulación del uso del big data e inteligencia artificial para los servicios públicos y la ciudad inteligente*, Valencia: Tirant lo Blanch, 193-2011.

TODOLÍ SIGNES, Adrián (2022): «Inspección de trabajo, algoritmos, inteligencia artificial y su uso en la planificación inspectora», en L. Cotino Hueso y A. Todolí Signes (coords.), *Explotación y regulación del uso del big data e inteligencia artificial para los servicios públicos y la ciudad inteligente*, Valencia: Tirant lo Blanch, 236-259.

VALERO TORRIJOS, Julián (2020): «The Legal Guarantees of Artificial Intelligence in Administrative Activity: Reflections and Contributions from the Viewpoint of Spanish Administrative Law and Good Administration Requirements», *ERDAL*, 1, 55-62.

— (2023): «Las singularidades del tratamiento de datos de carácter personal en entornos de inteligencia artificial en el sector público», en E. Gamero Casado (dir.), *Inteligencia artificial y sector público: retos, límites y medios*, Valencia: Tirant lo Blanch, 353-396.

VELASCO RICO, Clara (2019): «Vigilando al algoritmo. Propuestas organizativas para garantizar la transparencia», en B. Puentes Cociña y A. Quintiá Pastrana (coords.), *El derecho ante la transformación digital: oportunidades, riesgos y* garantías, Barcelona: Atelier, 73-90.

VESTRI, Gabriele (dir.) (2020): *La disrupción tecnológica en la Administración Pública. Retos y desafíos de la inteligencia artificial*, Cizur Menor: Aranzadi.

WISCHMEYER, Thomas y Timo RADEMACHER, Timo (eds.) (2020): *Regulating Artificial Intelligence*, Berlín: Springer.

Fecha de recepción:1 de julio de 2024.
Fecha de aceptación: 20 de octubre de 2024.

LA ÉTICA DE LOS ALGORITMOS

THE ETHICS OF ALGORITHMS

Fernando H. Llano Alonso
Catedrático de Filosofía del Derecho
Universidad de Sevilla

RESUMEN

El presente trabajo analiza varios casos prácticos de ética aplicada a la IA en los que, ya sea por error y/o introducción de sesgos en el diseño de los algoritmos, o bien por la comisión de alguna imprudencia en el proceso de supervisión humana del sistema algorítmico, resultaron afectados derechos y libertades de las personas que fueron víctimas de ese fallo técnico y/o humano. Por otra parte, se estudian los tres modelos éticos que pueden servir de referencia a los diseñadores y programadores de algoritmos para que, mientras no sea del todo autónoma y general, de acuerdo con el cambio de paradigma que presumiblemente conllevará la singularidad tecnológica, la IA siga manteniéndose fiel al legado ético del humanismo ilustrado, antropocéntrico y antropogénico.

PALABRAS CLAVE

Utilitarismo. Ética deontológica, ética de la virtud, vehículo autónomo, supervisión humana.

ABSTRACT

This work analyses several practical cases of ethics applied to AI in which, either by error and/or introduction of biases in the design of algorithms, or by the commission of some imprudence in the process of human supervision of the algorithmic system, the rights and freedoms of the people who were victims of this technical and/or human failure were affected. On the other hand, we study the three ethical models that can serve as a reference for algorithm designers and programmers so that, as long as it is not fully autonomous and general, in accordance with the paradigm shift that technological singularity will presumably bring, AI remains faithful to the ethical legacy of enlightened, anthropocentric and anthropogenic humanism.

KEYWORDS

Utilitarianism, deontological ethics, virtue ethics, autonomous vehicle, human supervision.

DOI: https://doi.org/10.36151/TD.2024.106

LA ÉTICA DE LOS ALGORITMOS

Fernando H. Llano Alonso
Catedrático de Filosofía del Derecho
Universidad de Sevilla

«Una fracción de segundo es tiempo más que suficiente para que una computadora estudie concienzudamente todas las posibilidades. La decisión dependerá de las prioridades que haya fijado el software»

(McEwan, 2019)

1. INTRODUCCIÓN

El 9 de noviembre de 2023 el diario *Financial Times* publicó un artículo sobre el caso Sarah Meredith una joven residente en Cambridge a la que le diagnosticaron fibrosis quística al nacer. Se trata de una enfermedad genética que a largo plazo causa daños graves en órganos vitales como los pulmones, el aparato digestivo o el hígado. Efectivamente, en el verano de 2021, recién cumplidos los 28 años, el hígado de Sarah empezó a fallar tras casi tres décadas de constantes tratamientos experimentales, hospitalizaciones y terapias para controlar su enfermedad crónica.

Sarah fue ingresada de urgencia en el hospital de Plymouth, donde los hepatólogos que la atendieron calcularon que a su hígado le quedaban aproximadamente dos años y le recomendaron un trasplante. Tras aceptar someterse a un trasplante, Sarah se inscribió en una lista nacional de pacientes que esperaban la donación de un hígado. Por su parte, el Servicio Nacional de Salud del Reino Unido (NHS, por sus siglas en inglés) le informó de que el promedio de espera era entonces de 68 días, si bien esta expectativa se veía ensombrecida con un dato inquietante: el hecho de que solo seis personas que padecían las mismas afecciones crónicas que Sarah habían sido incluidas en la lista para un trasplante de hígado en la historia del NHS.

Con el fin de reducir el número de pacientes que fallecían mientras esperaban un trasplante, en 2018 el NHS introdujo un algoritmo de asignación de trasplante hepático denominado National Liver Offering Scheme (NLOS, por sus siglas en inglés) que cruzaba datos de pacientes en lista de espera de trasplante con el total de hígados donados en toda Gran Bretaña para calcular la puntuación obtenida por cada paciente, estableciendo así un orden de prelación en la lista de espera, de modo que a quien obtuviera la puntuación más alta se le ofrecería el hígado.

Sin embargo, este algoritmo había sido cuestionado desde su puesta en marcha por la opacidad de su funcionamiento, la ausencia de un procedimiento de apelación y la inexistencia de un sistema de supervisión humana que permitiera corregir o anular la puntuación ante un hipotético error de cálculo del algoritmo. En efecto, cuando la familia de Sarah preguntó a varios hepatólogos del NHS Blood and Transplant si sabían cómo calculaba el algoritmo el Transplant Benefit Score (TBS, por sus siglas en inglés), es decir, cómo funcionaba el *software* a la hora de asignar la puntuación que le correspondía a cada paciente en espera de trasplante, ninguno pudo responder con seguridad, aunque se inclinaban a pensar que Sarah tendría más posibilidades de recibir el trasplante que la mayoría de pacientes debido a su edad.

Un buen día, la familia Meredith encontró en Internet un sitio web creado por Ewen Harrinson, profesor de cirugía y ciencia de datos en la Universidad de Edimburgo. El Dr. Harrinson, también cirujano práctico en trasplantes, había creado una versión sencilla y accesible del algoritmo NLOS que permitía entender cómo se calculaba el TBS. En esta especie calculadora *online* se introducían algunos datos del paciente como la edad, el sexo y algunas mediciones específicas de su química sanguínea, y, de este modo, se obtenía la puntuación que le correspondería al enfermo como probable beneficiario de un trasplante.

Tras introducir los datos de Sarah, la puntuación obtenida era inferior al número de puntos necesario para conseguir que se le trasplantara un hígado sano. El resultado de Sarah no mejoraba ni siquiera cuando se modificaban datos variables como el de su edad. La triste conclusión a la que llegaron Sarah y su familia después de probar el sistema NLOS para la asignación de órganos es que las posibilidades de que fuera seleccionada para un trasplante eran finalmente muy remotas.

En la comunidad de hepatólogos del Reino Unido fue creciendo la sensación de que el algoritmo presentaba un sesgo de edad que perjudicaba sobre todo a los pacientes más jóvenes, entendiendo por tales los menores de 45 años (ese era el caso de Sarah Meredith). Entre los expertos que habían analizado el algoritmo NLOS destacaban David Spiegelhalter (Universidad de Cambridge) y Palak Trivedi (Universidad de Birmingham), cuyos análisis revelaron que el algoritmo partía de una premisa errónea: generalizar e igualar la tasa de mortalidad de todos los pacientes en lista de espera, sin reparar en el estado de desarrollo de la enfermedad de cada paciente con independencia de su edad. De esta manera, el algoritmo priorizaba a los pacientes de mayor edad frente a los más jóvenes porque se presumía que, al tener una edad más avanzada, vivirían menos que los más jóvenes. Paradó-

jicamente, el algoritmo NLOS llegó a asignar una puntuación superior a muchos pacientes de la tercera edad cuyos órganos estaban en un estado aceptable en comparación con el de algunos pacientes más jóvenes cuya vida dependía de un trasplante urgente.

Por otra parte, el sistema no tenía en cuenta otros datos tan relevantes como los años de vida saludable perdidos por los pacientes más jóvenes, muchos de los cuales habían nacido ya enfermos o habían desarrollado la enfermedad durante la infancia, y tampoco contemplaba los resultados a largo plazo de los enfermos más jóvenes (cuyos hígados habían estado deteriorándose más tiempo que el de los pacientes de mayor edad, muchos de ellos enfermos hepáticos sobrevenidos debido a adicciones como el alcoholismo); en definitiva, no se tuvo en cuenta el progresivo empeoramiento de los pacientes más jóvenes debido al tiempo de espera ni la reducción de su esperanza de vida en general.

El algoritmo NLOS estaba destinado a asignar hígados disponibles a aquellos enfermos que, según las estadísticas, tenían más posibilidades de beneficiarse de los órganos donados; sin embargo, no existen aún medidores exactos para calcular con precisión cuál es el beneficio esperado, entre otras razones porque no hay datos comparables entre los pacientes que fueron trasplantados y los que no recibieron una donación.

Por si fuera poco, los problemas surgidos a raíz de la asignación hepática automatizada iban más allá de los meros defectos estadísticos. Al igual que sucede con otros procesos automatizados de toma de decisiones, el algoritmo NLOS también presentaba serios problemas de diseño humano. En primer lugar, porque la configuración del sistema restringía la capacidad de acción de los expertos humanos y les impedía impugnar sus decisiones automatizadas. Y en segundo término porque, debido a la falta de transparencia de su funcionamiento, no había forma de apelar a casos excepcionales como el de Sarah Meredith.

Afortunadamente, la historia acabó bien para Sarah Meredith, pues el 13 de septiembre de 2023 recibió por fin la llamada que estaba esperando desde hacía tanto tiempo: los cirujanos del Hospital Addenbrooke de Cambridge le informaron de que disponían de un hígado de un donante cuatro décadas mayor que ella que pudieron asignarle directamente, ya que el proceso de selección no se sometió al control del algoritmo NLOS, sino que se llevó a cabo sin puntuación automatizada, es decir, a través del acuerdo alcanzado por el equipo de hepatólogos del hospital tras estudiar minuciosamente el historial clínico de cada uno de los candidatos en lista de espera.

Como se desprende del análisis del caso Meredith, los algoritmos, especialmente los modelos derivados directamente de datos a través de un aprendizaje automático plantean diversos retos. No solo permiten a un número importante de agentes tomar decisiones sin intervención humana, sino que, por su naturaleza tan compleja y opaca, ni siquiera sus diseñadores pueden prever cómo se comportarán en muchas situaciones.

Por otra parte, cuanto menor es la implicación directa de las personas con el algoritmo o modelo final, menos conciencia adquieren aquellas de los posibles efectos secundarios no pretendidos —éticos, morales o de otra índole— de dichos modelos, que constituyen el centro de interés del presente trabajo.

2. LA SUPERVISIÓN HUMANA DE LOS ALGORITMOS: ¿EFECTO PLACEBO O MECANISMO EFICAZ DE REGULACIÓN?

El de Sarah Meredith es un caso paradigmático que viene a engrosar la lista de incidencias y fallos surgidos en torno a la toma de decisiones algorítmicas. A este respecto, parece inevitable que, al conocer los efectos colaterales e inesperados generados por un sistema de IA que analiza y actúa a ciegas sobre ámbitos tan sensibles como la salud, la seguridad y/o los derechos humanos, algunos tiendan a rechazar de plano todo lo relacionado con los algoritmos y el aprendizaje automático.

Ahora bien, aunque no deben tomarse a la ligera los riesgos potenciales que comporta la toma de decisiones dependiente de los algoritmos, tampoco sería acertado que, para evitarlos, se prescindiera de sus beneficios para confiarlo todo a la supervisión humana. En este sentido, cualquier sistema de IA que se apoye en última instancia, principal o exclusivamente, en la supervisión humana no conseguirá abarcar completamente el enorme volumen y la velocidad de la toma de decisiones algorítmicas. Por ende, si no quieren verse superados por la magnitud del problema y volverse insuficientes, los enfoques inspirados solo en la supervisión humana terminarán cediendo el testigo a las decisiones algorítmicas.

Tampoco parece que la sobrerregulación de la IA sea un remedio suficiente para impedir los múltiples daños que puede ocasionar la toma de decisiones algorítmicas. Si bien es cierto que las leyes y los reglamentos no solo resultan fundamentales para reducir los riesgos y los daños generados por mal uso o el funcionamiento defectuoso de los sistemas de IA, sino también, y sobre todo, para asegurar que las acciones y omisiones que se codifican en los algoritmos estén enraizadas en la ética humanista y la búsqueda del bien común de la sociedad, la solución a los problemas generados a raíz de la toma de decisiones algorítmicas debiera ser, en gran medida, también algorítmica. Al igual que sucede con la supervisión humana, un enfoque excesivamente legalista de la IA adolece del mismo defecto: no nos permite comprender plenamente la escala inconmensurable de los problemas que eventualmente puedan ser provocados por la ingente cantidad de datos y la celeridad con la que estos se procesan y gestionan en la toma de decisiones algorítmicas (Kearns y Roth, 2020: 263).

A propósito del exceso de confianza en el efecto tranquilizador de la supervisión humana y en el poder de la regulación para contener y controlar el uso sin restricciones de la IA, algunos autores han señalado que, en realidad, esta respuesta no entraña más que una ingenua presunción alejada de la realidad del mundo tecnológico de nuestro tiempo. Ciertamente, la regulación del mecanismo de supervisión humana podría ejercer sobre nosotros un efecto calmante, es decir, funcionaría como una especie de efecto placebo regulador para la sociedad; aunque hay que admitir que, en rigor, dicho efecto sería tal vez inútil en la práctica (Koulu, 2020: 720-735).

La cuestión, por tanto, no debería centrarse tanto en hallar un modo de limitar o incluso prescindir de los algoritmos cuanto en la precisión del diseño y la definición de estas herramientas. Como advierten Kearns y Roth el problema no es el algoritmo en sí, sino la falta de correspondencia entre las entradas al algoritmo y el mundo real, dado que el sesgo está in-

crustado en los datos. No cabe esperar que el algoritmo sea capaz de descubrirlo y corregirlo por sí mismo. *Mutatis muntandis*, se trataría de la constatación de una de las máximas más citadas de la ciencia informática: «garbage in, garbage out» («si entra basura, sale basura»); dicho en otras palabras, si en el proceso de aprendizaje automático introducimos datos con sesgos, el algoritmo hará inevitablemente su predicción mediatizada por dichos sesgos.

A fin de ilustrar el error consistente en alimentar con datos desviados las profecías autocumplidas de los algoritmos y la amplificación del sesgo introducido en los mismos, ambos autores proponen el siguiente ejemplo:

Imaginemos que la jefatura de policía de la ciudad X utiliza modelos estadísticos para predecir en qué barrios hay un mayor índice de delincuencia y enviar allí un mayor número de policías. Pues bien, supongamos a continuación que, a pesar de que los barrios A y B de esta ciudad X presentan unas cifras de delincuencia subyacentes muy parecidas, la jefatura de policía decide mandar más policías al barrio A que al B. Lógicamente, esta decisión provocará que se registrarán más delitos en A que en B. En este caso hipotético, si utilizásemos los datos comparativos recabados tras la actuación policial en los dos barrios, estos datos servirán para retroalimentar sesgadamente la siguiente operación del modelo: confirmarán que lo correcto fue enviar más patrullas policiales al barrio A que al barrio B, y reforzarán cada vez más esa tendencia, razón por la cual la amplificación del sesgo en una suerte de bucle *ad infinitum*, estaría garantizada (Kearns y Roth, 2020: 131-132).

De todo cuando antecede se desprende una primera conclusión: para cerciorarse de que los modelos de toma de decisiones cumplen con las normas sociales que se desea respetar, no solo es necesario diseñar con claridad estos objetivos directamente en los algoritmos con los que se va a trabajar, sino también situar sus metas en un primer plano. Hay que tener en cuenta, además, que la toma de decisiones complejas automatizada que aporta el aprendizaje de las máquinas tiene un perfil propio y diferente del de su diseñador. Tal vez este conozca bien el algoritmo que utilizó para encontrar el modelo de toma de decisiones, pero no el modelo en sí (Kearns y Roth, 2020: 23 y 26).

Los diseñadores, los programadores y los desarrolladores de las complejas redes de algoritmos que componen la IA son las personas que mejor conocen sus limitaciones, inconvenientes y peligros, y, posiblemente, también las que tienen al alcance la mejor manera de mitigarlos. Para ello, es preciso que los científicos informáticos y los tecnólogos especializados en el aprendizaje automático se comprometan e impliquen como actores centrales en los debates éticos que puedan surgir en torno a la toma de decisiones algorítmicas. A este respecto, tampoco estaría de más contar con equipos multidisciplinares en el diseño e implementación de las programaciones algorítmicas y, muy especialmente, integrar en estos equipos interdisciplinares a especialistas en la materia específica de dichos programas (Ponce Solé, 2023: 223).

Al hilo de lo anteriormente expuesto, cabe añadir que la perspectiva ético-jurídica de la IA se presenta como un enfoque indispensable porque contribuye a allanar el desnivel que en ocasiones se produce entre los principios y valores de la sociedad respecto a los criterios prácticos del mercado aplicados por las organizaciones, compañías y entidades que recopilan, almacenan y gestionan nuestros datos personales.

3. LA FUNCIÓN CORRECTORA DE LOS PRINCIPIOS DE LA ÉTICA DE LA IA EN EL DISEÑO Y DESARROLLO DE LOS ALGORITMOS

La IA está tan inserta en nuestra vida cotidiana y en el contexto social que resulta necesario tomar en consideración desde un punto de vista ético los efectos y las consecuencias que esta tecnología produce sobre el género humano en la era de la automatización. Esta es la principal razón por la que resulta oportuna una investigación crítica en clave ética que no solo se ocupe de la eficacia y el rendimiento de las nuevas tecnologías, sino también del bienestar de la humanidad ante el horizonte de la singularidad tecnológica (Schaich Borg, Sinnott, Armstrong y Conitzer, 2024: 190-191).

El objetivo de este enfoque ético es doble: por un lado, promover la concordancia o el alineamiento entre las intenciones de las distintas partes implicadas y los valores éticos pertinentes para el uso previsto; por otro lado, identificar, corregir o denunciar las aplicaciones que sirven a fines éticamente inaceptables que ignoran, o violan, valores fundamentales en relación con su ámbito de actuación. Este segundo objetivo (la vigilancia orientada a evitar el uso indebido de los sistemas de IA) tiene una gran relevancia porque, si se verifica un uso desalineado de los sistemas, entonces la denuncia del problema y la puesta en marcha de acciones dirigidas a su eliminación constituye un acto obligatorio y absolutamente necesario.

En realidad, la ética es un ingrediente básico para que la investigación tecnológica pueda avanzar tranquilamente, explorar sus propias posibilidades y ser beneficiosa para la humanidad, dado que propicia el mantenimiento de un alto nivel de confianza entre los actores implicados en el desarrollo de la IA. En última instancia, como sostiene Stefano Quintarelli, razonar sobre el posible impacto moral de la IA y asegurarse de que tenga efectos beneficiosos sobre la existencia de la humanidad no tiene porqué suponer un freno al pleno desarrollo de la tecnología sino; por el contrario, es la mejor receta para su éxito a largo plazo y para cosechar todos los beneficios que promete (Quintarelli, 2020: 84).

Sin embargo, para que haya una alineación entre las tecnologías y las expectativas éticas, estas deben estar claramente definidas. El establecimiento de un marco de valores lo más coherente y claro posible es, en efecto, crucial para que la investigación y el desarrollo tecnológicos sean éticamente conscientes y se practiquen a la luz de esta conciencia. A propósito de la definición de este marco axiológico relativo a la IA, cabe hacer mención de la gran labor ético-jurídica realizada a través de las declaraciones internacionales sobre los principios éticos de la IA. En este sentido, Floridi (2022: 94-95) ha destacado los tres instrumentos que, a su juicio, son los más influyentes hasta ahora en la definición del marco ético-jurídico de la IA:

– Los Principios de Asilomar. Se trata de un conjunto de 23 principios formulados bajo los auspicios del Future Life Institute en colaboración con los participantes en la Conferencia Asilomar, celebrada en Pacific Grove (California) en enero de 2017. Estos principios hacen referencia, entre otras cuestiones, a la seguridad y la protección, la transparencia y la responsabilidad, el desarrollo y uso de las tecnologías de IA de forma justa y equitativa, así como a la garantía de que los sistemas de IA se

diseñen y operen sin daño ni discriminación de ningún individuo o grupo (es decir, que todas las personas tengan acceso al uso y el beneficio de la IA).

– La Declaración de Montreal para un desarrollo responsable de la inteligencia artificial (2018), que enuncia diez principios que sientan las bases para fomentar la confianza de la sociedad en los sistemas de IA. El sexto principio es el de equidad, que precisamente en su primer apartado exhorta al diseño y el entrenamiento de los sistemas de IA de manera tal que «[…] no creen, refuercen ni reproduzcan patrones de discriminación basados en diferencias sexuales, étnicas, culturales o religiosas, entre otras».
– La Declaración sobre Inteligencia Artificial, robótica y sistemas «autónomos», publicada en marzo de 2018 por el Grupo Europeo sobre Ética de la Ciencia y las Nuevas Tecnologías de la Comisión Europea (en adelante, EGE), en la que se proponen un conjunto de principios éticos fundamentales y prerrequisitos democráticos basados en los valores establecidos en los Tratados y en la Carta de derechos fundamentales de la UE: a) dignidad humana; b) autonomía; c) responsabilidad; d) justicia, equidad y solidaridad; e) democracia; f) Estado de Derecho y rendición de cuentas; g) seguridad, protección, e integridad física y mental; h) protección de datos y privacidad; y i) sostenibilidad.

A diferencia de los Principios de Asilomar y de la Declaración de Montreal, el apartado d) de la Declaración de la EGE alude explícitamente a los sesgos discriminatorios y manifiesta que la IA debería contribuir a la justicia global y facilitar la igualdad de acceso a los beneficios y ventajas de la IA, la robótica y los sistemas «autónomos». Concretamente, el final del primer párrafo de este cuarto parágrafo proclama que:

> «[…] los sesgos discriminatorios en los conjuntos de datos utilizados para entrenar y ejecutar los sistemas de IA deben evitarse. De no ser posible, estos sesgos deben ser detectados, notificados y neutralizados en la etapa más temprana del proceso».

Consciente de la necesidad de reconocer y regular el impacto de la IA en el sistema de derechos fundamentales, la Unión Europea se ha situado a la vanguardia de la creación de un marco jurídico específico sobre IA. En este sentido, el 14 de marzo de 2017 el Parlamento Europeo aprobó una resolución sobre las implicaciones de los macrodatos en los derechos fundamentales: privacidad, protección de datos, no discriminación, seguridad y aplicación de la ley. La resolución, que marcó un hito en el comienzo de la normativa de la Unión Europea sobre IA, insiste precisamente en el hecho de que:

> «[…] los ciudadanos, los sectores público y privado, el mundo académico y la comunidad científica solo podrán aprovechar plenamente las perspectivas y oportunidades que brindan los macrodatos si la confianza pública en esas tecnologías se garantiza mediante la estricta observancia de los derechos fundamentales y el cumplimiento de la legislación vigente de la Unión en materia de protección de datos, así como la seguridad jurídica en relación con todas las partes interesadas».

A su vez, en el parágrafo 20 de esta misma resolución, el Parlamento insta a la Comisión, a los Estados miembros y a las autoridades encargadas de la protección de datos a que:

> «[…] definan y adopten las medidas que se impongan para minimizar la discriminación y el sesgo algorítmicos y a que desarrollen un marco ético común sólido para el tratamiento transparente de los datos personales y la toma de decisiones automatizada que sirva de guía para la utilización de los datos y la aplicación en curso del Derecho de la Unión[1]».

A propósito de las declaraciones de principios de la IA, Floridi ha advertido que existe una suerte de hipertrofia de los principios éticos que son proclamados cada vez que una organización internacional aprueba una declaración, aun a riesgo de incurrir en innecesarias repeticiones de —o superposiciones con— los que han sido enunciados en otras declaraciones precedentes, recurrencia que conduce a la confusión y la ambigüedad.

Frente a esta profusión de principios éticos, Floridi propone la idea de que la IA no debe ser entendida como un nuevo tipo de inteligencia, sino como una forma de actuar sin precedentes. Por ello, sostiene el autor, de todas las áreas de la ética aplicada, la bioética es la que más se parece a la ética digital en el tratamiento que esta hace de nuevas formas de agentes, pacientes y entornos (Floridi, 2013). Sin embargo, aunque los principios bioéticos se adaptan perfectamente a los nuevos retos éticos planteados por la IA, no es fácil explicarlos. A este respecto, Floridi añade a los cuatro principios clásicos de la bioética (*beneficencia, no maleficencia, autonomía y justicia*) un quinto principio: la *explicabilidad*, entendida como un principio que incluye tanto el sentido *epistemológico* de *inteligibilidad* como la respuesta a esta pregunta: «¿cómo funciona?», del sentido ético de responsabilidad (*accountability*) d respuesta a la pregunta: «¿quién es responsable del modo en que funciona?» (Floridi, 2022: 96).

Estos cinco principios son válidos tanto para expertos —por ejemplo, los diseñadores o los ingenieros de productos— como para los no expertos, entre ellos los pacientes o losclientes (Watson y Floridi, 2020: 1-32).

A continuación, comentaremos sucintamente y por separado cada uno de estos cinco principios.

i) Beneficencia

De los cuatro principios tradicionales de la bioética, el de beneficencia es el más fácil de interpretar por su claridad. La mayoría de las declaraciones de principios de la IA coinciden a la hora de identificar este término con la idea de «bienestar» de los seres humanos y de todas las criaturas sintientes (Declaración de Montreal), y conciben este principio como un criterio-guía que prescribe que los desarrollos de la IA se orienten al bien común, el beneficio de la humanidad y del planeta (Conferencia de Asilomar).

ii) No maleficencia

Este principio complementa al de beneficencia, en la medida en que advierte de las diversas consecuencias negativas que se derivan del uso abusivo e inadecuado de las tecnologías de IA, entre ellas las violaciones de la intimidad personal. En este sentido, la Declaración de Montreal sostiene que quienes diseñan y desarrollan los sistemas de IA deben asumir su responsabilidad y actuar contra los riesgos derivados de la innovación tecnológica. Sin embargo, el noveno principio de esta declaración (el de responsabili-

dad) introduce el siguiente matiz al respecto: cuando un sistema de IA provoca daños o perjuicios y se demuestra que, a pesar de haber sido utilizado según lo previsto, es responsable de un daño, no resulta razonable culpar a las personas involucradas en su desarrollo y uso.

iii) Autonomía

Este principio se relaciona con el anterior en la medida en que la autonomía funcional de la máquina inteligente implica una suspensión del control humano en la realización de determinada tarea automatizada. No parece razonable atribuir al usuario la responsabilidad de los eventuales efectos negativos generados por el funcionamiento erróneo del sistema. En otras palabras, en caso de erosión del control del funcionamiento del sistema, la autonomía funcional puede considerarse una buena razón para limitar, también, la responsabilidad de los individuos implicados en su diseño y desarrollo como programadores, constructores, etc. (Powers y Ganascia, 2021: 31-32).

Ahora bien, si ningún actor humano puede ser considerado plenamente responsable de los efectos producidos por el funcionamiento de la IA, y si no basta tampoco con imputarle la culpa al propio sistema de IA, ¿cómo puede resolverse este dilema? A este respecto, hay dos alternativas posibles:

En primer lugar, recuperar algún tipo de control significativo sobre el funcionamiento autónomo de nuestras herramientas de IA. En segundo lugar, ampliar el concepto de responsabilidad más allá del paradigma del control (Quintarelli, 2020: 90).

iv) Justicia

Pese a su importancia, el principio de justicia es el que más acepciones presenta en las diferentes declaraciones de principios de la IA. La Declaración de Montreal apela al principio de justicia para que «[…] no se creen, refuercen ni reproduzcan patrones de discriminación basados en diferencias sociales, sexuales, étnicas, culturales o religiosas, entre otras» (sexto principio), mientras que uno de los Principios de Asilomar —concretamente, el decimoquinto— identifica la justicia con la prosperidad económica compartida; por lo demás, en otras declaraciones el término justicia tiene otros significados (por ejemplo, se hace equivalente a la equidad). En suma, como afirma Floridi, las distintas formas de caracterizar la justicia remiten, en última instancia, a una falta de claridad más amplia sobre la IA como reserva de acción inteligente creada por los seres humanos (Floridi, 2022: 100).

v) Explicabilidad

Para salir de esta confusión asociada a los significados divergentes atribuidos al principio de justicia, Floridi propone un quinto principio específicamente referido a la IA que complementa a los principios clásicos de la bioética: el principio de explicabilidad. La agregación del principio de explicabilidad, que incluye tanto el sentido epistemológico de «inteligibilidad» como el sentido ético de «responsabilidad», es «[…] la pieza crucial que falta para completar el puzle ético de la IA».

En efecto, este quinto principio complementa los otros cuatro porque, para que la IA sea beneficiosa y no dañina, es necesario que estemos en disposición de comprender el bien o el daño que está haciendo a la sociedad y de qué manera; por otra parte, para que la IA promueva y no limite la autonomía humana, nuestra decisión sobre quién debe decidir tiene que estar informada por un conocimiento previo: cómo actuaría la IA en nuestro lugar y, en tal caso, cómo podrían mejorarse sus prestaciones; por último, para que la IA sea justa, necesitamos saber a quién responsabilizar ética o jurídicamente en caso de que produzca un resultado grave o negativo, determinación que, a su vez, requeriría una comprensión adecuada de las razones por las que se produjo tal resultado (Floridi, 2022: 100).

4. EL CASO HERZBERG. CRÓNICA Y ANÁLISIS DEL PRIMER ERROR ALGORÍTMICO FATAL EN LA HISTORIA DE LA IA

La noche del 18 de marzo de 2018 un vehículo de prueba de la empresa Uber conducido por robot circula a 69 km/h por una avenida de cuatro carriles en Tempe (Arizona). En el automóvil autónomo va sentada en el asiento de copiloto Rafaela Vásquez, cuyo único cometido era mantenerse atenta a la conducción del coche robotizado para vigilar que no cometiera ninguna infracción del código de circulación.

Alrededor de las 22 horas, en medio de la North Mill Avenue, aparece sorpresivamente, como salida de la oscuridad, una mujer empujando una bicicleta en la que lleva todas sus pertenencias. Se trata de Elaine Herzberg, una indigente de mediana edad que cruza la autovía ajena a la circulación. Cuando el sistema de detección del Uber autónomo percibe a la Sra. Herzberg y los algoritmos dan la orden de frenado de emergencia, ya es demasiado tarde para evitar el atropello de la viandante, que muere una hora más tarde en el hospital.

¿Cómo pudo producirse ese error algorítmico que causó la primera víctima de la IA? ¿Por qué no actuó antes la Sra. Vásquez para evitar el impacto del vehículo que debía supervisar contra la Sra. Herzberg? ¿De qué modo estaba programado el *software* responsable de tomar la decisión última para que fallara, hasta el punto de que confundió a la desdichada peatona con un falso positivo?

El atropello de Elaine Herzberg se debió a un cúmulo de circunstancias adversas y de errores humanos e informáticos. Veamos seguidamente la sucesión de contratiempos y fallos que provocaron el fatal accidente y que nos permitirán entender cómo fue posible.

En primer lugar, la Sra. Herzberg cruzó la avenida de forma temeraria, en medio de la oscuridad, porque quería llegar pronto al campamento de personas sin hogar en el que vivía. No en vano, hay que señalar que ese tramo de carretera es precisamente uno de los trayectos en los que Uber prueba sus coches robotizados (hay que tener en cuenta que este tipo de pruebas son posibles desde que, en 2015, un decreto de Doug Ducey, gobernador de Arizona, autorizó a circular por las carreteras de este estado a vehículos controlados por un robot, incluso sin necesidad de que hubiera un ser humano dentro; a cambio de la aprobación de esta legislación tan permisiva con los coches autopilotados, Uber instaló su flota de automóviles en el estado del Gran Cañón).

En segundo lugar, Uber aprovechó la flexibilidad de la ley de seguridad vial de Arizona para reducir el número de *robot babysitters*, es decir, controladores humanos de los coches autónomos (la empresa tenía 400 vehículos de este tipo circulando por las carreteras de Tempe). A diferencia de su principal competidor, Waymo —filial de los coches autónomos de Google—, cuyos vehículos contaban con dos personas por automóvil para garantizar la seguridad en la conducción, Uber —compañía más preocupada por recortar gastos por los costes derivados de los sueldos de sus *robot babysitters* para obtener beneficios que por garantizar la seguridad de los demás conductores y la integridad física de los peatones— redujo a solo uno el número de conductores humanos a bordo de sus Volvo xc90 robotizados.

En tercer lugar, cabe subrayar que la conducta de la conductora auxiliar encargada de vigilar el buen funcionamiento del sistema de conducción inteligente fue imprudente. La normativa del estado de Arizona obliga a los vigilantes de los coches autónomos a llevar las manos sobre el volante en todo momento y a mantener permanentemente la atención en todo cuanto suceda tanto en el interior del automóvil como en la carretera para reaccionar a tiempo ante cualquier imprevisto que pueda presentarse y sobre el que la máquina no sepa qué decisión tomar. Sin embargo, la conductora auxiliar, Rafaela Vásquez no fue diligente en el cumplimiento de su obligación de supervisar el buen funcionamiento del sistema de conducción inteligente del Volvo xc90 en el que se desplazaba, y tampoco estuvo pendiente de la carretera porque solo estaba pendiente de la pantalla del móvil que llevaba en su regazo, en la que veía un programa de televisión.

La señora Vásquez parecía confiar plenamente en los sistemas de IA que guían la conducción del coche autónomo. En efecto, la flota de Volvo xc90 robotizados de Uber disponía de un complejo sistema de herramientas de IA capaz de aprender por sí mismo. Incluía dispositivos de radar, cámaras delanteras y laterales, sensores de navegación y escáneres láser (el denominado 'sistema Lidar', un aparato instalado en el techo del vehículo, similar a una sirena de la policía, que va disparando a su alrededor millones de pulsos láser por segundo para detectar objetos, medir distancias y formar a partir de ellos imágenes tridimensionales).

Como se ve, ni siquiera las predicciones de los algoritmos de un sistema de conducción inteligente tan complejo como el que llevan instalados los coches Uber son infalibles. En realidad, el problema son los falsos positivos que, como en el caso del atropello mortal de Elaine Herzberg, distorsionan los datos recabados del exterior del automóvil con los que los algoritmos elaboran predicciones de lo que podría suceder en los segundos que siguen a la detección de un objeto y determinan en qué dirección se moverá, si se trata de un peatón o un animal que invade la calzada, el tiempo aproximado de reacción antes de que se produzca el impacto con el supuesto objeto…

En el caso que nos ocupa, el sistema de detección percibió la presencia de la Sra. Herzberg con su bicicleta en mitad de la calzada seis segundos antes del atropello, tiempo suficiente para desviar la trayectoria del coche o para frenar. Sin embargo, los algoritmos del automóvil de Uber que se desplazaba en modo de conducción automático por la North Mill Avenue de Tempe mientras la conductora visionaba un programa de televisión en su

móvil dudaron y no se pusieron de acuerdo. En primera instancia, el *software* del vehículo autónomo creyó haber detectado un objeto desconocido, un segundo más tarde estimó que se trataba de un coche y, solo cuando faltaban 1.3 segundos para el impacto, confirmó que es una bicicleta y dio la orden de ejecutar el frenado de emergencia.

Teniendo en cuenta el desarrollo tecnológico ya existente en 2018, suficiente al menos para garantizar a los usuarios de un vehículo Uber autónomo la seguridad en la conducción de un *software* equipado con algoritmos capaces de hacer predicciones y dar instrucciones precisas durante la conducción, la pregunta que cabe hacer es cómo no se activó antes el frenado de emergencia pese a la detección de un objeto en la carretera con un margen de 6 segundos, una eternidad para un sistema de IA.

Según las revelaciones de dos antiguos empleados de Uber, tiempo después de estos acontecimientos, la empresa habría elevado el nivel de tolerancia a los falsos positivos en aquellos tramos de carretera elegidos para realizar las pruebas de conducción autónoma de sus vehículos con el único fin de ser más competitiva que sus rivales directos en el mercado de los VTC (Google, General Motors e Intel). La reducción de la seguridad vial a cambio de la mayor fluidez y puntualidad en el recorrido de prueba realizado por el Uber, es decir, sin detenerse cada dos por tres a causa de presuntos obstáculos que pudieran aparecer en la carretera, era un sacrificio aceptable si con ello se conseguía convencer a los inversores y a los potenciales usuarios de que estaban ante el vehículo del futuro.

El riesgo de esta estrategia empresarial tan agresiva y temeraria es que los fabricantes de automóviles autónomos den por descontado que los accidentes de tráfico se producen indefectiblemente, que los algoritmos que gobiernan sus sistemas de conducción cometen errores y que dichos errores cuestan vidas humanas. A propósito del alto riesgo asumido por las empresas de vehículos autónomos, Anthony Levandowski, el ingeniero responsable del diseño del coche de conducción automática de Google que robó datos de esta empresa para fundar su propia empresa, adquirida con posterioridad por Uber, afirmó en la Cumbre AV de coches autónomos de 2019 que la industria del automóvil necesita un avance fundamental en IA para que se produzcan progresos significativos en la tecnología de este tipo de vehículos: si de verdad queremos impulsar una tecnología, advirtió en su ponencia Levandowski, entonces la seguridad no puede ser la prioridad dado que, si así fuera, no se conseguiría nada.

En relación con esta lógica del mercado tecnológico, despojada de principios éticos inspirados por el paradigma de la moral humanista, vemos que hay casos extremos y paradójicos en los que el comportamiento de un automóvil autónomo responde solo a los principios *éticos* internos de la empresa, que quedan plasmados en el *software* del vehículo (Colomba, 2019: 94).

La fuerza performativa de la tecnología, que, como en el caso de la primera víctima registrada de la IA, reduce drásticamente el abanico de opciones posibles de actuación e invita a plantear algunos interrogantes sobre las transformaciones estructurales y el condicionamiento que la autorregulación de las empresas tecnológicas impone a la libertad de actuación de las personas (Lessig, 1999: 7-8).

Desde un punto de vista operativo, la lógica del presupuesto ético-moral que justifica el automóvil autónomo sería la siguiente: en primer lugar, las carreteras por las que circulen los coches *self-driving* son más seguras a medio plazo que las carreteras por donde circulan los coches guiados solo por conductores humanos; esto implicaría que, *ceteris paribus*, es inmoral no apostar por la fabricación masiva de automóviles autónomos. De lo que antecede se desprende, en segundo lugar, la siguiente conclusión: la *ratio* moral que impulsa la producción de automóviles autónomos no es punitiva («sufre quien lo merece»), sino *utilitarista*: según este enfoque ético, en el caso que hemos estudiado el del atropello mortal de la señora Herzberg, primaría el cumplimiento del objetivo empresarial, a saber, la puntualidad en cubrir el recorrido del trayecto elegido para hacer la prueba sobre el principio ético-moral de seguridad (Casadei y Zanetti, 2019: 45-46).

Por otra parte, esta autorregulación basada en la ética empresas de vehículos autónomos como los que componen la flota de Uber, Google, Intel o General Motors, también pone de relieve la cuestión del grado de responsabilidad que cabe atribuir a los diseñadores y desarrolladores de los algoritmos, esto es, cuestiona la sostenibilidad ética de las decisiones que toman y de la elección de opciones morales que hacen, en la medida en que estos asumen una responsabilidad pública al decantarse por determinado resultado o por un modelo de funcionamiento específico (De Vanna, 2019: 82).

Llegados a este punto, algunos especialistas en ética de la IA se han preguntado qué criterio ético aplicar a los vehículos autónomos. ¿Debería establecerse una ética única a través de un pacto social? ¿O bien dejar que las empresas se autorregulen? ¿O tal vez consentir la coexistencia de diferentes enfoques? ¿En el futuro podremos elegir nuestros coches autónomos en función de que su *software* que se alinee más con nuestros valores y no en función de sus prestaciones? ¿Cómo desearíamos que fueran diseñados los algoritmos del sistema que conducen ese automóvil? ¿De manera altruista o de modo que preserven ante todo la vida de su dueño? (Sigman y Bilinkis, 2023: 183; Schaich, Borg, Sinnott, Armstrong y Conitzer, 2024: 17).

En suma, podríamos concluir este epígrafe coincidiendo con el parecer de quienes piensan que el asunto de la confiabilidad de las máquinas inteligentes va más allá de una simple cuestión ética. Es probable que en el futuro los vehículos automatizados no sean conducidos por personas ni por ordenadores, sino por empresas que actuarán a través de sus operadores humanos y máquinas. Una cuestión esencial para este campo —y para la inteligencia artificial en general— es cómo deben ganarse nuestra confianza las empresas que desarrollan y despliegan estas tecnologías.

Una empresa que es digna de confianza comparte sin problemas su filosofía de seguridad, hace una promesa al público y la cumple. Una empresa que comparte su filosofía de seguridad no tiene inconveniente en informar a los usuarios con total transparencia con un mensaje parecido a este: «Esto es lo que hacemos; por esto pensamos que es razonablemente seguro; y por esto puede Vd. creernos». Su promesa sería la siguiente: «Solo comercializamos lo que razonablemente creemos que es seguro; seremos sinceros sobre nuestras limitaciones y fallos; y cuando fallemos, lo corregiremos». Y la empresa cumple esa promesa gestionando adecuadamente las expectativas del público, supervisando el ciclo de

vida de su producto o servicio y mitigando los daños de forma rápida, completa y pública (Smith, 2020: 682).

5. UN DILEMA MORTAL Y TRES MODELOS ÉTICOS PARA LOS SISTEMAS ALGORÍTMICOS

En 1967, la filósofa inglesa Philippa Foot, que defendía una ética contemporánea de las virtudes inspirada por la ética aristotélica (de acuerdo con la cual una acción es éticamente correcta si hacerla fuera propio de una persona virtuosa), publicó un artículo en la *Oxford Review* sobre del problema del aborto y la doctrina del doble efecto. En este estudio, la autora proponía varios ejemplos de dilemas éticos para explicar la doctrina del doble efecto: esta doctrina se basa en una distinción entre lo que un hombre *prevé* como resultado de su acción voluntaria y lo que, en sentido estricto, *pretende*. En otras palabras, este agente moral pretende, *strictu sensu*, tanto las cosas que se propone en cuanto fines como las que se propone en cuanto medios para conseguir sus fines. Estos últimos pueden ser lamentables en sí mismos, aunque deseados en aras de la consecución del fin propuesto.

Entre los diversos ejemplos utilizados por Foot para ilustrar el dilema que plantea la doctrina del doble efecto, quizás el problema del conductor del tranvía (*the trolley driver problem*) es el más extrapolable a la cuestión de la ética del diseño y desarrollo de los algoritmos. En realidad, se trata de un dilema al que se recurre cada vez que se pretende generar patrones de comportamiento humano en la toma de decisiones de los vehículos autónomos, en la tecnología de los drones bélicos configurados con IA para minimizar daños colaterales, o en los protocolos de ética médica —como en el caso de Sarah Meredith— concebidos pata tomar decisiones que afecta a la vida humana (Campione y Pietropaoli, 2024: 133-135).

Pasemos ahora a analizar el referido dilema moral que propone Foot:

> «Supongamos que un juez o magistrado se enfrenta a alborotadores que exigen que se encuentre a un culpable de un determinado delito y amenazan con vengarse sangrientamente de un sector particular de la comunidad. Como se desconoce el verdadero culpable, el juez se ve a sí mismo capaz de evitar el derramamiento de sangre incriminando únicamente a una persona inocente y haciéndola ejecutar. Al lado de este ejemplo se coloca otro en el que un piloto cuyo avión está a punto de estrellarse está decidiendo si se dirige de una zona más a una menos habitada. *Para que el paralelismo sea lo más cercano posible, puede suponerse más bien que es el conductor de un tranvía fuera de control, que solo puede conducir de una vía estrecha a otra; cinco hombres trabajan en una vía y un hombre en la otra; cualquiera que esté en la pista en la que entra está destinado a ser asesinado*. En el caso de los disturbios, la turba tiene cinco rehenes, por lo que, en ambos ejemplos, se supone que el intercambio es la vida de un hombre por la vida de cinco» (Foot, 1967: 8).

Las respuestas posibles a este dilema variarán en función de la teoría ética por la que nos decantemos:

En primer lugar, si optásemos por un enfoque ético consecuencialista o teleológico, como el utilitarista, la acción más correcta sería la que genera el mayor bien posible o que

comporta más bien que mal. En el caso del dilema del conductor del tranvía fuera de control, una persona que actuase conforme a las directrices morales de la ética utilitarista no dudaría en responder que la mejor opción, la más deseable, sería la de desviar el tren a la vía en la que hay un solo hombre.

Por el contrario, si nuestro patrón moral fuera la ética deontológica y nos fijásemos más en el valor propio de la acción que en las consecuencias de la misma, entonces la solución al dilema del conductor del tranvía no dependería de la cantidad de bien o felicidad que comportara la decisión. Desde el punto de vista del paradigma ético principialista o deontológico, representado por la ética kantiana, el maquinista debería actuar conforme a unas reglas de comportamiento universalizables, al margen de las consecuencias que pudieran producirse de esa actuación. En todo caso, la ética formal de Kant está coronada por unos principios éticos humanistas, iusracionalistas, ilustrados y antropocéntricos que fomenten tanto la autonomía de la voluntad como la constitución de una sociedad abierta de personas racionales y de ciudadanos libres e iguales.

En este sentido, aunque la decisión de desviar el tranvía hacia la vía en la que solo hay una persona sería considerada como ilícita por la ética deontológica, en última instancia esta sería sensible a la necesidad de ponderar las consecuencias globales de acciones morales ilícitas si están dirigidas a un fin mayor. Para definir cuáles serían esos principios de justicia universales que inspiran las reglas de comportamiento de un buen ciudadano dentro de una sociedad abierta y bien ordenada, podría ser útil la teoría contractualista-liberal de John Rawls, en la que, al contrario de lo que sucede con el utilitarismo, las personas aceptan por anticipado actuar (previo acuerdo con los demás participantes en el pacto social) según un principio de igual libertad sin un conocimiento previo de sus fines más particulares, de manera que convienen en adecuar sus concepciones del bien a lo que prescriben los principios de justicia, que son el resultado del acuerdo original entre los individuos —principio de libertades o de distribución de igual número de esquemas de libertades para todos; y principio de diferencia, que expresa un sentido de amistad cívica y solidaridad moral que incluye la igualdad en la estimación social, y excluye los privilegios o el servilismo dentro de la sociedad— (Rawls, 1971: 60-62).

Por último, para la ética de la virtud, la moral surge de los rasgos internos de las personas, de las virtudes, que se contraponen tanto a los principios de la ética deontológica, en la que la moral proviene de las normas, y a los de la ética consecuencialista, en la que la moral depende de los resultados del acto. En consonancia con el enfoque moral de la ética de la virtud, Philippa Foot sostiene que el fin no puede justificar los medios, como propone el utilitarismo. Por lo tanto, para la teoría de la virtud la respuesta al dilema del conductor del tranvía pasaría por una solución opuesta a la de la teoría de la ética consecuencialista: el conductor debería dejar que el tranvía siga por la vía principal y no hacer ninguna maniobra de desvío, aunque fuera a costa de arrollar a los cinco operarios que trabajan en esa vía. Según Foot, si el conductor no interviene y se queda quieto, no sería responsable moral de las cinco muertes, mientras que, si desvía el tranvía a la pista secundaria en la que se encuentra un hombre, el conductor sería parcialmente responsable de su muerte y su actuación constituiría una participación en el mal moral.

Como se ha comentado antes, el dilema del conductor del tranvía es perfectamente aplicable al diseño y el desarrollo de los sistemas algorítmicos. Una buena muestra de su utilidad lo encontramos precisamente en el diseño del *software* para controlar automóviles autónomos (Hazel Si Min y Taeihagh, 2019: 5791).

Desde el año 2016, el Massachusetts Institute of Technologie (M.I.T, por sus siglas en inglés) viene desarrollando en su web un programa pedagógico de supuestos prácticos en los que se actualiza el clásico dilema del tranvía, ya lejano del imaginario de la sociedad tecnológica, y este se sustituye por el coche autónomo. Esta plataforma *online*, denominada *Moral Machine* (en adelante, MM), se presentan diferentes escenarios dilemáticos en los que un vehículo sin conductor debe elegir entre dos males. Basándose en las respuestas recabadas de los millones de personas que visitaron la web y participaron en esta prueba, se ha podido establecer una imagen muy aproximada de la compleja y heterogénea perspectiva de los seres humanos sobre las decisiones tomadas por máquinas inteligentes.

Veamos, a continuación, la representación gráfica y la explicación de la primera combinación de partida, que reproduce el dilema del automóvil automático sin frenos inspirándose en el dilema clásico del tranvía:

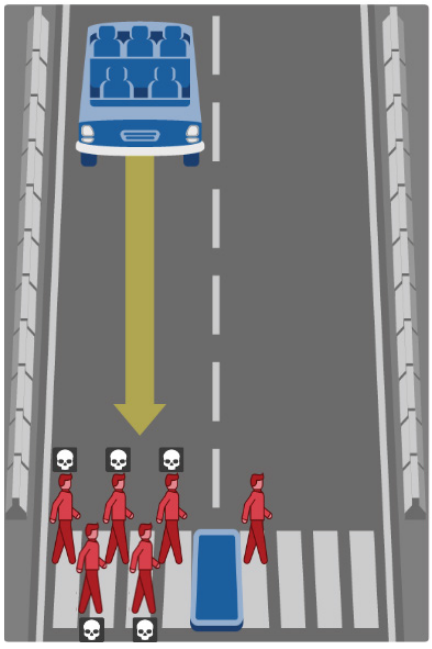

Figura 1. En este caso, el coche autónomo con fallo en los frenos sigue adelante y atraviesa el paso de cebra arrollando a los cinco peatones.

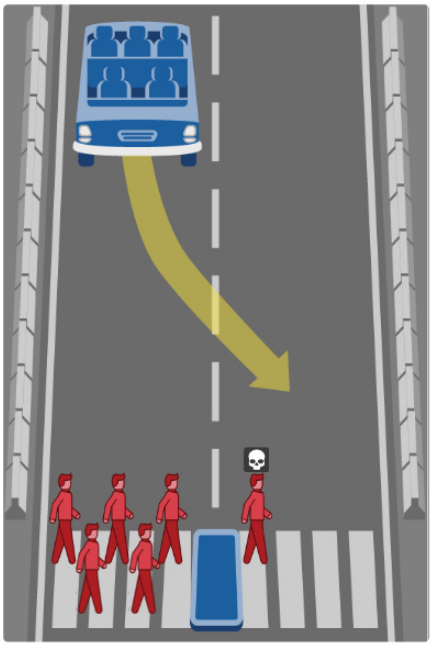

Figura 2. En este caso, el coche autónomo con fallo en los frenos gira, invade el otro carril, atraviesa el paso de cebra y arrolla al único peatón que cruza el paso de peatones

Además de su utilidad para la programación de la IA que toma decisiones, la finalidad de esta prueba de simulación se dirige principalmente a alinear los algoritmos morales

con los valores humanos mediante una «conversación global» sobre cuáles son esos valores aceptados y consensuados (Bonnefon, Shariff y Rawan, 2016).

La plataforma MM permite a sus usuarios diseñar una infinitud de escenarios para compartir, examinar y discutir *online* con otros usuarios. Las respuestas recopiladas proporcionan datos de gran valor, en la medida en que reflejan la percepción humana de la autonomía de las máquinas. Aunque el problema original del tranvía es un experimento mental útil que los filósofos utilizan para explorar las relaciones entre las intuiciones morales humanas y las predicciones de distintas teorías filosóficas, en puridad no es una guía útil para el diseño de robots corporales en el mundo físico. Para diseñar un robot ético (como un vehículo autónomo), hay que salir del estrecho marco figurativo del problema del tranvía y formular una norma social adicional.

Una situación figurada como la que nos plantea el dilema mortal del tranvía, que no ofrece opciones ni resultados éticamente buenos (en ambos caos se produce un atropello mortal), debería desencadenar un pensamiento contrafáctico, de modo que el sistema algorítmico de conducción autónoma aprenda que una situación que antes no tenía nada de especial, como entrar en una calle estrecha, exige una reducción de velocidad, para preservar la opción de realizar una parada de emergencia que salve la vida de una persona o de un grupo de personas que puedan cruzar la carretera (Kuipers, 2020: 429).

Dado que la IA actúa de manera diferente a los humanos en determinadas condiciones, tener idea de si una entidad es humana o IA hará que su comportamiento sea más predecible para los demás, aumentando tanto la eficacia como la seguridad (Walsh, 2017: 113-114).

Si una persona se pone delante de un vehículo que circula a 69 km/h, el conductor humano medio puede que sea incapaz de reaccionar con la rapidez suficiente para adoptar una maniobra evasiva, mientras que un sistema de IA sí podría hacerlo. Sin embargo, en otras situaciones, sobre todo las que requieren «sentido común», es probable que la IA (al menos de momento) sea muy inferior al ser humano. Los coches de IA pueden ser expertos en una autopista, pero evaluar la presencia de elementos complejos o inusuales —por ejemplo, una persona que empuja una bicicleta— puede ser más difícil para ellos. Del mismo modo que a un niño pequeño le hablamos de forma diferente, a una IA le daremos instrucciones distintas a las que damos a los humanos en aras de nuestra protección, pero también de la protección de la propia IA (Turner, 2019: 322).

6. CONCLUSIÓN

A lo largo de este recorrido por la casuística más reciente en torno a las implicaciones éticas de las decisiones tomadas por la IA, hemos podido comprobar que los modelos predictivos que sirven para el entrenamiento de los ordenadores en el procesamiento y análisis de datos se construyen no sólo a partir de los datos, sino también de las decisiones que tomamos sobre a qué datos prestar atención y cuáles omitir. Estas decisiones no solo tie-

nen que ver con la logística, los beneficios y la eficiencia. Son fundamentalmente morales (O'Neil, 2016: 218).

En primera instancia, puede parecer plausible modelizar sistemas éticos con técnicas de IA, ya que las prescripciones en las que se basan dichos sistemas han sido introducidas por humanos. Sin embargo, los intentos de modelizar el razonamiento ético han puesto de manifiesto las enormes dificultades a las que se enfrentan los investigadores para hacerlo. La primera dificultad proviene del modelado del razonamiento deóntico, es decir, el razonamiento sobre obligaciones y permisos. La segunda se debe a los conflictos de normas que se producen constantemente en el razonamiento ético. La tercera está relacionada con el entrelazamiento del razonamiento y la acción, que exige que estudiemos la moralidad del acto, *per se*, pero también los valores de todas sus consecuencias.

El reto técnico actual consiste en fusionar estos tres enfoques, es decir, crear uno que no sea uniforme, que pueda gestionar conflictos entre normas y que utilice modelos causales para evaluar las consecuencias de las acciones. Aunque existe un interés general en la creación de una máquina moral de este tipo (es decir, que se comporte de acuerdo con las normas de un enfoque moral), todas estas perspectivas abarcan distintos marcos normativos —el utilitarismo, la deontología y la ética de la virtud— que deben simularse. Los detalles de las simulaciones suelen resultar insuficientes, sobre todo para los filósofos. Además, surgen cuestiones sobre la utilidad práctica de tales máquinas morales, así como sobre las dificultades para ponerlas en práctica (Powers y Ganascia, 2021: 40-41).

En realidad, la ética es un ingrediente básico para que la investigación tecnológica pueda avanzar sosegadamente, para que explor sus propias posibilidades y para que sea beneficiosa para la humanidad, en la medida en que propicia el mantenimiento de un elevado nivel de confianza entre los actores implicados en el proceso de la IA. En última instancia, como sostiene Stefano Quintarelli, razonar sobre el posible impacto moral de la IA y asegurarse de que tenga efectos beneficiosos sobre la existencia de la humanidad no tiene porqué suponer un freno al pleno desarrollo de la tecnología; por el contrario, como ya se ha dicho, es la mejor fórmula para que alcance su éxito a largo plazo y, por ello, para que cosechemos todos los beneficios que promete (Quintarelli, 2020: 84).

NOTAS

1. Resolución del Parlamento Europeo, de 14 de marzo de 2017, sobre las implicaciones de los macrodatos en los derechos fundamentales: privacidad, protección de datos, no discriminación, seguridad y aplicación de la ley (2016/2225(INI)), §§ 1 y 20.

BIBLIOGRAFÍA

BONNEFON, Jean-François, Azim SHARIFF e Iyad RAHWAN (2016): «The Social Dilemma of Autonomous Vehicles», *Science*, 352 (6293), 1573-1576.

CAMPIONE, Roger y Stefano PIETROPAOLI (2024): *Los artefactos de la inteligencia jurídica: personas y máquinas*, Madrid: Dykinson.
CASADEI, Thomas y Gianfrancesco ZANETTI (2019): «Tra dilemmi etici e potenzialità concrete: le sfide dell´autonomous driving» en S. Scagliarini (dir.), *Smart Roads e driverless cars: tra diritto, tecnologie, etica publica*, Turín: G. Giappichelli Editore, 41-54.
COLOMBA, Vittorio (2019): «Driverless cars e intelligenza artificiale» en S. Scagliarini (dir.), *Smart Roads e driverless cars: tra diritto, tecnologie, etica publica*, Turín: G. Giappichelli Editore, 87-95.
DE VANNA, Francesco (2019): «Autonomous driving e questione della responsabilità: alcuni nodi teorici», en S. Scagliarini (dir.), *Smart Roads e driverless cars: tra diritto, tecnologie, etica publica*, Turín: G. Giappichelli Editore, 77-86.
FLORIDI, Luciano (2022): *Etica dell´intelligenza artificiale. Sviluppi, opportunità, sfide,* Milano: Raffaelo Cortina.
FOOT, Philippa (1967): «The Problem of Abortion and the Doctrine of the Double Effect», *Oxford Review*, 5, 5-15.
HAZEL SI MIN, Lim y Araz TAEIHAGH (2019): «Algorithmic Decision-Making in AVs: Understanding Ethical and Technical Concerns for Smarts Cities», *Sustainability*, 11(20), 5791.
KEARNS, Michael y Aaron ROTH (2020): *El algoritmo ético. La ciencia del diseño de algoritmos socialmente responsables*, Madrid: La Ley-Wolters-Kluwer.
KOULU, Riikka (2020): «Proceduralizing control and discretion: Human oversight in artificial intelligence policy» *Maastricht Journal of European and Comparative Law*, 27(6), 720-735.
KUIPERS, Benjamin (2020): «Perspectives on Ethics of AI: Computer Science», en M. D. Dubber, F. Pasquale y S. Das (eds.), *The Oxford Handbook of Ethics of AI*, Oxford: Oxford University Press, 421-441.
LESSIG, Lawrence (1999): *Code and Other Laws of Cyberspace*, Nueva York: Basic Books.
O´NEIL, Cathy (2016): *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy*, Londres: Penguin Books.
PONCE SOLÉ, Juli (2023): «Seres humanos e inteligencia artificial: discrecionalidad artificial, reserva de humanidad y supervisión humana», en E. Gamero Casado (dir.), *Inteligencia Artificial y sector público. Retos, límites y medios* Valencia: Tirant lo Blanch, 196-225.
POWERS, Thomas y Jean-Gabriel GANASCIA (2021): «The Ethics of the Ethics of AI», en M. D. Dubber, F. Pasquale y S. Das (eds.), *The Oxford Handbook of Ethics of AI* (ed.), Oxford, Oxford: University Press, 27-51.
QUINTARELLI, Stefano (2020): *Intelligenza Artificiale. Cos´è davvero, come funciona, che effetti avrà*, Milán: Bollati Boringhieri.
RAWLS, John (1971): *A Theory of Justice*, Cambridge (Massachusetts) – London: The Belknap Press of Harvard University Press.
SCHAICH BORG, Jana, Walter SINNOTT y Vincent ARMSTRONG-CONITZER (2024): *Moral AI and How We Get There*, Milton Keyes – Dublin: Penguin Random House.
SIGMAN, Mariano y Santiago BILINKIS (2023): *Artificial. La nueva inteligencia y el contorno de lo humano*, Barcelona: Debate.
SMITH, Bryant Walker (2020):«Ethics of Artificial Intelligence in Transport» en M. D. Dubber, F. Pasquale y S. Das (eds.), *The Oxford Handbook of Ethics of AI*, Oxford: Oxford University Press, 672-683.
TURNER, Jacob (2019): *Robot Rules. Regulating Artificial Intelligence*, Londres: Palgrave Macmillan.
WALSH, Toby (2017): *Android Dreams*, Londres: Hurt&Co.
WATSON, David S. y Luciano FLORIDI (2020): «The explanation game: A formal framework for interpretable machine learning», *Synthese*, 198, 9211-9242.

Fecha de recepción: 1 de julio de 2024.
Fecha de aceptación: 20 de octubre de 2024.

DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL EN EL SEGURO: MANIFESTACIONES Y ENCAJE LEGAL*

DIGITALISATION AND ARTIFICIAL INTELLIGENCE IN INSURANCE: MANIFESTATIONS AND LEGAL FRAMEWORK

María Luisa Muñoz Paredes

Catedrática de Derecho mercantil
Universidad de Oviedo

RESUMEN

Este trabajo ofrece una panorámica de los cambios que está experimentando el contrato de seguro y su distribución debido a la introducción de la IA, que afectan a todas sus fases, desde la publicidad hasta la selección de riesgos, pasando por la comunicación con los asegurados, el control del riesgo durante el contrato, la gestión de siniestros y la prevención de fraudes. La evaluación de riesgos y la fijación de precios en los ramos de salud y vida es considerada como de alto riesgo por el Reglamento de inteligencia artificial con toda lógica, pues los riesgos de exclusión y discriminación entre asegurados requieren un control especial. Por otro lado, este fenómeno se manifiesta también en la aparición de nuevos tipos de contratos (telemáticos, *on-demand*, paramétricos) que se alejan netamente de los seguros clásicos. Por su parte, el embebimiento de seguros en las contrataciones *online* supone una ruptura con el modo de distribución tradicional. En fin, también se estudia el efecto que puede tener la aplicación de la EU Data Act en los seguros telemáticos.

PALABRAS CLAVE

Digitalización del seguro, IA en el seguro, Reglamento de inteligencia artificial, seguros paramétricos, seguros basados en el uso, seguros *on-demand*, seguros embebidos, Reglamento de Datos de la UE.

ABSTRACT

This paper provides an overview of the changes that the insurance contract and its distribution are undergoing due to the introduction of AI, which affect all phases of the contract, ranging from advertising and risk selection to communication with policyholders, risk control during the contract, claims management and fraud prevention. Risk assessment and pricing in health and life insurance is logically considered as high risk by the artificial intelligence Regulation, as the risks of exclusion and discrimination between policyholders require special control. This phenomenon is also reflected in new types of contracts (telematic, on-demand, parametric) that are a clear departure from traditional insurance. For its part, the embedding of insurance in online contracts represents a clash with the traditional mode of distribution. Finally, it also examines the effect that the application of the EU Data Act may have on telematic insurance.

KEYWORDS

Digitalization of insurance, AI in insurance, EU Artificial Intelligence Act, parametric insurance, usage-based insurance, on-demand insurance, embedded insurance, EU Data Act.

DOI: https://doi.org/10.36151/TD.2024.107

---

* Este trabajo se ha realizado en el marco del Proyecto de Investigación MCIU-22-PID2021-127094OB-I00, «El seguro en la era digital» (DIGINSURANCE).

DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL EN EL SEGURO: MANIFESTACIONES Y ENCAJE LEGAL

María Luisa Muñoz Paredes
Catedrática de Derecho mercantil
Universidad de Oviedo

1. INTRODUCCIÓN

En los últimos años, el seguro (entendido como contrato) y el propio mercado asegurador han experimentado cambios constantes motivados, de un lado, por el hecho de que no son ajenos al proceso de digitalización que se ha impuesto en las relaciones comerciales, y, de otro, porque se han visto notablemente afectados por la utilización de la inteligencia artificial. Esto último no es extraño. Una compraventa puede celebrarse de manera esporádica y, en muchas ocasiones, no requiere más que el acuerdo de las partes sobre la cosa y el precio, mientras que el de seguro sigue siendo, posiblemente, el contrato mercantil más

complejo. Tradicionalmente, su celebración presuponía una fase precontractual (art. 10 de la Ley 50/1980, de 8 de octubre, de Contrato de Seguro, en adelante LCS) en la que, a partir de la sola base de las declaraciones del potencial tomador, la aseguradora tenía que evaluar el riesgo para decidir si lo asumía o no y, en su caso, en qué condiciones. Durante la vigencia del contrato, y partiendo de la premisa de que la aseguradora no tiene acceso directo al riesgo, la LCS (art. 11) impone al asegurado (que es el que lo tiene cercano) la obligación de comunicar las agravaciones que sufra a lo largo del curso del contrato para que aquella pueda decidir si le compensa continuar cubriéndolo o no y, eventualmente, reclame un aumento de la prima. Hasta hace poco tiempo, una vez producido un siniestro, si la cobertura pactada se extendía a los daños en las cosas, no surgía sin más el derecho al cobro de una cantidad pactada (salvo en las llamadas «pólizas estimadas» previstas en el art. 28 LCS, y solo si el siniestro era total y se había elegido el valor estimado como suma asegurada), sino que se procedía a la averiguación por peritos de los menoscabos efectivamente derivados del accidente.

Así, la indemnización a cargo de la aseguradora de daños tenía tres límites: *i)* el valor del interés en el momento inmediatamente anterior al acaecimiento del siniestro; *ii)* la suma asegurada, que era la cantidad que constituía el límite legal de responsabilidad de la aseguradora (art. 27 LCS) y, además, la base de la prima; y *iii)* los concretos daños sufridos de acuerdo con la valoración pericial. Pues bien, incluso estos elementos básicos del seguro (la selección de los riesgos, el control del riesgo durante el contrato y la liquidación del siniestro), que se habían mantenido inalterados en nuestra historia reciente, aparecen ahora sustancialmente transformados. Por una parte, las aseguradoras ya no dependen únicamente de los datos que le proporcione el tomador y cuentan con más fuentes de información. Por otra, gracias a la utilización de la inteligencia artificial, las compañías van creando nuevas formas de aseguramiento que se adaptan mejor a las necesidades de los asegurados y que, en algunos casos, les permiten controlar directamente el riesgo a través del uso de dispositivos telemáticos y, en otros (tal es el caso del llamado «seguro paramétrico», del que luego nos ocuparemos), abandonar incluso la exigencia de que el asegurado acredite que ha sufrido un daño como requisito para el pago de la indemnización, a pesar de que se trate de un seguro de daños.

En otro orden de cosas, la European Insurance and Occupational Pensions Authority (EIOPA, 2021: 21) ha manifestado su preocupación por el progresivo aumento de personas que se quedan sin seguro porque son calificadas como de alto riesgo, una tendencia que contrasta con el pasado reciente, en el que la mayoría de las personas podían acceder a una póliza porque se afinaba menos en el proceso de selección de riesgos. Por ello, el organismo europeo ha solicitado a las aseguradoras que desarrollen no solo políticas de no discriminación —concretadas, principalmente, en la exclusión de factores que no influyan realmente en el riesgo a la hora de calcular las primas—, sino también políticas de justicia orientadas a que las personas y los grupos especialmente vulnerables no sean privadas del aseguramiento. Asimismo, ha indicado que, en el marco de su responsabilidad social corporativa, las empresas deben tener en cuenta los problemas de inclusión financiera y considerar diversas vías para evitar la reproducción o el incremento de las desigualdades

preexistentes, especialmente en el caso de aquellos productos que son socialmente beneficiosos. Es oportuno observar, en todo caso, que la precisión en la estimación del riesgo no siempre supone un aumento del número de personas que quedan sin seguro. Determinados colectivos que tradicionalmente se consideraban de alto riesgo —típicamente, los conductores jóvenes, evitados por las aseguradoras debido al factor etario— pueden, en el nuevo escenario, acceder a la contratación de seguros mucho más fácilmente y en mejores condiciones económicas siempre que acepten que la aseguradora controle el riesgo a través de un dispositivo conectado al vehículo. Así, estos asegurados pueden demostrar que su riesgo no es alto simplemente por el hecho de que son jóvenes.

La evolución del control del riesgo es esencial y ha conferido una nueva función al seguro: tradicionalmente, este se contrataba en un momento determinado y la aseguradora solo se comunicaba con el asegurado cuando cobraba la prima en los sucesivos vencimientos y cuando este le reclamaba el pago de un siniestro; normalmente, la empresa aseguradora no hacía un seguimiento de las variaciones del riesgo entre estos momentos y su función se limitaba a la reparación de los daños sufridos en caso de siniestro. Esta dinámica ha cambiado radicalmente: en la actualidad, la compañía puede perfectamente ofrecer seguros de todo tipo (de vida, de salud, de auto o de hogar) en los que el asegurado acepta que la compañía sepa en cada momento en qué estado está «su» riesgo y esta tenga la posibilidad de anticiparse a la producción de un siniestro, o al menos evitar que los efectos perjudiciales del mismo sean mayores. La aseguradora no se limita ya a constatar las variaciones del riesgo, sino que, además, aconseja o prescribe la observancia de determinadas pautas de comportamiento al asegurado (relativas, entre otros aspectos, a los hábitos de vida saludable o al comportamiento en la conducción de vehículos) para que este se implique en su reducción. Así, la función meramente reparadora da paso a una función *preventiva* que amplía la virtualidad del seguro. Si bien es cierto que en el nuevo panorama existen variaciones en la concreción de determinados extremos —entre ellos, la configuración contractual de este tipo de seguro o la alternativa entre la eventual reducción de la prima por la reducción del riesgo y el mero ofrecimiento de ventajas promocionales temporales al asegurado por este motivo—, en términos generales la idea de que el seguro no solo actúe *a posteriori*, reparando, sino que se anticipe preventivamente a los siniestros e implique al asegurado en la reducción de los riesgos es realmente un cambio muy positivo.

Lo expuesto hasta aquí es solo una aproximación general que sirve para ilustrar la revolución que ha experimentado el contrato de seguro y su mercado. Los cambios que afectan al seguro afectan a todas sus fases y también en su distribución, dado que han surgido nuevos agentes y formas de venta a las que la vigente regulación nacional (Real Decreto-Ley 3/2020, de 4 de febrero) no presta la atención debida. De todas formas, la digitalización del seguro es un proceso en marcha que convive con la contratación tradicional. Quizás por ello sea aún más interesante tratar de ver qué efectos produce y cuáles son las ventajas del nuevo entorno, los potenciales riesgos que comporta y su encaje en la legislación nacional anterior a esta transformación, y en la normativa internacional, que, en contraste con nuestro ordenamiento, sí ha afrontado directamente los problemas que

estos cambios generan. de todo ello pretendemos dar cuenta en estas páginas, aunque sea brevemente.

2. ESTADO ACTUAL DE LA DIGITALIZACIÓN Y DEL USO DE IA EN EL SEGURO

El proceso de digitalización en el ámbito de los seguros es completamente natural y era totalmente esperable, dado que ya había comenzado en otros sectores del mercado financiero —la banca y la bolsa— y se aceleró en el contexto de la pandemia de COVID-19, debido a la necesidad abrupta de sustituir las interacciones presenciales por las relaciones y al surgimiento de cambios en los modos de vida y nuevas necesidades a las que el seguro había de dar respuesta.

La digitalización afecta al seguro en múltiples aspectos: desde la publicidad hasta el proceso de contratación, pasando por la información precontractual y contractual, las comunicaciones entre los clientes y las compañías, la presentación de quejas y reclamaciones y las liquidaciones de siniestros. En principio, en todos estos elementos es perfectamente posible sustituir el trato personal por una relación puramente *online*. La forma de contratar, en definitiva, se simplifica y se acomoda al modo de vida actual, en el que un creciente número de actividades se realizan cotidianamente a través del teléfono móvil o del ordenador. La digitalización del seguro se convierte, así, en un modo de acercar un contrato poco querido y árido en una figura jurídica algo más próxima y cercana, al menos para los jóvenes.

Esta tendencia de cambio también tiene sus costes, dado que, si bien facilita la asunción de seguros por jóvenes, también comporta el incremento del riesgo de exclusión de las personas vulnerables por razones de edad o por falta de recursos —entre otros, los de naturaleza digital—. Así, la autoridad de supervisión del seguro europea (EIOPA, 2024: 14) ha puesto de manifiesto que los canales de distribución digital son utilizados con mayor frecuencia para contratar seguros por clientes jóvenes de sexo masculino, con un alto nivel educativo y elevados ingresos, y ha llamado la atención sobre el alto porcentaje de clientes con ingresos bajos que durante los últimos dos años no han adquirido productos en el mercado asegurador de los últimos dos años, dato que podría indicar la existencia de una «brecha de protección» («protection gap») que afecta especialmente a este grupo de clientes. En esta línea, la EIOPA también ha señalado que la población de mayor edad (55 o más años) utiliza con menor frecuencia los canales digitales (23 % de los seguros contratados) en comparación con los clientes de entre 25 y 39 (38 %), por lo que podrían verse más impactados por la digitalización del sector asegurador.

En todo caso, el último informe de EIOPA (2024: 8) también subraya que la plena digitalización está lejos de ser una realidad, y que su implantación aún es desigual. De las 209 aseguradoras de toda Europa que participaron en la encuesta que sirvió como base para la elaboración del informe, el 77 % declararon que habían puesto en marcha una estrategia de digitalización en marcha —ya fuera actuando en solitario en el mercado

o integrándose en un grupo empresarial—, mientras que el 23% de ellas reconocieron seguir utilizando los medios tradicionales de contratación y comunicación con la clientela. Por otro lado, es reseñable que la mayoría de las empresas encuestadas expresaran que su principal motivo para embarcarse en este proceso era, precisamente, mejorar las relaciones con la clientela.

No es casual que el ámbito la comunicación con los futuros clientes y los ya asegurados sea el campo en el que se ha intensificado la aplicación de los distintos usos de IA. Así, centrando la atención en nuestro mercado —que en este punto no es diferente al mercado europeo—, UNESPA (junto con AEFI, 2023: 16) ha identificado las siguientes tendencias: *i)* el análisis del sentimiento del centro de llamadas, el examen de las causas de las rutas, los guiones dinámicos y la asignación de agentes; *ii)* el autoservicio del cliente a través de múltiples canales mediante la utilización del procesamiento del lenguaje natural (PNL), el reconocimiento de voz, los mapas ontológicos de seguros y los *chatbots*; *iii)* la automatización de procesos robóticos (RPA), incluido el reconocimiento óptico de caracteres (OCR), para extraer información de los documentos (por ejemplo, correo electrónico con preguntas sobre reclamaciones, etc.) y dirigirlos al departamento correcto; y *iv)* la promoción y la venta cruzada de servicios relacionados basados en los datos del consumidor procedentes de los sistemas que gestionan las relaciones con el cliente.

La propia IA favorece también la personalización de la publicidad de cualquier producto o servicio, incluidos los seguros. Como advierte la Comisión Europea (2024: 5), «[…] las empresas pueden recopilar información sobre nuestras preferencias e intereses en las páginas web que visitamos, saber a qué le damos "me gusta", los enlaces que seguimos, así como la información personal que proporcionamos sobre nosotros mismos, como nuestra edad o nuestro lugar de residencia. Las plataformas utilizan algoritmos e inteligencia artificial en la elaboración de perfiles a partir de los datos para decidir qué anuncios mostrarnos y tener el mayor impacto posible en cada uno de nosotros. Algunas plataformas en línea obtienen dinero cada vez que compramos productos gracias a estos anuncios». Sin embargo, hay límites legales: el artículo 26.3 del Reglamento de Servicios Digitales[1] contempla la posibilidad de que la publicidad se base en la elaboración de perfiles, pero excluye aquella que se nutre de datos confidenciales de los asegurados. Además, el artículo 28 prohíbe a las plataformas que hagan publicidad basada en la elaboración de perfiles cuando dispongan de medios razonables para saber que el destinatario es un menor.

También está en el punto de mira de los reguladores y del propio Reglamento de Servicios Digitales el uso de los llamados *dark patterns* para influir en las decisiones del destinatario de la publicidad. Como recuerda EIOPA (*Consumer Trends Report*, 2023: 37 y 2024: 44-45), este término hace referencia a los elementos del diseño de un interfaz que pueden inducir a un consumidor a tomar una decisión (contratar o ceder datos personales) que interesa al proveedor y que aquel no habría adoptado si hubiese sido informado correctamente y se le hubiesen ofrecido alternativas viables. Un ejemplo en el ámbito del seguro es la estrategia orientada a incitar al consumidor a la suscripción de una póliza sin que este comprenda aspectos importantes de las condiciones generales de la cobertura utilizando el reclamo de un descuento temporal de la prima que perdería en caso de demorarse en tomar

la decisión. Las vías para influir en la voluntad del consumidor son muy diversas: desde el oscurecimiento de la información relevante para promover una opción específica hasta la predeterminación de la opción más beneficiosa para el proveedor, pasando por la manipulación emocional del diseño de producto, la obligación de aceptar *cookies*, etc. Aunque estas prácticas también se llevan a cabo en la contratación con presencia física de las partes, son más comunes en las plataformas digitales, incluidos los sitios web de medios sociales, compras y aplicaciones móviles; con frecuencia, los consumidores no son conscientes del modo en que afectan a sus decisiones (*vid.*, más ampliamente, los ejemplos encontrados por EIOPA en el mercado y que recoge en «Dark patterns in insurance: practices that exploit consumer biases» para advertir de los riesgos que comportan estas medidas para los usuarios).

El artículo 25.1 del Reglamento de Servicios Digitales establece en estos términos la prohibición de este tipo de prácticas: «Los proveedores de plataformas en línea no diseñarán, organizarán o gestionarán sus interfaces en línea de forma que engañen o manipulen a los destinatarios de su servicio o de forma que distorsionen o menoscaben materialmente la capacidad de los destinatarios de su servicio para tomar decisiones libres e informadas». Por su parte, el artículo art. 5.1. a) del Reglamento de IA prohíbe directamente las que consisten en «la introducción en el mercado, la puesta en servicio o la utilización de un sistema de IA que se sirva de técnicas subliminales que trasciendan la conciencia de una persona o de técnicas deliberadamente manipuladoras o engañosas con el objetivo o el efecto de alterar de manera sustancial el comportamiento de una persona o un colectivo de personas, mermando de manera apreciable su capacidad para tomar una decisión informada y haciendo que tomen una decisión que de otro modo no habrían tomado, de un modo que provoque, o sea razonablemente probable que provoque, perjuicios considerables a esa persona, a otra persona o a un colectivo de personas».

En todo caso, la preocupación de EIOPA (*ibid.*) persiste, especialmente en relación con las personas vulnerables, a las que el Reglamento de IA brinda una protección reforzada frente al uso de métodos que puedan explotar su vulnerabilidad (cfr. art. 5.1.b): debido a la forma en que se presenta la información o a la utilización de determinadas técnicas de venta agresivas, los consumidores con un bajo nivel de conocimientos financieros podrían comprar productos de inversión basados en seguros (IBIPs) sin entender en qué están invirtiendo y/o podrían tener dificultades para comprender la cobertura que están contratando. Por lo que respecta a la privacidad de los datos de los consumidores, EIOPA señala que el 17 % de los consumidores de la UE declararon que se les pedían datos personales innecesarios para comparar ofertas de pólizas de seguros. Frente a la oscuridad de estos patrones, se exige transparencia, habiendo autores (De Novaes Da Silva, 2024) que proponen el *legal design* como antídoto.

La IA permite, por otro lado, precisar mucho más en la selección de riesgos, lo que, en principio, supone que el precio solicitado se ajuste más al riesgo realmente sufrido por cada asegurado, entre otras razones porque, como quedó dicho, además de los datos que utilizaban tradicionalmente para fijar la prima, hoy en día las aseguradoras disponen de otras vías para obtener información de los asegurados. En algunos casos, estos datos los suministra el

asegurado voluntariamente a través de dispositivos incorporados a su automóvil, su hogar o que él mismo lleva en su propia muñeca, y en otros lo hace de manera involuntaria: por ejemplo, cuando deja un rastro al visitar la web de la aseguradora. Ahora bien, en la práctica esto no siempre supone que el precio que se pide a cada cliente sea el más ajustado a su riesgo, dado que las compañías recurren a prácticas como la optimización de precios, que tienden a tratar de obtener el mayor beneficio posible de cada cliente, y utilizan de factores para su fijación que no guardan una relación de causalidad con el riesgo. Con todo, esto no debe llevar a pensar que la discriminación de los asegurados es, sin más, una consecuencia de la aplicación de la IA al seguro. Se trata de un peligro que ha existido siempre porque también en la valoración de los riesgos hecha al modo tradicional —es decir, por un ser humano— las empresas también utilizan estos recursos para tergiversar los precios. Sencillamente, ahora pueden hacerlo de manera cada vez más masiva —a veces sin que la propia aseguradora sea consciente de ello—, posibilidad que depende la complejidad de los algoritmos que utilice. La legislación española todavía no ha incorporado una medida ya prevista en otros ordenamientos más avanzados —tal es el caso de Bélgica, que la ha incluido en la Ley relativa a los seguros, de 4 de abril de 2014 (§§ 45 y 46 en relación con el § 43)— orientada contrarrestar los efectos perniciosos de aquellas prácticas, a saber, la obligación impuesta a las compañías para que hagan públicos (en Bélgica deben hacerlo en su web) no los algoritmos empleados para fijar los precios —cuyo conocimiento poco aporta al consumidor medio—, sino algo mucho más útil y comprensible: los factores de riesgo que utilizan para calcular las primas. Esta es la manera más eficaz para que un asegurado pueda comprobar si los factores que se han utilizado en su caso se corresponden con el riesgo que ha trasladado a la aseguradora o si, por el contrario, se ha utilizado algún factor de naturaleza subjetiva.

En el plano de la gestión de siniestros, AEFI y UNESPA (2023: 16) señalan que la IA sirve para: *i)* mejorar el análisis de fraude (puntuación de siniestros, detección de anomalías, análisis de redes sociales y modelización del comportamiento); *ii)* estimar el valor de las pérdidas, en particular en siniestros de alta frecuencia; iii) evaluar los costes de reparación en seguros de hogar, locales comerciales y automóviles mediante el reconocimiento de imágenes; iv) segmentar de manera automatizada las reclamaciones por tipo y complejidad; y *v)* automatizar la verificación y el pago de facturas.

Además de a la IA, debemos hacer referencia al rol que la *Distributed Ledger Technology* (DLT, más conocido como *blockchain*) está desempeñando en el ámbito del seguro atendiendo a los datos que aportan tanto la EIOPA (2024: 30 ss.) como AEFI-UNESPA (2023: 19). Esta tecnología comprende una base de datos integrada en una cadena y protegida por una clave criptográfica. Así, los usuarios pueden consultar los datos almacenados en la base sin que su contenido pueda ser alterado (aunque cabe la posibilidad de que añadan información a la cadena). De esta forma, se alcanza el objetivo de disponer de información veraz e inalterable. Dado que esta tecnología permite acceder a la información guardada, en el ámbito de los seguros se utiliza, sobre todo, en el proceso de gestión y tramitación de reclamaciones, así como en la detección de fraudes. Asimismo, es el punto de partida para configurar *smart contracts*, es decir, contratos cuya información se encuen-

tra integrada en el sistema *blockchain* y cuya ejecución es automática si se cumplen las condiciones o requisitos especificados en su programación. De hecho, los ya mencionados seguros paramétricos (que analizaremos más abajo) normalmente son diseñados a partir de esta tecnología, que también se usa en determinados seguros «basados en el uso» (*usage-based insurance*), aunque se trata de el aspecto menos interesante de esta modalidad, a la que también prestamos atención más adelante. De todas formas, en Europa, el uso de DLT en el seguro es aún incipiente en Europa. Según datos de EIOPA (2024: 30-31), solo un 15 % de las aseguradoras encuestadas había recurrido a esta tecnología y la mitad de los casos de uso reportados estaban aún en fase de pruebas.

3. LA POSTURA DEL REGLAMENTO DE IA EN RELACIÓN CON EL SEGURO

En relación con la implantación de la IA en el ámbito del seguro, durante la prolongada elaboración del Reglamento de IA de la UE se discutió extensamente si, considerando algunos de sus usos, los sistemas de IA deberían calificarse como «de alto riesgo». Haciéndose eco de la pretensión de las aseguradoras que operan en España, la UNESPA manifestó que ninguna de estas actividades debería ser considerada como tal con el fin de liberarlas de las obligaciones impuestas a los sistemas de algo riesgo (AEFI-UNESPA, 2023: 17). Aunque la Comisión[2] secundaba este planteamiento, en su réplica a la propuesta de este órgano el Consejo incluyó entre los sistemas de IA de alto riesgo los que se utilizan para evaluar los riesgos y fijar los precios a la luz de las características de las personas físicas en los seguros de vida y de salud, excepción hecha de los sistemas de IA puestos en servicio por las microempresas y las pequeñas empresas[3]. En la fase posterior de formulación de enmiendas por el Parlamento (junio de 2023), se presentó la enmienda n.º 723, cuyo objeto era, precisamente, incorporar los sistemas utilizados en los seguros de salud y de vida para tomar decisiones o influir materialmente en las determinaciones relativas a la selección de las personas físicas al conjunto de sistemas de IA de alto riesgo[4]. En la enmienda n.º 67 al considerando 37, el Parlamento justificó su criterio en estos términos: «Los sistemas de IA destinados a ser utilizados para tomar decisiones o influir materialmente en las decisiones sobre la elegibilidad de las personas físicas para el seguro de salud y de vida también pueden tener un impacto significativo en los medios de subsistencia de las personas y pueden vulnerar sus derechos fundamentales, por ejemplo, limitando el acceso a la asistencia sanitaria o perpetuando la discriminación basada en características personales. Por lo tanto, estos sistemas deben clasificarse como de alto riesgo».

Finalmente, el texto definitivo del Reglamento (cuya publicación en el DOUE, a la fecha de cierre de este trabajo, está prevista para el 12 de julio) incluyó entre los sistemas de IA de alto riesgo los «[…] destinados a ser utilizados para la evaluación de riesgos y la fijación de precios en relación con las personas físicas en el caso de los seguros de vida y de salud» (Anexo III. 5.c), dado que «[…] pueden afectar de un modo considerable a los medios de subsistencia de las personas y, si no se diseñan, desarrollan y utilizan debidamente, pueden vulnerar sus derechos fundamentales y pueden tener graves consecuencias

para la vida y la salud de las personas, como la exclusión financiera y la discriminación» (considerando 58). Obsérvese que no se hace salvedad alguna para favorecer a las pequeñas empresas.

Esta calificación (sistemas de alto riesgo) es coherente con el texto del Anexo III.5.b, que atribuye semejante condición a «[…] los sistemas de IA usados para evaluar la calificación crediticia o solvencia de las personas físicas», ya que, como señala el considerando 58, «[…] deciden si dichas personas pueden acceder a recursos financieros o servicios esenciales como la vivienda, la electricidad y los servicios de telecomunicaciones. Los sistemas de IA usados con estos fines pueden discriminar a determinadas personas o grupos y perpetuar patrones históricos de discriminación, como por motivos de origen racial o étnico, género, discapacidad, edad u orientación sexual, o generar nuevas formas de discriminación». Es importante hacer notar que se consideran de alto riesgo los sistemas de IA destinados a determinar la calificación crediticia de una persona para su eventual acceso a los recursos elementales que garanticen su bienestar. Ahora bien, debemos añadir que sería inadmisible que esa calificación crediticia se utilizara como factor para determinar la concesión o no de seguros de cualquier tipo, así como su precio, salvo en aquellos seguros en los que el riesgo cubierto sí tenga una relación directa con la solvencia del asegurado, como los de caución. Baste recordar el problema, revelado por *Consumer Reports* en 2015 y dado a conocer por O'Neil (2016: 164 ss.), que planteó en Estados Unidos el uso de la calificación crediticia de los potenciales asegurados en la concesión de seguros de automóviles para determinar la decisión de la aseguradora sobre la asunción de la cobertura, método que, a su vez, propiciaba la exigencia del pago de primas centradas antes en este último aspecto que en el riesgo de accidente del vehículo asegurado. No puede confundirse el riesgo asegurado (por ejemplo, el robo, el incendio o la responsabilidad civil) con el riesgo de impago de primas.

Frente a los casos expuestos, el Reglamento (Anexo III.5.b *in fine*) excepciona de la consideración de sistema de «alto riesgo» a los que se utilizan para «detectar fraudes financieros». Consideramos, por ello, que habría que entender que están fuera del ámbito de los sistemas de «alto riesgo» aquellos que habitualmente usan las aseguradoras para evitar la comisión de fraudes en el seguro. En materia de solvencia, el arriba citado considerando 58 señala que esta exclusión se extiende a los sistemas de IA utilizados «[…] para calcular los requisitos de capital de las entidades de crédito y las empresas de seguros».

4. NUEVAS FORMAS DE ASEGURAMIENTO: SEGUROS *USAGE-BASED* TELEMÁTICOS, *ON-DEMAND* Y PARAMÉTRICOS

La introducción de la IA y de las tecnologías a ella asociadas en el campo de los seguros seguro ha supuesto un antes y un después. No hay duda de que esta incorporación es ventajosa, dado que favorece el diseño de productos más adecuados a las necesidades de los potenciales asegurados, entre ellos los seguros basados en el uso (como los *pay-as-you-drive* y *pay-as-you-live*, por citar los más difundidos), los denominados *on demand*, que se activan y desactivan cuando lo decide el asegurado y le permiten no pagar la prima por los

períodos de ausencia de cobertura, y los paramétricos o basados en índices. Para ampliar la información sobre estas modalidades, remito a Muñoz Paredes (2022 y 2023).

4.1. SEGUROS BASADOS EN EL USO Y *ON-DEMAND*

Una de las manifestaciones más claras de la aplicación de la IA en el seguro es la creciente oferta de seguros telemáticos, entendiendo por tales aquellos en los que el riesgo se controla durante toda la vigencia del contrato a través de un dispositivo incorporado al bien asegurado o portado por el propio tomador del contrato. Este tipo de seguros son una aplicación del llamado «internet de las cosas» (*Internet of Things*, IoT), expresión que la autoridad supervisora del seguro de la UE (EIOPA 2024: 56) define como «[…] la conexión en red de dispositivos telemáticos, vehículos, edificios y otros objetos dotados de electrónica, *software*, sensores, *wearables* y conectividad de red que permiten a estos objetos (a) recoger e intercambiar datos y (b) enviar, recibir y ejecutar órdenes».

Cabe señalar, por otra parte, que los seguros telemáticos o monitorizados son un subtipo de los seguros basados en el uso o *usage-based* (UBI). Su denominación da cuenta del rasgo que los distingue del seguro ordinario. En este, la cuantía de la prima normalmente no se hace depender del mayor o menor uso que se va a dar al bien ni, en general, del comportamiento del asegurado en relación con el riesgo cubierto durante el curso del contrato. En cambio, los seguros *usage-based* (*usage-based insurance*, UBI) se caracterizan precisamente porque en ellos la prima aumenta o se reduce en función del uso efectivo del bien —si son seguros de cosas— o del comportamiento del asegurado —si lo son de personas—. En un seguro ordinario, el uso efectivo que se haga del bien durante el curso del contrato no influye normalmente en la determinación de la prima (por ejemplo, el asegurado no obtiene descuentos por usar poco el automóvil o por conducir bien), mientras que en los seguros basados en el uso ese es justamente el factor determinante para que el precio del seguro suba o baje incluso durante un solo periodo de vida de la póliza.

Ese uso efectivo del bien (o el comportamiento del asegurado en relación con el riesgo) puede controlarse a través de diversas vías a efectos de establecer su influencia en la prima (y otras condiciones del contrato). Así, inicialmente, en los seguros basados en el uso más primitivos, el control del riesgo se realizaba inicialmente sobre la base de factores como el consumo de carburante o las distancias recorridas. Por tanto, en este proceso de evaluación se consideraba que cuanto menor era el uso, tanto menor era menor riesgo. Posteriormente, la aplicación de nuevas tecnologías al seguro abrió la posibilidad de controlar el riesgo a través de dispositivos introducidos en el propio automóvil que permitían medir otros parámetros, entre ellos la brusquedad en la conducción. De este modo, como una derivación de los primeros seguros basados en el uso, nacieron los seguros telemáticos, aunque siguen existiendo seguros basados en el uso que no son telemáticos (por ejemplo, los más elementales, basados en la distancia recorrida).

Pueden distinguir varios tipos de seguros de vehículos basados en el uso. Los más difundidos los siguientes:

i) *Pay-Per-Mile* (PPM): en él, la prima depende total o parcialmente de las distancias recorridas, que son controladas sin recurrir a la utilización de dispositivos telemáticos.

ii) *Pay-As-You-Drive* (PAYD) or *Pay-As-You-Go* (PAYG): en este caso, la prima se fija atendiendo a la distancia recorrida por el automóvil; en contraste con el anterior modelo contractual, en este tipo de seguro la distancia queda registrada en un dispositivo telemático, por regla general un *smartphone*.

iii) *Pay-How-You-Drive* (PHYD): se trata de la modalidad más compleja, dado que la prima se calcula tomando en consideración tanto la distancia recorrida como el estilo de conducción —caracterizado por variables tales como la velocidad, el tipo de carreteras y las franjas horarias en las que más se usa el vehículo—, variables reflejadas en los datos recogidos por un dispositivo incorporado al vehículo.

Los dos primeros tienen un rasgo en común: están basados en la distancia recorrida, aunque el segundo es telemático y el primero no. El tercero se diferencia de los anteriores por el hecho de que combina distancia recorrida y comportamiento al volante. Al igual que el segundo, siempre es telemático.

Por otra parte, la aplicación del nuevo concepto del control de riesgo —especialmente cuando este empezó a realizarse a través de dispositivos telemáticos— se extendió a otros ámbitos en los que tenía una utilidad análoga: dada su similitud con los seguros del sector automotriz, el primer ramo en el que se verificó esta ampliación fue el del hogar, ampliación facilitada por la difusión de las instalaciones domóticas en las viviendas. Más allá de la cobertura de riesgos en las cosas, también en los seguros de personas se percibió la utilidad potencial del uso de los datos que pudieran recopilarse automáticamente en las pulseras de actividad (y, posteriormente, en los *smartwatches* o los *smartphones*) para controlar la actividad física del asegurado. La idea básica era incentivar el autocuidado físico del asegurado. Las aseguradoras crearon aplicaciones para que los tomadores del seguro introdujeran datos de diversa índole, entre ellos los relativos a la alimentación, con el fin de implicarlos en la reducción del riesgo. Dado que no hay mejor método para conseguir este objetivo que la observancia continuada de las pautas sugeridas por las compañías, las nuevas tecnologías son un aliado perfecto de las aseguradoras.

Los seguros telemáticos se han difundido en los ramos del automóvil, el hogar, la salud y la vida, si bien su aplicación está más extendida en el primer sector citado. De acuerdo con los datos de EIOPA (2024: 33), de las 209 aseguradoras que operan en Europa y que participaron en la encuesta que sirvió de base para la elaboración de ese informe, el 17 % ofrece un seguro de vehículos vinculado al internet de las cosas, mientras este porcentaje se reduce al 10 % en hogar y al 7 % en salud. No obstante, la propia EIOPA señala que es previsible que aumente la difusión los seguros ligados al uso de dispositivos, pues cerca del doble de las empresas que participaron en esa encuesta manifestaron su intención de añadirlos a su cartera en el plazo de tres años. El informe no contiene datos relativos al seguro de vida, pese a que en los países del mercado europeo —sin ir más lejos, España— hay oferta de seguros telemáticos de vida. De hecho, las distribuidoras que en nuestro país

comercializan seguros de salud que prevén el seguimiento telemático del asegurado para controlar sus hábitos —y, de esa manera, le inducen a reducir el riesgo de empeoramiento de su salud— suelen ofertar seguros de vida con condiciones similares. Es claro que este tipo de control persigue el mismo fin en los dos casos y resulta útil en ambos ramos. Por ejemplo, Generali oferta sus seguros de salud y vida vinculados al programa de bienestar «Vitality», e incluso algunos proveedores (entre ellos, la correduría de El Corte Inglés, con su programa «Vida movida») solo comercializan el seguro de vida.

Las ventajas de este tipo de seguros son evidentes. En el caso de los seguros de vehículos, pueden resumirse del siguiente modo: *i)* permiten que cada asegurado pague la prima que le corresponde, en la medida en que esta se adapta al uso real del vehículo; *ii)* incitan a reducir el riesgo (aunque en los PPM a veces es sencillamente difícil porque el asegurado no rebasa el kilometraje pactado en la póliza); *iii)* dado que miden el riesgo real, pueden servir para asegurar a conductores tradicionalmente calificados como de «alto riesgo», entre ellos los jóvenes, que ahora disponen de un medio para demostrar que su conducción no es temeraria; *iv)* la utilidad de los dispositivos para el asegurado no se agota en el control del riesgo para la determinación de la prima. En caso de accidente, el dispositivo avisa a los servicios de emergencias e indica la ubicación del vehículo; si en el accidente hay terceros afectados, la grabación puede servir para determinar si el asegurado es responsable del siniestro; en caso de robo, sirve para localizar el coche; asimismo, permite controlar el fraude. Por lo que respecta a los seguros de hogar, vida o salud telemáticos, las principales ventajas son la implicación del asegurado en la reducción del riesgo y la función preventiva del seguro. Lógicamente, ambas utilidades se despliegan en función de la naturaleza de los bienes asegurados y los riesgos cubiertos (incendio, robo, inundación y otros riesgos en el caso de los seguros del hogar y diversas contingencias que suponen una amenaza para la salud del asegurado en los seguros de vida y salud).

Con todo, este tipo de seguros también tiene sus inconvenientes. Los contratos no siempre contemplan la obligación de la compañía de reducir la prima a medida que se reduce el riesgo, un problema que afecta a los seguros telemáticos de todos los ramos. En efecto, en ocasiones, la única ventaja económica ofrecida al asegurado es su participación en un programa de retos con premios cuyo valor es meramente promocional, que la aseguradora puede cancelar cuando lo estime conveniente. Ahora bien, entendemos que el hecho de que el asegurado acepte integrarse en un programa de este tipo no supone una renuncia a su facultad para solicitar la reducción de primas prevista en el artículo 13 LCS. Por otro lado, el asegurado no siempre es consciente de que las agravaciones del riesgo pueden dar lugar a aumentos de prima y de que la aseguradora puede conocerlas fácilmente cuando el seguro es telemático.

En esta misma línea argumental, cabe señalar que una de las mayores preocupaciones que plantean los seguros telemáticos es precisamente el potencial menoscabo de la intimidad del asegurado, dado que los datos recopilados por el dispositivo son personales y, en algunos casos, sensibles. Esto es muy evidente en los seguros de vida o salud —en los que el asegurado comparte con la aseguradora numerosos datos relativos a su modo de vida, su estado físico y las actividades deportivas que realiza—, pero también es claro en el seguro

de hogar —pues la aseguradora puede recabar datos de la vida diaria del asegurado— y en el de automóvil —dado que el dispositivo no solo está programado para registrar todos los recorridos hechos por el vehículo y determinar qué sitios frecuenta el asegurado y a qué hora, sino también para grabar imágenes y/o sonidos—. Conscientes de la posibilidad de que se produzcan estos efectos indeseables, algunas compañías aclaran que el dispositivo les envía los datos de los viajes del asegurado de forma agregada para ajustar la prima, por lo que solo el asegurado conocerá el detalle de los viajes (así, Generali en su primitivo seguro «Pago como conduzco», que ya no comercializa) o que los datos personales de conducción se alojan en otra compañía y son separados de los datos del seguro tradicionales (así, la alemana Huk Coburg, *vid.,* John, 2022). Otras compañías informan al cliente de que el dispositivo que instalan no graba las conversaciones que el conductor mantiene en el interior del coche y que, además, el asegurado puede desconectar la grabación cuando quiera y viajar en «modo privado», es decir, sin que aparezca la ubicación del coche en la aplicación (tal es el caso de Hello Auto: https://helloauto.com/privacidad).

Es importante señalar, por otro lado, que las aseguradoras son uno de los blancos más sensibles a los ciberataques, en la medida en que almacenan todo tipo de datos, personales o no, de sus asegurados. Si, además, el asegurado se conecta con la compañía por vía telemática, el riesgo es mayor. En el caso de que la empresa ceda los datos u opere a través de otra empresa en la cobertura de este tipo de seguros, una práctica bastante frecuente, el riesgo se incrementa. Por ello, las entidades aseguradoras que operan en la UE están actualmente sometidas a las medidas de prevención que el Reglamento DORA[5] impone a todas las empresas del sector financiero.

Junto a los tradicionales seguros basados en el uso, la aplicación de las nuevas tecnologías al seguro ha facilitado la (más reciente) difusión en el mercado asegurador de los seguros *on demand*, que se toman y dejan cuando el asegurado lo decida (de ahí que se conozcan también como seguros *on-off*). Este tipo de seguro adopta diferentes modalidades. Así, en el ramo de auto, son muy habituales los seguros por días, que se contratan al alquilar coches eléctricos a través de una plataforma de economía colaborativa o, sencillamente, por mediación de una empresa de alquiler. En otros casos, el seguro es anual, pero el asegurado solo paga por los períodos en los que lo «activa» y no por el tiempo durante el que lo «desactiva». Lógicamente, la prima establecida para el periodo en el que el seguro está «activo» es más alta de lo normal, pero este mecanismo puede compensar al asegurado en función del uso que haga de la cobertura. Los seguros *on-demand* se han difundido para todo tipo de bienes de uso: además del ramo de los automóviles, que es el ejemplo paradigmático, también se aplican a las motos, las bicicletas y a todo tipo de dispositivos electrónicos, y se comercializan en ramos tan variados como el transporte, los accidentes o la salud.

Los seguros *on-demand* son totalmente digitales y se contratan rápidamente a través del teléfono móvil, medio que también se usa para comunicar a la aseguradora el posterior «apagado» y eventual nuevo «encendido» del seguro en función de las variaciones de la exposición al riesgo cubierto. Si se cubren los derivados del uso de una cosa, bastará activar el seguro en ciertas situaciones (generalmente cuando aquella se esté utilizando). Esta

fórmula guarda, por tanto, una evidente similitud con la de los seguros *usage-based* más conocidos: por ejemplo, en el sector del seguro de auto, los *pay-as/how-you-drive* o los *pay-per-mile*. Las principales características de los seguros *on-demand* son las siguientes: *i)* están concebidos para lapsos breves de tiempo o para periodos con intermitencias, dependiendo del uso que va a hacer el usuario del bien o servicio cuyos riesgos cubren; *ii)* el tomador (el usuario del bien o servicio) es el que decide voluntariamente su «activación» y «desactivación»; y *iii)* son completamente digitales.

Como señala Ruiz Echauri (2020), en contraste con otros países, entre ellos Brasil, donde la autoridad de control ha dictado recientemente una circular relativa a los seguros con vigencia reducida o con período intermitente[6], no contamos con un pronunciamiento semejante en nuestro sistema. No obstante, hay que entender que, en principio, este tipo de seguro puede contratarse válidamente, dada la flexibilidad que ofrece la Ley de Contrato de Seguro. Compartimos esta opinión: en la medida en que la modalidad de seguro beneficie al asegurado-consumidor más que las previsiones de la LCS, está amparada por el art. 2 LCS. Cuestión distinta son las dificultades con que las pueda encontrarse el asegurador (y que el propio Ruiz Echauri señala con detenimiento) a la hora de entender cumplidas sus obligaciones de información al tomador en la fase precontractual y de prueba del contrato. Las disposiciones de nuestro ordenamiento relativas a este punto aún no se adaptan bien a la contratación digital ajena al papel y eminentemente rápida.

A semejanza de los seguros *usage-based*, la utilidad de los seguros *on-demand* se puso de manifiesto de manera muy notoria en los confinamientos que en mayor o menor grado se decretaron en todos los países para frenar los contagios durante la pandemia de COVID-19 (2020-2021). Muchos asegurados de auto que contaban con seguros ordinarios —que cubrían todos los riesgos o la responsabilidad civil derivada del uso— se vieron obligados a pagar la prima entera de un bien que no podían usar, por lo que la posibilidad de acogerse una cobertura que discriminara si había exposición al riesgo o no fue una fórmula ideal para satisfacer sus necesidades en aquella coyuntura. Por otro lado, la modalidad *usage-based* es el mecanismo más adecuado para asegurar las flotas de alquiler de vehículos debido a que el cambio de asegurados en relación con un solo automóvil es constante. Considerando la tendencia creciente de los consumidores a optar por el uso compartido de bienes —el automóvil es un claro exponente de ello—, es previsible que se produzca una expansión del seguro *on-demand*. Ahora bien, no todo son luces en la realidad práctica de este seguro. Las aseguradoras de Estados Unidos (NAIC, 2023) han detectado ya un riesgo de fraude, consistente en que las compañías solo «activan» el seguro para reclamar el pago de un siniestro.

4.2. SEGUROS PARAMÉTRICOS

El paramétrico (*parametric* o *index-based insurance*) es un seguro de daños que se caracteriza por que la aseguradora se obliga a pagar una suma acordada en el contrato, monto que corresponde a los daños que previsiblemente se le derivarían al asegurado de producirse un evento de una cierta magnitud, siempre que este se produzca, atendiendo (normal-

mente) a la información procedente de un tercero independiente, y, aunque no siempre, con independencia del daño efectivamente sufrido por el asegurado. Por tanto, a diferencia del seguro de daños tradicional, no cubre necesariamente los daños reales derivados del siniestro ni un riesgo (incendio, terremoto, etc.) de forma genérica, pues solo hay siniestro si el evento alcanza la envergadura determinada en el contrato. Esa magnitud, denominada «índice» o «parámetro», es el elemento diferenciador de esta modalidad y da nombre al propio seguro.

Veamos un ejemplo: la aseguradora se obliga a pagar a un asegurado 5 millones de euros si se produce un terremoto de magnitud 7.0 en una zona geográfica identificada en el contrato en la que el asegurado tiene algún tipo de interés. Normalmente, es un tercero independiente de las partes (por ejemplo, el Centro Sismológico Euromediterráneo) el que determina que efectivamente se ha alcanzado ese umbral.

De manera simplificada, se entiende que la sola ocurrencia de ese índice o parámetro desencadena el pago. Por eso, el índice se denomina gráficamente en inglés *parametric trigger*, *index trigger* o simplemente *trigger*. También se afirma (*vid*., por ejemplo, Elliott, 2019) que este seguro es, en esencia, un *if-then*: si, en nuestro caso, el terremoto alcanza o supera la magnitud pactada, entonces se pagará sin más la cantidad estipulada.

Esta modalidad de aseguramiento se utiliza para la cobertura de riesgos de producción de catástrofes naturales de todo tipo (huracanes, terremotos, inundaciones o sequías), contingencias en las que es particularmente necesario un pronto pago de los daños para la recuperación de los lugares afectados. En economías poco desarrolladas y carentes de un mercado de seguros estable, el seguro paramétrico es la principal fórmula que, con el apoyo de los poderes públicos y organizaciones como Naciones Unidas y el Banco Mundial, se está utilizando desde hace décadas para proporcionar cobertura de los riesgos que amenazan a las producciones agrícolas. Actualmente, se utiliza en todo tipo de economías para cubrir cualesquiera riesgos generalmente desatendidos por el seguro tradicional de daños (tal sería el caso, por ejemplo, de las pérdidas derivadas del cierre de un establecimiento que no haya sido provocado por un daño material en el mismo, como el cierre por pandemia), y para colmar los vacíos que deja abiertos el seguro de daños común como las franquicias previstas en este.

El diseño de los seguros paramétricos es aparentemente sencillo y flexible (esto es, son personalizables) y generalmente su contratación tiene lugar a través de medios digitales. Precisamente, el mecanismo de evento desencadenante-pago permite que haya seguros paramétricos configurados como *smart contracts*; en este caso, se automatiza y objetiva de manera completa todo el funcionamiento del seguro, incluido el pago de siniestros.

4.2.1. Estructura y funcionamiento

El seguro paramétrico siempre se construye sobre la base de dos elementos: *i)* el evento desencadenante, y *ii)* el mecanismo de pago, que «se activa» cuando se produce el evento.

i) Evento desencadenante

El evento desencadenante tiene que ser de una magnitud concreta dentro de una escala conocida (por ejemplo, un terremoto de nivel 6.0 en la escala Richter). Además, dado que el seguro no es una apuesta, ese parámetro o índice objetivo ha de estar relacionado con el riesgo concreto del asegurado, dicho esto en el sentido de que, si se alcanza ese parámetro (el nivel 6.0 en nuestro ejemplo), es previsible que el asegurado sufra unos daños cuyo valor puede determinarse de antemano con cierto grado de precisión.

A tal efecto, es preciso que la póliza determine el área concreta de influencia del evento que va a cubrir el seguro (algunos condicionados generales usados actualmente la denominan «área de medición»), identifique los bienes que el asegurado tenga en la zona y declare su valor. Sobre esa base, se fija la cuantía de la indemnización.

Así, esta cuantía (que, como se ha dicho, es una cantidad fija pactada de antemano, no un máximo) se calcula cuidadosamente en cada caso concreto atendiendo a los daños que previsiblemente sufriría el asegurado si en el área delimitada en el contrato se produce un evento de la magnitud prevista. En modo alguno es una suma fijada arbitrariamente, como en los seguros de sumas o de vida.

Por otro lado, la determinación del área de medición tiene una gran relevancia: en la medida en que en el seguro paramétrico se pone el foco en el riesgo —esto es, en el posible acaecimiento del evento en el área marcada— más que en los daños concretos que sufra el asegurado, la cobertura no se limita a las propias instalaciones del tomador, sino que también puede incluir otras infraestructuras críticas para la cadena de suministro del cliente, pues los daños en estas pueden suponer la interrupción de su actividad empresarial y las pérdidas de beneficios que de ellos se le deriven pueden ser cubiertas por el seguro sin necesidad de que haya ningún daño material previo en su propio establecimiento. Esta es la característica distintiva del seguro paramétrico respecto al seguro de interrupción de actividad que se comercializa en el mercado español (y en otros) y constituye la principal ventaja de este tipo de seguros.

Es más, hay seguros en los que la indemnización acordada no cubre los daños materiales del establecimiento. Un ejemplo de este supuesto es el seguro paramétrico Insur8 de SwissRe: cuando el Hong Kong Observatory registra un tifón de nivel 8, el asegurado tiene derecho a cobrar una indemnización que abarca: *i)* la pérdida de beneficios por la ausencia de clientes durante el periodo en el que se verifiquen los efectos del tifón; *ii)* el reembolso de costes por la cancelación o el aplazamiento de un evento cancelado; y *iii)* la pérdida de beneficios derivada de la reducción de la productividad. En este sentido, el seguro paramétrico actúa como un mecanismo complementario del seguro que cubre los daños materiales del establecimiento del asegurado, pero en modo alguno lo sustituye.

Un parámetro o índice adecuado es una medida objetiva correlacionada con un riesgo específico y, en última instancia, con un daño patrimonial para el asegurado. En otras palabras, el desencadenante paramétrico puede adoptar diversas formas, pero es necesario que se trate de un estándar independiente, objetivamente medible, verificable (inmediatamente

después de producirse el siniestro) y que guarde una correlación estadísticamente significativa con las pérdidas reales (Sengupta y Kousky, 2020).

En síntesis, el índice cuyo rebasamiento supone la ocurrencia del siniestro no suele fijarse de manera genérica para una masa de asegurados, sino que se determina tomando en consideración las necesidades concretas de cada cliente (al igual, naturalmente, que la cuantía de los daños que se prevé pueda sufrir por un siniestro). Por lo tanto, y como hemos adelantado, normalmente se trata de seguros personalizados.

ii) Mecanismo de pago

Cualquier parámetro o índice que se utilice como base para una solución paramétrica debe ser verificable de manera independiente. En los seguros tradicionales, la ocurrencia del siniestro debe ser comunicada por el asegurado; los peritos de la aseguradora comprueban si se trata de un siniestro cubierto o no y, si ese es el caso, determinan en qué medida debe indemnizar los daños causados. En cambio, en el seguro paramétrico las partes se someten de antemano a la información que sobre el evento suministre un tercero ajeno a ellas —que puede ser una entidad pública o una empresa privada especializada— para determinar la producción de un siniestro durante el curso del contrato. En general, se recurre a terceros independientes incontestables e inaccesibles, y con frecuencia la comunicación del siniestro es prescindible, dado que la aseguradora puede conocer la existencia del siniestro por sus propios medios y de manera inmediata.

También hay seguros paramétricos en los que no se requiere la intervención de ese tercero porque la aseguradora controla el riesgo durante el curso del contrato y puede conocer de manera inmediata que se ha producido un siniestro y es ella misma la que se lo comunica al asegurado, no al revés. Un ejemplo de esta modalidad es el seguro Parametrix lanzado por Lloyd's en 2020 para cubrir los riesgos derivados de la interrupción de la actividad que pueden sufrir las pymes por la interrupción de sus servicios en la nube ocasionada, entre otras causas, por ciberataques.

Sea como fuere, una vez que un tercero o la propia compañía (si es ella la que controla el riesgo) estiman que hay un siniestro cubierto —algo que se sabe de inmediato—, surge la obligación de pago. Por eso suele decirse gráficamente que el pago «se activa» con la sola producción del evento. En muchos seguros se pacta que el asegurado recibirá de manera automática el pago de la cuantía pactada en su aplicación de cartera del móvil. En otros, en cambio, la aseguradora se reserva el derecho de realizar una evaluación del siniestro para comprobar si el asegurado ha sufrido daños reales o no. Sobre esto volvemos más adelante.

4.2.2. Riesgo de base positivo y negativo

Como puede inferirse de lo expuesto hasta aquí, en el funcionamiento del seguro paramétrico existe un riesgo especial, a saber, que la cuantía pactada no se ajuste a los daños que puedan derivarse del siniestro. Es lo que se denomina habitualmente *basis risk*, que puede ser «negativo» —si el asegurado recibe del seguro una cuantía inferior al daño sufrido— o «positivo» —si la indemnización excede del daño experimentado—.

El riesgo de base negativo puede encontrarse también en un seguro tradicional de daños, dado que el asegurado corre con el riesgo de que la suma asegurada, que siempre es límite de la indemnización (art. 27 LCS), sea inferior al valor del interés en el momento del siniestro, dando lugar a la aplicación de la regla proporcional si no se ha excluido en el contrato (art. 30 LCS), o a la posibilidad de que la compañía le aplique descubiertos, franquicias o límites de indemnización con arreglo a lo pactado. Sin embargo, en el seguro paramétrico el problema se agudiza porque, además de que no siempre se acierta al predecir las consecuencias dañosas de un eventual siniestro, se considera que hay siniestro únicamente si el evento alcanza la magnitud pactada, de modo que, si esta es inferior, el asegurado no tiene derecho a cobrar indemnización alguna, aunque haya sufrido daños. En un seguro tradicional de daños esta eventualidad no suele ocurrir, pues generalmente se estima que hay siniestro cuando se produce la verificación del riesgo, sea cual sea su magnitud (sin perjuicio de que, cuando superan una determinada intensidad, la cobertura de ciertos riesgos queda a cargo del Consorcio de Compensación de Seguros y no de la compañía aseguradora), y lo que se cuestiona, en cambio, es la cuantía de los daños, que habrá de calcularse en cada caso tras el siniestro. Por ello, cada vez es más común que en los seguros paramétricos se establezcan distintos *triggers*, unos más elevados que otros, y se prevea el pago de prestaciones distintas, que serán más elevadas cuanto mayor sea el siniestro y más reducidas cuanto menor sea este.

Como se ha dicho, también cabe la posibilidad de que el riesgo de base sea positivo, pues la cuantía indemnizatoria se pacta de antemano y, llegado el caso, puede exceder de los daños sufridos. Para evitar que el seguro paramétrico dé lugar a enriquecimiento, en los condicionados de algunas aseguradoras pueden encontrarse cautelas de diversa naturaleza que alejan al contrato de la estructura lógica de funcionamiento *if-then* expuesta anteriormente. Así, no es raro que las compañías exijan al asegurado que pruebe que sufrió daños como requisito para recibir el pago, aunque no se lleve a cabo una peritación como en el seguro tradicional. Es más, hay también aseguradoras que incluso se reservan expresamente el derecho a comprobar las pérdidas sufridas por el asegurado mediante la intervención de peritos, por lo que el siniestro puede acabar liquidándose como en un seguro tradicional. Otro mecanismo al que también se recurre para frenar el posible enriquecimiento es introducir en las pólizas una cláusula en virtud de la cual se obliga al asegurado a invertir la indemnización obtenida por esta vía en la compensación de pérdidas derivadas del siniestro y se fija un plazo (por ejemplo, de un año) para ello. Como es sabido, las indemnizaciones obtenidas en los seguros de daños son, en principio, de libre aplicación, como las derivadas de seguros de sumas, pero esa libertad puede limitarse. La estipulación de una cláusula como la que se ha descrito no es una novedad. Es una aplicación a este campo de la cláusula de reconstrucción o reposición, que suele incorporarse en el contrato de seguro a valor nuevo precisamente para que el asegurado no se beneficie abiertamente del exceso de la cuantía de la indemnización obtenida sobre el valor real de su interés en el momento del siniestro (para un amplio análisis del uso de esta cláusula en el seguro a valor nuevo, *vid.* Muñoz Paredes, 1998: 335 ss.). La estipulación prevista en el seguro paramétrico tiene una finalidad claramente análoga. También hay cláusulas que simplemente limitan o reducen

la responsabilidad de la aseguradora sin que su presencia en la póliza esté justificada por el mantenimiento de la finalidad indemnizatoria del seguro. En este sentido, hay pólizas en las que, aunque en principio se prevea que la indemnización se fijará atendiendo al «valor convenido» (entendiendo por tal la suma indemnizatoria pactada) con independencia de los daños realmente sufridos, ese «valor convenido» opera como un máximo, pues junto a él se fija un límite de indemnización, de manera que, acaecido el siniestro, la indemnización será el menor valor entre *a)* el valor convenido y *b)* el límite de indemnización, una vez deducidos los pagos realizados para eventos ocurridos anteriormente durante la vigencia del seguro.

4.2.3. Calificación jurídica y compatibilidad con el principio indemnizatorio

En relación con la calificación jurídica de este seguro, cuestión escasamente estudiada, suele afirmarse (Jerry, 2023) que es un tipo de seguro con valor convenido o estimado —o, si se quiere, una póliza estimada—. Las propias pólizas utilizan esta expresión para referirse al método de fijación de la cuantía indemnizatoria. Ahora bien, en nuestro ordenamiento, el hecho de que el contrato haga referencia al «valor convenido» no habilita para estimar que el asegurado tiene derecho, sin más, al cobro de toda la prestación con independencia del daño sufrido. Esta conclusión no deriva solo de la prohibición general de enriquecimiento injusto (art. 26 LCS), dado que, en materia de pérdida de beneficios (el tipo de riesgo en el que mayor aplicación tienen los productos paramétricos), el artículo 67 del mismo cuerpo normativo dispone: «Si el contrato tuviera exclusivamente por objeto la pérdida de beneficios las partes no podrán predeterminar el importe de la indemnización». Esta norma es imperativa porque su finalidad es, precisamente, salvaguardar el principio indemnizatorio, y refleja el temor a que las pólizas estimadas y la cobertura del lucro cesante puedan servir para que el asegurado obtenga un enriquecimiento, aunque entre nosotros hay autores (Tirado Suárez, 2010: 1450-1452; Girgado, 2015: 31) que apoyan la validez de la cláusula de estima también para los beneficios esperados contra lo dispuesto en el precepto arriba transcrito, dado que es más favorable para el asegurado *ex* art. 2 LCS, siempre, eso sí, que se mantenga la finalidad indemnizatoria del seguro. Tampoco podemos olvidar que, en otros ordenamientos, entre ellos el alemán, una prohibición idéntica a la del art. 67 LCS está siendo reemplazada por interpretación más flexible de aquel principio, aunque lo cierto es que en el nuestro no se ha derogado. Por otro lado, cabría argumentar que el seguro en cuestión no cubre únicamente la pérdida de beneficios y que el artículo 67 LCS se aplica solo a los seguros que cubren «exclusivamente» este tipo de daños, pero entonces toparíamos con otro escollo: la póliza estimada (art. 28 LCS) tiene como solo efecto predeterminar el valor indemnizable en caso de siniestro total (pero no libera al asegurado de la obligación de probar que el siniestro se ha producido ni de la prueba de la preexistencia del interés asegurado en el momento del siniestro) y, si el siniestro es parcial, como ocurre en numerosas ocasiones, la indemnización se limitará a los daños efectivamente sufridos.

Con todo, un seguro en el que la compañía se comprometa al pago con total independencia de los daños sufridos por el asegurado no encaja en el concepto de póliza estimada y, si apuramos, ni siquiera en el de seguros de daños. Más bien se aproxima al seguro de sumas, dado que en este no se exige en modo alguno que el siniestro haya producido un daño. Ahora bien, como hemos visto, si algo caracteriza al seguro paramétrico es la variabilidad de sus cláusulas, de modo que en él pueden incluirse medidas que lo mantengan dentro de los márgenes conceptuales del seguro de daños sin que pierda completamente su esencia y su eficacia. Por ejemplo, si el contrato exige acreditar que el siniestro afectó al asegurado sin que este requisito suponga la realización de una tasación pericial de las pérdidas, estaríamos razonablemente dentro de los límites del seguro de daños y, al mismo tiempo, no se perdería la utilidad del seguro paramétrico como sistema de pronto pago. Además, a diferencia de los seguros de sumas, en los que la suma asegurada se elige de manera arbitraria, el monto previsto en el seguro paramétrico como indemnización se predetermina, pero de forma objetiva, es decir, sobre la base de las pérdidas que, de acuerdo con datos fiables, previsiblemente sufriría el asegurado si se produjera un evento de la magnitud pactada. No hay que olvidar, por otra parte, que, a diferencia del seguro de daños tradicional, el seguro paramétrico suele cubrir riesgos —como las catástrofes naturales— en los que el asegurado no puede influir en la producción del siniestro, factor que supone por sí solo un medio de salvaguarda del principio indemnizatorio, pues uno de los fundamentos de este principio es, precisamente, el temor a la causación voluntaria del siniestro.

Debe recordarse, además, que el principio indemnizatorio no ha permanecido estático ni en su formulación legal ni en la interpretación de las diferentes normas que lo han proclamado y regulado. En nuestro ordenamiento, la evolución histórica del principio indemnizatorio es clara. En sus orígenes (desde las antiguas Ordenanzas marítimas de Barcelona (1435), Burgos (1538) y Bilbao (1560)) prohibía asegurar las cosas (los buques) por todo su valor real, y establecía que una parte debía quedar a cargo del asegurado. Posteriormente, permitió la cobertura plena de ese valor, sin descubiertos (gracias al derogado art. 751 del Código de comercio de 1885, que, frente al art. 854 del Código de 1829, permitía excluirlos si así se pactaba). Más recientemente, el principio ha posibilitado, sin perder su esencia, que la indemnización alcance el valor nuevo del interés. Es más, esta forma de asegurar los bienes de uso (de todo tipo) se ha convertido en el mecanismo ordinario y ha relegado al seguro a valor real a un papel totalmente secundario. Más aún: las cautelas que fueron introduciéndose en el seguro a valor nuevo para evitar el posible enriquecimiento del asegurado se han flexibilizado e incluso se han eliminado de los condicionados a medida que las aseguradoras han ido comprobando que no se producía el temido aumento del riesgo subjetivo.

Por otro lado, la formulación del principio indemnizatorio en nuestro ordenamiento es lo suficientemente abierta como para acoger nuevas formas de aseguramiento, entre ellas el paramétrico. A diferencia de otros sistemas legales —por ejemplo, Chile y Colombia—, que, al consagrar de manera estricta el principio indemnizatorio en sus regulaciones del seguro de daños, se han visto obligados a reformar varias normas en 2023 para que se puedan contratarse seguros paramétricos en su mercado sin que se ponga en duda su licitud, el

artículo 26 de la Ley de Contrato de Seguro es flexible. En efecto, esta norma no prohíbe que el asegurado obtenga cualquier tipo de enriquecimiento, sino que únicamente proscribe el enriquecimiento «injusto», de modo que da cierto cobijo legal al enriquecimiento que puede considerarse «justo» o justificado (Sánchez Calero, 2010: 608; y Girgado, 2005: 133 y ss.), como yo misma sostuve en mi tesis doctoral (Muñoz Paredes, 1997: 366-369). Aparte del seguro a valor nuevo, si hay un caso en el que está justificado cierto enriquecimiento es precisamente el del seguro paramétrico. Ello porque esta modalidad ofrece indemnizaciones casi automáticas para todo tipo de desastres que los propios seguros tradicionales de daños no cubren y, por tanto, permite que el seguro cumpla su función económica de forma mucho más plena. El papel fundamental que desempeña como complemento del seguro tradicional —y como seguro principal para muchos asegurados que no pueden acceder a pólizas comunes— justifica esa conclusión. En definitiva, si para evitar todo enriquecimiento hubiera de proceder en todo caso a una liquidación pericial como en un seguro tradicional, perdería la gran utilidad que tiene y que ha impulsado extraordinariamente su expansión.

5. CAMBIOS EN LA DISTRIBUCIÓN: SEGUROS EMBEBIDOS (*EMBEDDED INSURANCE*)

Actualmente, muchos de los productos o servicios que contratamos *online* a través de plataformas llevan embebido un seguro que se contrata a la vez, sin salirse del itinerario de compra de la página que se está utilizando. Esto permite que muchos de los riesgos que llevan aparejados sean cubiertos, a diferencia de lo que sucedía anteriormente: dado que era necesario contratar un seguro por separado, en muchos casos sencillamente quedaban al descubierto.

Los embebidos son seguros de contratación digital y se comercializan por una empresa *que no es aseguradora ni mediadora de seguros tradicional*, sino que se dedica a la colocación en el mercado de productos o servicios *de otra naturaleza* para ofrecer a sus clientes la cobertura de riesgos generalmente relacionados con su actividad principal. En este sentido, suponen un cambio profundo en la distribución de los seguros y son un ejemplo claro de la adaptación del seguro a las necesidades de las personas, pues les permite obtener una cobertura inmediata —y sin pérdidas de tiempo— de los riesgos que afectan al bien o servicio que les interesa. Pueden introducirse en cualquier proceso de compra y en distintos ámbitos —finanzas, telecomunicaciones, energía, viajes o comercio electrónico—. Son canalizados por la propia empresa comercializadora del producto o servicio, que es la que aparece ante el cliente, no siempre consciente de que es otra empresa la que asume la cobertura aseguradora o de que, en realidad, la empresa comercializadora vende seguros por cuenta de un mediador de seguros.

Los seguros embebidos tienen precedentes muy claros en la práctica anterior a la digitalización del sector asegurador. Son, en realidad, una derivación de los seguros que se comercializaban de forma vinculada a la adquisición de bienes o la contratación de servicios

con presencia física de las partes antes de la digitalización: por ejemplo, el que ofrecían (y ofrecen) habitualmente el concesionario del automóvil al comprador o la gran superficie al adquirente de un ordenador o un electrodoméstico. La diferencia es que el seguro embebido es digital. Además, el concesionario o el vendedor del ordenador no necesariamente ofrecen un seguro concreto de una aseguradora específica, sino que lo hacen a través de una correduría.

Otro antecedente claro es el seguro ligado al alquiler de un automóvil ofrecido por la empresa de alquiler al cliente en el propio mostrador en el que se cierra la contratación. En este supuesto, a menudo el cliente no tiene posibilidades de elegir ni la aseguradora, pero sí es habitual que se le ofrezca determinar el alcance de la cobertura y, por tanto, el precio, siempre dentro de la panoplia de ofertas de la aseguradora que colabora con la empresa de alquiler. Este tipo de compañías ofrecen la misma opción (seguros embebidos en el proceso de compra) cuando el alquiler se contrata *online* a través de la web de la aseguradora (es el caso, por ejemplo, de los seguros que oferta Europcar en su propia web).

Ahora bien, cada vez hay más seguros embebidos en la contratación de todo tipo de bienes o servicios *online*. Se ofrecen como un añadido antes del pago y el cliente no tiene realmente más opciones que la de contratar o no un seguro en las condiciones que se le ofrecen. Si decide no suscribirlo, continúa el proceso de contratación del bien o servicio y, si decide hacerlo, la prima se añade al precio final de venta. En este último caso, es habitual que el comprador no tenga la posibilidad de comparar ofertas de distintas aseguradoras ni de elegir cierto nivel de cobertura entre los que ofrezca la aseguradora elegida por la comercializadora. Es más, si acepta contratar la cobertura, ni siquiera sabrá cuál es su aseguradora ni qué condicionados generales constituirán la base de su póliza.

Dada su forma de contratación, no todos los seguros pueden ser embebidos: tienen que ser productos simples, transparentes y fáciles de entender, y deben prever un proceso simple de reclamación de pago de daños derivados de siniestros (Instech London, 2021: 5). Sin embargo, se considera previsible que cada vez se vayan incorporando a la tendencia seguros más complejos, incluidos los del ramo de vida (Dutch authority for the Financial Markets, 2023: 15). Ya hay *insurtechs* (como la alemana Embea) que han empezado a diseñarse seguros de vida para embeber en la contratación servicios tan diversos como la reserva de un vuelo o los préstamos de dinero.

Por otro lado, en la medida en que conviven la contratación física y la digital de los mismos productos y servicios, se mantendrán en el mercado los seguros accesorios contratados de forma tradicional y los seguros embebidos en el itinerario digital, aunque se espera un crecimiento muy relevante de los embebidos en los próximos años.

Los seguros embebidos presentan indudables ventajas: sirven como instrumento para cerrar la denominada *protection gap*, dado que cubren riesgos que, a falta de este tipo de seguros, quedarían habitualmente al descubierto: es sabido que lo que interesa al asegurado es el contrato principal, no el accesorio, y que no siempre es consciente del riesgo asociado a la operación que realiza o se siente inclinado a acudir a una aseguradora o mediador a que le cubra ese riesgo a través de la adaptación de un contrato estándar. En cambio, los segu-

ros embebidos son personalizados y de muy sencilla contratación. No obstante, también plantean algunos problemas. En el caso del mercado español, uno de los más evidentes es el regulatorio. En efecto, la vigente regulación de la distribución de seguros (Real Decreto-Ley 3/2020, de 4 de febrero) no se ajusta a la expansión de este tipo de seguros, dado que, al igual que la Directiva que transpone, deja fuera de su aplicación la generalidad de este tipo de operaciones (cfr. art. 130.2), a las que considera de menor entidad, pese a que su número puede ser muy elevado. Semejante exclusión supone que los mediadores que las realizan, que encajan en la categoría de «mediadores de seguros complementarios», no tendrán que inscribirse en el registro de mediadores. Si no están inscritos, es claro que la DGSFP no puede controlar ni quiénes son ni las actividades que realizan. Considerando que estos mediadores pueden realizar un elevadísimo número de operaciones a través de las plataformas que permiten el acceso a un número masivo de clientes, no parece en absoluto adecuado que actúen al margen de la supervisión del regulador. Este déficit deberá revisarse y es esperable que se haga si, tal y como se prevé, el «embebimiento» de seguros sigue creciendo en los próximos años.

6. CASOS DE INTERSECCIÓN ENTRE LAS DISTINTAS MODALIDADES DE SEGUROS

Los seguros *usage-based*, *on-demand* y paramétricos son nuevas formas de configuración del seguro de daños. Por su parte, los seguros embebidos se caracterizan por su forma de distribución. Por ello, nada impide que un solo seguro pueda ser, al mismo tiempo, *usage-based*, *on-demand* o paramétrico y embebido. Veamos algunos ejemplos (más ampliamente, Muñoz Paredes, 2024).

6.1. SEGUROS EMBEBIDOS QUE SON A LA VEZ PARAMÉTRICOS

Un sector en el que se están imponiendo los seguros paramétricos embebidos es el de la cobertura de retrasos y cancelación de vuelos. Es el caso de los siguientes:

– El seguro de Wakam y Koala para viajes: Wakam (wakam.com/es), aseguradora digital, se asoció en 2019 con la *insurtech* Koala (https://hikoala.co/products/) para diseñar seguros paramétricos de retraso y cancelación de vuelos. Como todos los seguros paramétricos, se contratan digitalmente, su diseño es muy sencillo y ofrecen el pago automático de siniestros sin requerir prueba alguna al asegurado, dado que Koala está conectada en vivo con la información sobre vuelos y puede conocer la ocurrencia del evento asegurado de manera inmediata. El producto de seguro se embebe a través de APIs en las webs de reservas de las aerolíneas y el cliente puede elegir si lo añade o no cuando contrata el vuelo.

– El seguro de Wakam y la *insurtech* Setoo (setoo.com) para retrasos de viajes en ferry: se trata de un seguro paramétrico que puede embeberse en las webs de reservas de

viajes, de manera que, si se produce un retraso, Setoo pagará automáticamente la compensación prevista al usuario asegurado sin mayores comprobaciones (sobre estos concretos seguros, *vid.*, Deloitte, 2023; y Gale, 2021).

6.2. SEGUROS EMBEBIDOS QUE SON A LA VEZ *USAGE-BASED* Y *ON-DEMAND*

– El seguro de Flock (flockcover.com) para flotas de vehículos: Flock es una aseguradora británica que diseña específicamente seguros para que puedan embeberse en plataformas de alquiler de vehículos. Estos seguros son flotantes, dado que a través del portal de esta compañía la empresa de alquiler de vehículos da aviso a Flock de cada vehículo que se alquila para que se active la cobertura, y lo mismo debe hacer cada vez que se acaba el alquiler, para que cese aquella. Se trata de un seguro *on-demand*, pues la empresa asegurada paga la prima por los períodos concretos de cobertura que ella misma decide en cada momento. Además, es un seguro telemático; la empresa puede obtener las características de la conducción de cada usuario, registrar las distancias recorridas y localizar el vehículo en cualquier momento. Así es el seguro a todo riesgo que ha contratado con Flock la empresa The Out (theout.com), una compañía de alquiler de coches Jaguar Land Rover que opera en el Reino Unido, y que embebe un seguro cada vez que pone a disposición de sus clientes un vehículo de su flota. La tecnología de Flock permite a The Out llevar a cabo un control constante de cada vehículo, lo que le facilita cobrar al arrendatario un sobreprecio si recorre más millas diarias de las previstas en el contrato de alquiler y localizar el vehículo si este no es devuelto, sin que ello implique la necesidad de adaptar la prima del seguro al uso real (más o menos seguro) de cada vehículo, más allá de la activación y desactivación de la cobertura por el asegurado. Es, por tanto, un seguro telemático embebido y, a la vez, *on-demand*. Además, The Out también aprovecha la tecnología de mitigación de riesgos de Flock, diseñada para reducir los accidentes de tráfico, aunando de este modo prevención y seguro.

– El seguro de impago de facturas de DAS y LISA: «Aseguratufactura.es» es un seguro digital que cubre el riesgo de impago de facturas. Ha sido concebido como un dispositivo asegurador basado en el uso, pues no se paga una prima en general para el riesgo de impago de créditos, sino una cantidad fija por cada factura que el asegurado decide asegurar. Por lo que respecta a su configuración técnica, se ha diseñado para que pueda embeberse en cualquier proceso de facturación de los clientes (un producto similar en el mercado británico es el de Barclays y Nimbla al que se refiere Instech London, 2020: 14).

7. CAMBIOS DERIVADOS DE LA APLICACIÓN DEL REGLAMENTO UE DE DATOS (*EU DATA ACT*)

A pesar de que, como se ha dicho uno de los riesgos asociados a los seguros telemáticos de todo tipo es el de la pérdida de la intimidad del asegurado —pues permiten a la aseguradora acceder a todo tipo de datos de su vida cotidiana— una reciente encuesta de la compañía Allianz (2023) realizada en 5 países europeos muestra que la mayoría de los conductores estarían dispuestos a ceder sus datos a la aseguradora si, a cambio, esta les ofrece un mejor servicio: en Alemania, se mostró a favor el 53 % de los encuestados y en Reino Unido el 61 %. En este clima favorable a la compartición de datos, se ha aprobado recientemente en la UE el Reglamento de Datos, que entró en vigor el 11 de enero de 2024 y será aplicable a partir de septiembre de 2025.

De acuerdo con la Comisión Europea (2023), la finalidad del Reglamento es «[…] potenciar la economía de datos de la UE y fomentar un mercado de datos competitivo haciendo que los datos (en particular los industriales) sean más accesibles y utilizables, fomentando la innovación impulsada por los datos y aumentando su disponibilidad». Para lograrlo, el texto normativo garantiza la equidad en la asignación del valor de los datos entre los distintos agentes implicados, y «[…] aclara *quién* puede utilizar *qué* datos y *bajo qué* condiciones».

La finalidad de esta regulación es proporcionar a los usuarios de productos conectados (empresas o individuos que tienen un bien de estas características en propiedad, *leasing* o alquiler) el control sobre los datos que generan. Así, aunque es un asegurado cuyos datos son recopilados por un dispositivo con arreglo a lo previsto en el contrato, es claro que se le reconoce el control sobre ellos con independencia de que el dispositivo sea de su propiedad (por ejemplo, su *smartphone*) o de la aseguradora (como ocurre cuando la compañía pone a disposición del asegurado un dispositivo específico; tal es el caso del seguro llamado «Hello Connect», de la compañía Hello Auto) sin requerir un precio a cambio o, aunque lo exigiese, a modo de renta. Lo mismo puede decirse en relación con otros sujetos que puedan obtener datos del uso del bien conectado; por ejemplo, el fabricante de un automóvil.

Esta previsión va a tener distintas consecuencias jurídicas. Veámoslas:

El usuario del automóvil podrá exigir al fabricante el acceso gratuito a sus datos y solicitar que los transfiera a un tercero (por ejemplo, una aseguradora), pero el fabricante estará facultado para requerir al asegurado una compensación económica «justa, razonable y no discriminatoria» (art. 8.1 del Reglamento).

En el caso de la aseguradora disponga de los datos de un asegurado (porque los ha obtenido directamente mediante el uso de dispositivos o porque se los ha proporcionado un tercero con consentimiento del asegurado), consideramos que el asegurado podrá hacer uso de ellos para cambiar de aseguradora y obtener descuentos en la prima.

Además, dado que los asegurados podrán requerir de la aseguradora el acceso a sus datos en cualquier momento, es probable que haya menos sorpresas a la hora de reclamar el pago de la indemnización derivada de un siniestro.

Desde el punto de vista de las aseguradoras, el cambio puede ser muy relevante, pues, por ejemplo, en el ramo de automóvil, los clientes podrán acceder a los datos de los fabricantes de los vehículos. Esta posibilidad les permitirá crear seguros mejor adaptados a los intereses y los riesgos cubiertos, y prestar, al mismo tiempo, un servicio preventivo y de auxilio muy útil para el asegurado, especialmente si ha sufrido un accidente. Como apunta la compañía Allianz (2023), en caso de accidente la compañía podrá registrar automáticamente el alcance de los daños en tiempo real, evaluando los datos de posición y los de los sensores de colisión. La aseguradora podría informar inmediatamente a un servicio de grúa, reservar un coche de alquiler, pedir las piezas de recambio adecuadas y concertar una cita para que el asegurado acuda a un taller de reparaciones. En caso de accidente grave, podría incluso prestarse asistencia médica. Para acceder a los datos del automóvil, las compañías necesitarán, en todo caso, el consentimiento del asegurado y pagar al fabricante el precio que les pida, un monto que, como hemos visto arriba, la norma limita de forma indeterminada.

Como puede observarse, el servicio que cumple la aseguradora se extiende más allá de los límites del seguro y abarca servicios complementarios que realmente quedan fuera de su ámbito. No se trata únicamente de que el seguro haya pasado de cumplir un rol reparador a asumir una función preventiva, sino que, al ensanchar su papel en ambos aspectos y proporcionar medios al asegurado para evitar o mitigar los efectos adversos de un accidente, se convierte en un instrumento fundamental para el bienestar de las personas y la salvaguarda de sus patrimonios. Es previsible que este efecto se refuerce con la aplicación del Reglamento de datos a partir de septiembre de 2025.

NOTAS

1. Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo de 19 de octubre de 2022, relativo a un mercado único de servicios digitales, DO L 277 de 27.10.2022.

2. 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), de 21 de abril de 2021, COM (2021) 206 final.

3. Orientación general del Consejo (6/12/2022; 15698/22) a la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial.

4. Vid., Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts [COM (2021)0206 – C9-0146/2021 – 2021/0106(COD)]).

5. Reglamento (UE) 2022/2554 del Parlamento Europeo y del Consejo de 14 de diciembre de 2022 sobre la resiliencia operativa digital del sector financiero y por el que se modifican los Reglamentos (CE) 1060/2009, (UE) 648/2012, (UE) 600/2014, (UE) 909/2014 y (UE) 2016/1011, DO L 333 de 27.12.2022, 1-79.

6. Superintendencia de Seguros Privados. Circular n.º 592, de 26 de agosto de 2019. Disponible en: https://www.in.gov.br/en/web/dou/-/circular-n-592-de-26-de-agosto-de-2019-213190421.

BIBLIOGRAFÍA

AEFI-UNESPA (2023): *Informe sobre la digitalización de la industria aseguradora* [en línea] <https://www.unespa.es/main-files/uploads/2023/03/Informe-Digitalizacion-Industria-Aseguradora_feb_2023-DEF_V5.pdf>. [Consulta: 17/03/2024.]

ALLIANZ (2023): «Vehicle data can become a turbo for the European digital economy» [en línea], <https://www.allianz.com/en/press/news/business/insurance/231017-allianz-vehicle-data-can-become-a-turbo-for-the-european-digital-economy.html>. [Consulta: 23/01/2024.]

COMISIÓN EUROPEA (2023): «Data Act explained» [en línea], <https://digital-strategy.ec.europa.eu/en/factpages/data-act-explained>. [Consulta: 17/01/2024.]

— (2024): «El Reglamento de Servicios Digitales (Digital Services Act- DSA) explicado» [en línea], <https://www.aepd.es/documento/el-reglamento-de-servicios-digitales-dsa-explicado.pdf>. [Consulta: 21/02/2024.]

EIOPA (2021): «Artificial intelligence governance principles: towards ethical and trustworthy artificial intelligence in the european insurance sector», [en línea] <https://www.eiopa.europa.eu/sites/default/files/publications/reports/eiopa-ai-governance-principles-june-2021.pdf>. [Consulta: 21/02/2024.]

— (2023): *Consumer Trends Report* 2023 [en línea] <https://www.eiopa.europa.eu/document/download/015404b4-a289-41a2-a044-17fa6a96799b_en?filename=EIOPA-BoS-23-470-%20Consumer%20Trends%20Report%202023.pdf>. [Consulta: 13/03/2024.]

— (2024): *Report on the Digitalisation of the European Insurance Sector* [en línea] <https://www.eiopa.europa.eu/document/download/6ca9e171-42b9-44d7-a2e6-beaf0134ecb8_en?filename=Report%20on%20the%20digitalisation%20of%20the%20European%20insurance%20sector.pdf>. [Consulta: 28/01/2024.]

— «Dark patterns in insurance: practices that exploit consumer biases», https://www.eiopa.europa.eu/tools-and-data/behavioural-insights-insurance-and-pensions-supervision/dark-patterns-insurance-practices-exploit-consumer-biases_en

DE NOVAES DA SILVA, Anthony, (2024): «Overcoming Dark Patterns in Insurance by Legal Design», en M. L. Muñoz Paredes (dir.), *Digitalización y Seguro* (dir.), Granada: Comares, 241-257.

DELOITTE (2023): «The ecosystem imperative. Embedded finance: customer relationships and value web dynamics» [en línea] <https://www.deloitte.com/content/dam/assets-shared/docs/iif-deloitte-embedded-finance-report-3-august.pdf>. [Consulta: 21/03/2024.]

DUTCH AUTHORITY FOR THE FINANCIAL MARKETS (2023): «Technology towards 2033. The future of insurance and supervision» [en línea] <https://www.afm.nl/en/sector/actueel/2023/april/kansen-risico-digitalisering-verzekeringsmarkt>. [Consulta: 20/03/2024.]

ELLIOTT, Michael W. (2019): «Rethinking the Limits of Parametric Insurance», *Insurance Journal* [en línea] <https://www.insurancejournal.com/news/international/2019/09/24/540903.htm>. [Consulta:18/03/2024.]

GALE, Henry (2021): «Wakam's embedded parametric solutions for travel insurance» [en línea] <https://www.linkedin.com/pulse/wakams-embedded-parametric-solutions-travel-insurance-henry-gale/>. [Consulta: 18/03/2024.]

GIRGADO PERANDONES, Pablo (2005): *El principio indemnizatorio en los seguros de daños*, Granada: Comares.

— (2015): *La póliza estimada. La valoración convencional del interés en los seguros de daños*, Madrid: Marcial Pons.

INSTECH LONDON (2021): «Insurance: to Embed, or not to Embed», [en línea] <https://www.instech.co/insight/insurance-to-embed-or-not-to-embed>. [Consulta: 18/12/2023.]

JERRY, Robert H. (2023): «Understanding Parametric Insurance: A Potential Tool to Help Manage Pandemic Risk», en M. L. Muñoz Paredes y A. Tarasiuk (eds.), *Covid-19 and Insurance*, *AIDA Europe Research Series on Insurance Law and Regulation*, vol.7, Berlín: Springer, 17-62.

JOHN, Daniel (2022): «Telematics in Motor Insurance and AI Governance» [en línea] <https://www.eiopa.europa.eu/document/download/d3ff96a6-b080-4476-b104-0dae6e192d67_en?filename=AI%20governance%20in%20motor%20insurance%20telematics%20%20-%20Huk%20Coburg.pdf&prefLang=es>. [Consulta: 15/06/2023.]

LIN, Xiao y W. Jean KWON (2019): «Application of Parametric Insurance in Principle/Regulation-Compliant and Innovative Ways» [en línea] <https://ssrn.com/abstract=3426592>. [Consulta: 2/10/2023.]

MUÑOZ PAREDES, M.ª Luisa (1997): *El seguro "valor nuevo"*, tesis doctoral depositada en la Universidad de Oviedo.

— (1998): *El seguro a valor nuevo*, Madrid: Civitas.

— (2022): «Seguros "usage-based": luces y sombras», en A. Veiga (dir.), *Seguro de personas e inteligencia artificial*, Cizur Menor: Civitas Thomson Reuters, 245-273.

— (2023): «El seguro paramétrico o basado en índices», en J.A. García Cruces (coord.), *De iure mercatus. Libro Homenaje al Prof. Dr. Dr. h. c. Alberto Bercovitz Rodríguez-Cano*, Valencia: Tirant lo Blanch, 3371-3415.

— (2024): «Seguros embebidos (embedded insurance)», en M.L. Muñoz Paredes (dir.), *Digitalización y Seguro*, Granada: Comares, 51-82.

NAIC. CENTER FOR INSURANCE POLICY AND RESEARCH (CIPR) (2023): «On-demand insurance», [en línea], <https://naic.soutronglobalnet/Portal/Public/en-GB/RecordView/Index/26558>. [Consulta: 15/02/2024.]

O'NEIL, Cathy (2016): *Weapons of Math Destruction*, London: Penguin UK.

RUIZ ECHAURI, Joaquín (2020): «Seguros on demand en España», [en línea], <https://www.perezllorca.com/wp-content/uploads/2020/09/nota-juridica-seguros-on-demand-en-espana.pdf>. [Consulta: 2/02/2024.]

SÁNCHEZ CALERO, Fernando *et* al (2010): Ley *de Contrato de Seguro*, 4ª ed., Cizur Menor: Aranzadi.

SENGUPTA, Rohini y Carolyn Kousky (2020): «Parametric insurance for disasters» [en línea] <https://riskcenter.wharton.upenn.edu/wp-content/uploads/2020/09/Parametric-Insurance-for-Disasters_Sep-2020.pdf>. [Consulta: 5/03/2024.]

SWISSRE CORPORATE SOLUTIONS (2020): «What is the most popular parametric solution? Cat in a box explained», [en línea] <https://corporatesolutions.swissre.com/insights/knowledge/what-is-the-most-popular-parametric-solution-cat-in-a-box-explained.html>. [Consulta: 02/04/2024.]

— (2023): «What is parametric insurance?» [en línea] <https://corporatesolutions.swissre.com/insights/knowledge/what_is_parametric_insurance.html>. [Consulta: 8/02/2024.]

TIRADO SUÁREZ, F. JAVIER (2010): en SÁNCHEZ CALERO, Fernando, *et al.,* (2010): *Ley de contrato de seguro,* 4.ª edición, Cizur Menor: Aranzadi.

Fecha de recepción: 1 de julio de 2024.
Fecha de aceptación: 20 de octubre de 2024.

DERECHO DEL TRABAJO, INTELIGENCIA ARTIFICIAL Y TUTELA DE LAS PERSONAS*

LABOUR LAW, ARTIFICIAL INTELLIGENCE AND THE PROTECTION OF PERSONS

Miguel Rodríguez-Piñero Royo

Catedrático de Derecho del trabajo
Universidad de Sevilla
Senior Counsellor, PwC

«Oh, P.I. stuff mostly» Hackworth said. Supposedly Finkle-McGraw still kept up with things and would recognize the abbreviation for pseudo-intelligence, and perhaps even appreciate that Hackworth had made this assumption. Finkle-McGraw brightened a bit. «You know, when I was a lad, they called it A.I. Artificial intelligence».

(Neal Stephenson, The Diamond Age)

RESUMEN

El Derecho del trabajo ha estado siempre condicionado por la evolución de la tecnología. En este momento, el contexto tecnológico se caracteriza por la digitalización y por la utilización de la Inteligencia Artificial (IA), avances que ha impactado en el ámbito del trabajo en general, no solo en el empleo asalariado. El objetivo de este artículo es doble: por un lado, analizar cómo afectan los sistemas de IA a las relaciones laborales y al mercado de trabajo; por otro, estudiar y valorar la respuesta del ordenamiento jurídico a este fenómeno. Los poderes públicos han reaccionado mediante la construcción de un verdadero «Derecho algorítmico» que tiene una relevante dimensión laboral, dado que las personas trabajadoras se ven potencialmente afectadas por la generalización del uso de sistemas de IA en las organizaciones en las que están empleadas. Se trata de una rama emergente del Derecho que se nutre de componentes de distinto origen y naturaleza y que presenta un alto grado de complejidad y sofisticación técnica.

PALABRAS CLAVE

Personas trabajadoras, gestión de personas, control algorítmico, derechos de los trabajadores, discriminación

ABSTRACT

Labour Law has always been conditioned by the evolution of technology. At this moment, the technological context is characterized by both digitalization and the use of Artificial Intelligence (AI). These advances have impacted the field of work in general, not only salaried employment. The objective of this work is twofold: on the one hand, to analyze how AI systems affect labor relations and the labor market; on the other hand, to study and assess the response of the legal system to this phenomenon. Public authorities have reacted building a true «Algorithmic Law», which has a labor dimension, given that as workers are potentially affected by the widespread use generalization of AI systems in the organizations that employ them. This emerging branch of Law is nourished by components of different origin and nature and presents a high degree of complexity and technical sophistication.

KEYWORDS

Workers, personnel management, algorithm management, workers' rights, discrimination.

DOI: https://doi.org/10.36151/TD.2024.108

* Este trabajo es un resultado científico del Proyecto de Investigación «La huida del mercado de trabajo y la legislación social en España (TRABEXIT), PID2022-141201OB-I00», de la Convocatoria 2022 – «Proyectos de Generación de Conocimiento», en el marco del Plan Estatal de Investigación Científica, Técnica y de Innovación 2021-2023.

DERECHO DEL TRABAJO, INTELIGENCIA ARTIFICIAL Y TUTELA DE LAS PERSONAS

Miguel Rodríguez-Piñero Royo
Catedrático de Derecho del Trabajo
Universidad de Sevilla
Senior Counsellor, PwC

1. PRESENTACIÓN: DERECHO DEL TRABAJO, TECNOLOGÍA E INTELIGENCIA ARTIFICIAL

El Derecho del trabajo, entendido no como la regulación del trabajo asalariado y dependiente (que ha existido desde momentos históricos muy anteriores), sino como un conjunto normativo aplicable a grandes sectores de la población y construido sobre la base del Derecho de la Codificación —que, sin embargo, se aleja de este en lo que hace a la intensidad de su tratamiento regulatorio y a su carácter pro-operario—, nació condicionado por su contexto tecnológico. Los cambios tecnológicos introducidos en la Revolución Industrial generaron la explotación laboral, que dio lugar a la cuestión social, a la reacción de las personas que la sufrieron y, tras largos años de años de lucha, a la construcción de un

Derecho propio diferenciado del Derecho civil que primero fue de clase y, posteriormente, se aplicó a todas las personas que prestan servicios para otras.

El Derecho del trabajo surgió, por tanto, en un contexto industrial específico y sus principales instituciones se diseñaron tomando en consideración la tecnología de aquel contexto. Desde entonces, sucesivas revoluciones tecnológicas han puesto a prueba su capacidad de adaptación. En general, el Derecho del trabajo ha evolucionado con retraso respecto a estas, circunstancia que ha provocado situaciones crónicas de obsolescencia regulatoria y déficits de protección. El enfoque regulatorio basado en los derechos fundamentales, propio de las últimas décadas, ha permitido corregir en parte estas carencias mediante la aplicación en sede jurisdiccional de diversos derechos, entre ellos los derechos a la intimidad y a la protección de datos, así los nuevos derechos digitales.

Cuando estábamos construyendo una regulación de las relaciones laborales adaptada a un sistema productivo altamente digitalizado, nos sorprendió el impacto de una nueva realidad, la inteligencia artificial (IA), tecnología que se encuentra en pleno desarrollo, que se extiende aceleradamente y que ha irrumpido en nuestras sociedades para cambiarlo todo. Uno de los ámbitos en los que la IA está provocando efectos más intensos es precisamente el del trabajo en general, no solo el del empleo asalariado, en todas sus facetas y dimensiones (Sanguineti Raymond y Muñoz Ruíz, 2024; Gil de Alburquerque, 2023; Lozano, 2023; Tolosí Signes, 2023).

El objetivo de este trabajo es doble: por un lado, analizar cómo afectan los sistemas de IA a las relaciones laborales y al mercado de trabajo; por otro, estudiar y valorar las respuestas que el ordenamiento jurídico ofrece para hacer frente al impacto de este fenómeno.

Dado que se trata de analizar un proceso de construcción regulatoria, este trabajo recurre a una metodología jurídica y prescinde de los instrumentos de otras ciencias sociales.

Disponemos de una enorme cantidad de fuentes de información, jurídicas o no, sobre la realidad que aquí será objeto de estudio. Por lo que respecta a la parte regulatoria, trabajaré tanto con los materiales legislativos —de *hard* y *soft law*— como con los estudios elaborados sobre ellos. En términos generales, manejaré la producción de la doctrina iuslaboralista española, que se ha ocupado extensa e intensamente de las cuestiones relacionadas con el impacto de la IA y que suministra materiales más que suficientes para construir este discurso.

Las hipótesis de partida son las siguientes: la IA está afectando ya a las personas que trabajan en múltiples aspectos, y ha generado importantes riesgos para sus derechos como trabajadores y como ciudadanos. Frente a esta realidad, los poderes públicos han reaccionado de manera decidida, diseñando instrumentos jurídicos *ad hoc* y elaborando conjuntos regulatorios completos y actualizados. Estamos, pues, en presencia del nacimiento de un verdadero «Derecho algorítmico» que tiene una relevante dimensión laboral, dado que las personas trabajadoras se ven potencialmente afectadas por la generalización del uso de sistemas de IA en las organizaciones en las que están empleadas. Este «Derecho algorítmico laboral» tendría su origen en el «Derecho digital del trabajo» elaborado para afrontar la

revolución digital que hoy se superpone a la iniciada por la IA. No obstante, el Derecho algorítmico laboral se diferencia del Derecho digital del trabajo en aspectos esenciales.

2. IA Y LAS PERSONAS QUE TRABAJAN

2.1. EL IMPACTO DE LA IA EN EL EMPLEO

El impacto más inmediato de la generalización del uso de sistemas de IA en las empresas tiene lugar en el ámbito del empleo. Si bien esta es una característica común a todos los cambios tecnológicos, la irrupción de la IA está mostrando unas implicaciones, una capacidad de difusión y, sobre todo, una imprevisibilidad sin precedentes. La imposibilidad de calibrar los efectos reales de la IA es consecuencia del ritmo de la innovación tecnológica, del acelerado diseño de nuevas utilidades y del propio funcionamiento de estos sistemas, cuyos *outputs* son en gran medida imponderables.

Las grandes preocupaciones que suscita el desarrollo de la IA son la pérdida de puestos de trabajo y la obsolescencia de las capacidades profesionales. La Guía del Mercado Laboral 2024 de la consultora Hays España indica que el 47 % de los trabajadores está preocupado por el impacto de la inteligencia artificial en el empleo. Sin embargo, el 24 % considera que contribuirá a crear nuevas oportunidades laborales, mientras que un 50 % lo estima que es útil para mejorar su productividad (Hays, 2024).

Los estudios sobre la IA son numerosos. En general, evitan hablar de destrucciones masivas de empleo y hacen hincapié en su efecto en el incremento de la competitividad y la creación de nuevos empleos y ocupaciones. Incluso la Organización Internacional del Trabajo ha negado que sea previsible un «apocalipsis de empleos», y ha sostenido que la IA provocará más bien una gran movilidad (OIT, 2023; OECD, 2023). Por su parte, la consultora PwC ha afirmado que «[…] el mercado laboral no volverá a ser el de ayer, pero habrá grandes oportunidades para quienes se adapten a la era de la IA» (PwC, 2024).

Según la consultora AFI, se observan principalmente dos impactos contrapuestos: el efecto complementariedad, que supone que la IA contribuye a mejorar aquellas tareas que ya se realizaban, por lo que impacta positivamente en el crecimiento del empleo; y el efecto sustitución, en virtud del cual la IA reemplaza determinadas tareas (AFI, 2024). No obstante, parece claro que los efectos potenciales pueden ser inmensos, y que estos dependerán tanto de las estrategias de uso de estas herramientas por las empresas como de la capacidad del Estado y de los interlocutores sociales para regularlo.

Recientemente se ha acuñado un concepto muy expresivo, el de 'exposición', que hace referencia modo en que puede afectar la introducción de sistemas de IA en una profesión o un empleo, en particular a sus posibilidades de supervivencia. El concepto sirve para graduar a los trabajos en función del previsible alcance de este impacto. Se habla, así, de trabajos muy, poco o nada expuestos (AFI, 2023).

La experiencia de un sector pionero en su uso, el de las plataformas digitales, evidencia tanto su potencial de creación de empleo como sus efectos «deslaboralizadores», dado que una parte importante de las personas prestan servicios a título de trabajadores autónomos. El trabajo por cuenta propia puede ser real o fraudulento, pero el efecto de su expansión es el mismo, a saber, la ausencia de un empleador responsable del cumplimiento de las medidas de tutela laboral.

2.2. LA DISCRIMINACIÓN ALGORÍTMICA

Uno de los mayores riesgos vinculados a la utilización de sistemas de IA en la gestión del personal laboral es la discriminación algorítmica, que es consecuencia de la presencia de sesgos en las decisiones que toma el programa. Los sesgos perjudican de manera injustificada e ilegal a las personas que pertenecen a ciertos grupos con determinadas características (Asquerino Lamparero, 2022*a*: 335; Rodríguez-Piñero, 2022*a*; Calvo Gallego, 2022). Se trata de una realidad contrastada en la práctica que se explica por una pluralidad de factores. La generalización del uso de la IA en aquellas fases de la dirección de personas, en las que tradicionalmente se producen conductas discriminatorias, explica su incidencia, que ha provocado una auténtica alarma en un momento en el que, como se sabe, el ordenamiento jurídico es especialmente sensible a la discriminación y se están ampliando tanto los motivos prohibidos como los instrumentos legales para combatirla.

Se han identificado los factores que dificultan la lucha contra estas prácticas: la invisibilidad de esta forma de discriminación; su extensión e intensidad; su complejidad técnica; y la autoridad matemática de los sistemas de IA que la provocan, al menos hasta ahora (Sáez Lara, 2020: 41; Soriano Arnanz, 2021: 11; Fernández García, 2023: 169; Rivas Vallejo, 2022: 295). La tutela frente a la discriminación algorítmica se enfrenta también al problema de la identificación de los sujetos responsables, bien por la pretendida automatización y objetivación de la decisión, en la que no habría intervención humana, bien por la interacción de una pluralidad de sujetos en la actuación discriminatoria (el diseñador de la herramienta, el responsable de su aplicación, los responsables de la empresa).

En la práctica, y a pesar de su pretendida objetividad, la IA puede potenciar la discriminación porque los datos con los que trabaja suelen reflejar pautas discriminatorias que la IA tiende a considerar adecuadas; puede haber sesgos no solo en los datos, sino también en las variables y los *proxys* utilizados, todo lo cual se traduce en un impacto discriminatorio que generalmente solo es identificable a través del concepto de discriminación indirecta, es decir, atendiendo al impacto desproporcionado de la medida pretendidamente neutral sobre el colectivo particular que resulta discriminado. Finalmente, la generalización del uso de los mismos programas en multitud de empresas hace que el daño potencial que estos puedan provocar se multiplique[1].

La Agencia Europea de los Derechos Fundamentales, que ha elaborado propuestas para protección de los derechos fundamentales en el desarrollo y uso de algoritmos, ha incluido entre estas la una medición del «potencial de discriminación» en relación con motivos diferentes como el género, la edad, la etnia, religión, y la orientación sexual y política

(EUAFR, 2024). Esta y otras medidas de transparencia y evaluación deben de verificarse tanto respecto a calidad de los datos como al modo en que funcionan los algoritmos («caja negra») y las facultades predictivas de los mismos.

2.3. LA GESTIÓN ALGORÍTMICA DE PERSONAS

La gestión algorítmica de personas se ha generalizado, y los programas diseñados para llevarla a cabo se utilizan ya en todos los ámbitos y fases de la realización de esta función, que corre a cargo de la empresa (Mercader Uguina, 2022; Ginès i Fabrellas, 2024*a*). Cabe distinguir la gestión algorítmica en sí misma, que supone la adopción de decisiones en materia de recursos humanos, y otros instrumentos como las medidas de vigilancia y control, en la que los sistemas de IA se han integrado con los instrumentos tradicionales de videovigilancia, auditoría informática, biometría y demás (Raimu *et al.*, 2022; OIT, 2022).

Este tipo de gestión algorítmica está especialmente presente en los procesos de selección, en los que con frecuencia deben gestionarse grandes volúmenes de solicitudes, desarrollarse pruebas y adoptarse decisiones arriesgadas debido al desconocimiento de muchos aspectos de la personalidad y la capacidad profesional de los candidatos. El ejercicio de esta función es la actividad en la que se denuncian los mayores riesgos. Se detectan importantes sesgos, errores e incluso discriminaciones directas en una situación en la que, además, las personas son especialmente vulnerables y no tienen a su disposición mecanismos de tutela efectiva. Me remito a lo ya dicho sobre la discriminación algorítmica, que se verifica con mayor frecuencia en esta fase del proceso de contratación.

A diferencia de las tecnologías anteriores, los sistemas de IA no interactúan físicamente con las personas en sus puestos, por lo que su impacto en los riesgos laborales debería ser limitado. Sí se combinan con otros soportes tecnológicos —entre ellos, robots y drones—, pueden ser una fuente de riesgos para la salud y la seguridad. También pueden tomar decisiones sobre el momento y lugar de prestación del trabajo, o sobre la carga, el ritmo y los contenidos del desempeño laboral, determinaciones que sí afectan directamente a la salud y la seguridad. Se ha hecho hincapié su capacidad de incrementar los riesgos psicosociales, como la fatiga mental producida por la monitorización constante, el estrés por las cargas de trabajo exageradas o la incomprensibilidad de las decisiones[2]. Al mismo tiempo, los sistemas de IA pueden ser enormemente útiles para mejorar la seguridad en el trabajo, dado que pueden predecir accidentes o enfermedades, identificar a las personas especialmente susceptibles ante algunos riesgos, y asignar eficientemente los recursos preventivos (Llorens Espada, 2024)[3].

Estos sistemas pueden servir para soportar decisiones de gestión relativas, entre otros aspectos, a los ascensos, los incrementos retributivos, la movilidad y las oportunidades formativas. También se utilizan para despedir, bien tomando la decisión del despido porque el trabajador no alcanza determinados niveles de productividad o para anticiparse a los riesgos, bien seleccionando a las personas que van a verse afectadas por esta medida. En todos estos cometidos, la gestión algorítmica tiene sus ventajas y sus riesgos. Entre las primeras podemos citar la posibilidad de manejar un gran número de variables, la objetivación y

despersonalización de la decisión, o la capacidad de aprendizaje y mejora continua de las herramientas. En general, la gestión algorítmica facilita la acción el área de dirección de personas, acelerando los procesos y eliminando sesgos humanos.

Por lo que respecta a los riesgos, la IA aplicará su propia programación, que puede incluir sesgos y errores, y que también puede ignorar los criterios y derechos reconocidos por la legislación y los convenios colectivos aplicables a la empresa. Es posible recurrir a decisiones automatizadas para dar una apariencia de objetividad, ocultar motivos reales y generar justificaciones. En general, la gestión algorítmica puede dificultar el control de los poderes empresariales, que se ejercen apoyándose en herramientas de difícil comprensión, elaboradas por terceros y, en muchos casos, opacas. Por no hablar de la difuminación de las responsabilidades entre el empleador y los responsables del programa (Rodríguez San de Galdeano, 2024).

2.4. IA E INTERVENCIÓN PÚBLICA EN LAS RELACIONES LABORALES

Los juristas, y los iuslaboralistas no somos una excepción, hemos adoptado una actitud recelosa ante la expansión de la IA y, por hecho, hemos centrado nuestra atención antes en los posibles riesgos derivados de su utilización que en sus posibles efectos positivos. Tenemos razones para ello, pues la experiencia de la implantación de tecnologías precedentes no ha sido especialmente buena debido a que no se tuvo en cuenta adecuadamente su impacto sobre las personas, una desatención que obligó a realizar intervenciones legislativas correctoras y reparadoras posteriores. Ahora bien, como ya hemos indicado, la IA tiene utilidades importantes para las empresas que redundan en beneficio de las personas que trabajan. También, y de esto me ocuparé ahora, son muy efectivas para mejorar la actuación de los distintos sujetos públicos que operan en el mercado de trabajo a través de la prestación de servicios y la tutela de sus derechos. En general, podemos identificar tres grandes ámbitos de la intervención estatal en los que ya se está utilizando la IA de manera sistemática: las políticas activas de empleo;la protección social pública; y la potestad sancionadora de la Administración laboral.

En el ámbito de las políticas activas de empleo, los distintos procesos requieren la toma de decisiones basadas en grandes volúmenes de información, en un contexto caracterizado por el gran número de demandantes de servicios, la enorme diversidad de estos y la gran cantidad de datos que deben ser manejados, amén de la carencia crónica de recursos en los servicios públicos de empleo. Estos rasgos explican que los organismos que tienen a su cargo la gestión pública del empleo recurran a herramientas *ad hoc* para realizar las tareas de diversa índole incluidas en su cartera de servicios. Ello es especialmente cierto en el caso de los servicios de intermediación, que buscan casar la oferta y la demanda de trabajo mediante la gestión de la información que tienen sobre ambas para identificar candidatos idóneos y oportunidades de colocación adecuadas (Rodríguez-Piñero, 2022*b*, 19; Pérez Anaya, 2022: 117; Asquerino Lamparero, 2022*b*: 21; García Coca, 2022: 383; Fernández García, 2024).

Últimamente se está extendiendo la aplicación de técnicas de perfilado estadístico individualizado a los demandantes de empleo para realizar estimaciones sobre sus posibilidades reales de colocación e identificar sus necesidades específicas en materia de formación, movilidad, etc. (Sánchez López, 2023: 63).

Por su parte, el sistema público de Seguridad Social utiliza técnicas de decisiones administrativas automatizadas tanto en materia de recaudación como de reconocimiento de prestaciones, así como para controlar el disfrute de estas, identificando abusos o usos indebidos (Goerlich Peset, 2021: 22).

Por lo que se refiere a la potestad sancionadora de la Administración, la Inspección de Trabajo utiliza desde hace años una «Herramienta de Lucha contra el Fraude», que en sus sucesivas versiones ha ido incorporando distintos instrumentos de IA predictiva tanto para detectar incumplimientos como para orientar la actividad inspectora hacia aquellas empresas en las que, de acuerdo con los cálculos realizados por el programa, el número de infracciones puede ser más elevado.

Una práctica frecuente en la actualidad es el envío de comunicaciones a los empleadores —que no llegan a ser apercibimientos— porque los sistemas de IA han detectado que existe una alta posibilidad de que se estén produciendo incumplimientos de la legislación laboral a partir de los datos que la Administración tiene de las empresas a las que los remite. El objetivo de esta medida es claramente preventivo y disuasorio, dado que en ella no está contempla la imposición de sanciones inmediatas. Los incumplimientos que se detectan— se producen en aquellas áreas hacia las que la Inspección dirige sus campañas, entre ellas el fraude en la contratación temporal o el control de la jornada laboral. Recientemente, el Gobierno ha planteado una reforma del registro de jornada que permita a la Administración laboral acceder directamente a la información de cada empresa al objeto de identificar posibles incumplimientos mediante el uso de instrumentos de IA.

3. LA CONSTRUCCIÓN DE UN DERECHO ALGORÍTMICO LABORAL

La respuesta del ordenamiento jurídico a las cuestiones que plantea la generalización de la utilización de la IA en las relaciones laborales se está produciendo de una manera muy rápida (Del Rey Guanter, 2023: 115) a través de la aprobación de un conjunto normas específicas y actualizadas al ya se denomina como «Derecho algorítmico laboral». Este sería una modalidad diferenciada (una *spin-off*, diríamos hoy) de otra rama del Derecho del trabajo elaborada en los últimos años, el Derecho del trabajo digital, que reúne las distintas normas que se ocupan del impacto de la digitalización en las relaciones laborales, incluyendo los derechos de nueva generación a los que llamamos «derechos digitales». Aunque ambas comparten origen y código genético, las notables diferencias que presentan justifican que se traten de manera diferenciada.

3.1. UN ENFOQUE PREVENTIVO EN EL TRATAMIENTO DE LA IA

A diferencia de lo ocurrido con otras revoluciones tecnológicas de mayor o menor alcance, en el caso de la IA la adaptación regulatoria se está produciendo de manera temprana, pues esta tecnología aún no ha terminado de desarrollarse y sus efectos en el trabajo laborales son todavía limitados debido a su escasa difusión en la esfera laboral. Puede afirmarse incluso que el tratamiento legal del uso de la IA ha adoptado un enfoque preventivo, en parte porque sus efectos potenciales generan no pocas preocupaciones, en parte porque es una realidad tecnológica que se está todavía conformando. El legislador ha sido consciente de los riesgos desde un primer momento, cuando se estaban poniendo las bases técnicas para el uso de los sistemas algorítmicos, y ha actuado en consecuencia. A este respecto, OCDE ha señalado la necesidad de una «urgente necesidad de actuar» para asegurar que la IA se utiliza de una manera responsable y fiable en los centros de trabajo (OCDE, 2023: 4).

Existía, de esta manera, una «conciencia situacional» sobre la IA, que es la que ha llevado a regularla antes incluso de que surgieran los problemas. Esta anticipación no es habitual en el mundo del Derecho, y refleja las particularidades de este cambio tecnológico. El papel del jurista no ha consistido, como en las mutaciones tecnológicas del pasado, identificar *a posteriori* la obsolescencia de la regulación o sus carencias para proponer mejores y actualizaciones, sino que las reglas se han diseñado y elaborado en el mismo origen con la finalidad de condicionar el propio desarrollo técnico. En la medida en que no existían construcciones jurisprudenciales o normas pioneras sobre para fundamentar el cambio normativo, el grado de innovación ha sido muy elevado.

3.2. LA DESCODIFICACIÓN DEL DERECHO DEL TRABAJO

La regulación de la IA está teniendo un efecto de descodificación del Derecho del trabajo, fenómeno que no es nuevo y que no se produce solo debido al surgimiento del Derecho algorítmico. Me refiero a la tendencia hacia la creciente reubicación de las normas laborales, por su contenido y sus destinatarios, fuera de los cuerpos legislativos propios e idiosincrásicos de este sector del ordenamiento. En efecto, el Derecho del trabajo español no está reunificado en último grado, al no existir un Código del Trabajo que, a semejanza de otros ordenamientos (entre ellos el francés, el portugués, el chileno y el ruso), reúna todas las normas que se aplican al trabajo asalariado. Sí existe, sin embargo, una cierta estructuración de sus distintos sectores desarrollada a través de grandes leyes que actúan como enclaves normativos para cada uno de ellos, un paquete de reglamentos y alguna legislación espacial. Hay, así, un Estatuto de los Trabajadores, una Ley de Prevención de Riesgos Laborales, una Ley de Empleo, una Ley Reguladora de la Jurisdicción Social… estas leyes, de gran extensión, agrupan los aspectos fundamentales de la regulación, dotándola *de facto* de cierto nivel de codificación.

La tendencia a ubicar contenidos laborales en cuerpos normativos no laborales comenzó con el Derecho Digital, y ahora está ocurriendo con el de la IA. La técnica de la regulación de la IA está generando grandes textos transversales que incluyen algunos con-

tenidos claramente laborales que no llegan a plasmarse en su normativa propia. El ejemplo paradigmático es el de los derechos digitales, ausentes de una norma legal que se llama a sí misma «Estatuto de los Trabajadores» y que contiene un artículo sobre «derechos de los trabajadores». Aquellos están regulados en un título monográfico de la Ley Orgánica de Protección de Datos y Garantía de los Derechos Digitales. Lo mismo ocurre con la creciente regulación de los aspectos relacionados con el trabajo de algoritmos y decisiones automatizadas, presentes en el Reglamento General de Protección de Datos y en la Ley de Inteligencia Artificial de la Unión Europea y, en el caso español, en su norma de trasposición, pero no reflejados en el Estatuto de los Trabajadores.

Se multiplican, de esta manera, las «normas paralaborales», según la terminología propuesta por Del Rey Guanter, construidas desde una perspectiva que no es la del Derecho del trabajo (Del Rey Guanter, 2023*b*). Lo mismo ocurre con las numerosas declaraciones y textos de *soft law* que se han adoptado en los últimos años.

Sin embargo, esta tendencia puede cambiarse si los partidos que conforman la coalición electoral surgida de las elecciones de junio de 2023 finalmente cumplen su compromiso de aprobar el «Estatuto del Trabajo del Siglo XXI». Este texto legal, que pretende ser más que una reforma o una reescritura del Estatuto de los Trabajadores aprobado en 1980, se concibe como como un cuerpo normativo no solo afecta al trabajo asalariado, sino también a otros supuestos, es decir, «[…] una red básica de derechos para todos aquellos que prestan actividades profesionales, desde los autónomos hasta los cooperativistas», tal y como señala el acuerdo de legislatura. Por lo que respecta a su contenido, el nuevo Estatuto deberá adaptar la regulación del trabajo por cuenta ajena para incluir dos realidades: la transición digital, que deberá incorporar la regulación del gobierno de los algoritmos; y la transición verde, que habrá de prever fórmulas de negociación colectiva que garanticen la sostenibilidad. Como se ve, se reconoce y prioriza la atención al impacto de la IA en las relaciones laborales, y se afirma que esta regulación deberá emplazarse en su lugar natural, que al menos a mi juicio no es otro que el cuerpo legislativo central de las relaciones de trabajo en España, el Estatuto previsto en el artículo 35 de la Constitución Española.

3.3. LA COMPLEJIDAD DEL DERECHO ALGORÍTMICO LABORAL

La respuesta del Derecho a la IA a los retos que esta genera adopta formas jurídicas particulares. De nuevo este fenómeno tiene sus orígenes inmediatos en el precedente directo de esta rama emergente, el Derecho digital, que le antecede solo unos años. Al igual que el Derecho digital, el Derecho algorítmico laboral se está construyendo a través de la agregación de normas jurídicas diversas, que incluyen manifestaciones tanto de *soft law* (entre ellas, declaraciones y cartas no vinculantes) como de *hard law*. Desde la perspectiva de los sistemas jurídicos continentales europeos de raíces romanistas, algunas de estas formas de expresión resultan poco habituales, pero no lo son para sistemas de distinta estirpe. El impacto de otras formas de entender lo jurídico puede explicarse por el hecho de que estamos ante un fenómeno global que está generando respuestas también globales por parte de la comunidad internacional y las organizaciones supranacionales que han asumido formas de

regulación de distintas tradiciones. Lo cierto es que estamos ante una nueva manifestación, más avanzada quizás, de la globalización jurídica.

A esta combinación del *hard law* y el *soft law* se une el papel atribuido a ciertos organismos especializados, agencias y grupos de trabajo que, aunque tener competencias para legislar, acaban creando Derecho cuando interpretan las normas legales o aportan guías y sugerencias para su aplicación.

El Derecho algorítmico del trabajo está configurándose a través un esquema multinivel en el que coexisten y colaboran normas elaboradas por diversas instituciones. Esto es lo común en el Derecho del siglo XXI. En el caso que nos ocupa, el centro de gravedad es el nivel supranacional, en el que se han aprobado los textos fundamentales en esta materia. Debe, no obstante, destacarse el interesante diálogo entre el nivel de la UE y el estatal, que se ha verificado en el caso español: España aprueba una ley sobre el trabajo en plataformas digitales que inspira al Parlamento Europeo a elaborar una directiva sobre esta realidad; por su parte, el Parlamento Europeo aprueba un Reglamento General de Protección de Datos que lleva a nuestro país a aprobar una Ley Orgánica que lo implementa (sin ser ello estrictamente necesario) y aprovechar para reconocer los derechos digitales.

Estas normas no solo son diversas, sino también técnicamente complejas, dado que son el resultado de unos trabajos preparatorios que han buscado soluciones adecuadas y efectivas para hacer frente a las nuevas realidades utilizando una terminología precisa. El conocimiento de la materia es completo, lo que se traduce en regulaciones potencialmente eficientes, de modo que, además de su tradicional función reguladora, desempeñan un papel didáctico, que se concreta en la incorporación de terminología y definiciones legales para facilitar su uso por los distintos operadores jurídicos.

Un factor adicional de complicación tiene que ver con la coyuntura en la que se ha empezado a elaborar el Derecho algorítmico. Coincidiendo con la preocupación por estas cuestiones se produjo en todo el mundo la extensión del trabajo en plataformas, especialmente las de reparto de comida a domicilio. Se trata de un fenómeno complejo, con múltiples variantes que tienen en común el hecho de que se basan en un soporte tecnológico, la plataforma digital, que utiliza sistemas de IA para asignar encargos, fijar precios y evaluar a los prestadores de servicios.

El principal problema jurídico que plantearon fue el de la calificación de los servicios que se prestaban por este soporte debido a las innovaciones que incorporaba este modelo en contraste con los tradicionales rasgos de ajenidad y dependencia, pero también a la extensión de los formatos de negocio basados en la contratación no laboral para incrementar flexibilidades y reducir costes (Baylos Grau, 2022: 95-113; Rodríguez Piñero, 2023: 13-35). Asimismo, el uso generalizado de sistemas de IA por estas plataformas —que más que un trabajo digital es un verdadero trabajo algorítmico— ha generado nuevas preocupaciones. La necesidad de resolver el primero de estos problemas (la calificación de la prestación laboral) ha motivado la aprobación de normas específicas sobre esta forma de empleo, que también se han ocupado de regular la cuestión del uso de la IA en ella. El resultado es una notable dualidad regulatoria: muchos ordenamientos disponen de dos conjuntos norma-

tivos que se ocupan de la IA, uno común para todas las empresas, y otro *ad hoc* para las plataformas digitales (todas o solo las de algunos sectores, como ocurre en España). Esta dualidad se explica por la prioridad que se dio a la resolución de la cuestión de los *riders* de plataformas de *delivery,* consecuencia, entre otros factores, del activismo del colectivo. Las plataformas, además, han sido una especie de campo de pruebas tanto para analizar el funcionamiento de la gestión algorítmica como para adoptar las medidas legislativas orientadas a regular esta modalidad. Aun así, una vez decidido que las cuestiones algorítmicas requerían un tratamiento completo y unitario, considero que el mantenimiento de dos líneas de regulación dejó de tener sentido.

3.4. LA GUERRA CULTURAL EN EL DERECHO DE LA IA

Lo que podemos calificar como una verdadera «guerra cultural» en la regulación de ciertas materias no es un fenómeno circunscrito al ámbito de la IA, sino que, por el contrario, se manifestándose en múltiples ramas del ordenamiento jurídico. Soy consciente de que utilizar la expresión 'guerra cultural', cuyos orígenes se remontan a la *Kulturkampf* de la Alemania guillermina, resulta tremendista, pero creo que existen motivos para recurrir a esta licencia. Con esta expresión me refiero a la existencia de diferentes formas de entender la regulación de las relaciones sociales en los distintos sectores de un ordenamiento cuyos elementos identitarios se reflejan en el tratamiento que dan a las materias que regulan y que pugnan entre sí por inspirar regulaciones concretas en las que se produce conflicto o concurrencia. Como jurista del trabajo, soy muy consciente de las particularidades de esta rama del Derecho en multitud de materias.

Antes que en las regulaciones concretas, los conflictos de cultura jurídica se encuentran allí donde el Derecho del Trabajo se superpone con otros sectores del ordenamiento jurídico. Una de sus primeras manifestaciones fue la regulación del contrato de trabajo internacional, tarea acometida por un Derecho internacional privado del trabajo originalmente poco sensible a las particularidades de esta figura frente a otros contratos. El amplio margen reconocido a las partes para elegir la legislación aplicable no encajaba bien con la concepción tradicional del ordenamiento laboral, más tuitiva y *pro operario*, y, justamente por ello, más desconfiada respecto a las declaraciones de voluntad de la parte débil de la relación laboral. Por ello, el Reglamento Roma I impone en su artículo 8 importantes restricciones a la libertad de elección que rigen con carácter general para los contratos.

El conflicto de culturas se puso de manifiesto con mucha claridad en dos aspectos esenciales el Derecho del trabajo digital. El primero es el papel del consentimiento más o menos informado, que es uno de los mecanismos de tutela general en materia de protección de datos. Cuando se trata de una persona que trabaja por cuenta ajena, es decir, que está sometida a una dependencia jurídica y económica, su validez no deja de ser discutible. Ya en 2017 un dictamen del Grupo de Trabajo del artículo 29 afirmó que los trabajadores casi nunca están en condiciones de dar, denegar o revocar el consentimiento libremente, habida cuenta de la dependencia que resulta de la relación empresario/trabajador. Considerando el desequilibrio de poder entre los dos sujetos de la relación contractual, los trabajadores solo

pueden prestar su libre consentimiento en circunstancias excepcionales, cuando la aceptación o el rechazo de una oferta no tiene consecuencias. Pues bien, el Derecho algorítmico parece haber limitado este papel. Dado que la regulación está basada en los riesgos y que considera algunos de ellos inaceptables, está conformada por normas imperativas que dejan menos margen a la voluntad de las personas afectadas.

El segundo aspecto es el insuficiente reconocimiento de la dimensión colectiva característica de las relaciones laborales tradicionales —de hecho, uno de los rasgos genéricos de la regulación laboral—. La magra presencia de esta dimensión se detecta en el Reglamento General de Protección de Datos y en el Reglamento de Inteligencia Artificial. Con todo, en ambas regulaciones pueden encontrarse referencias a las instituciones centrales del ámbito jurídico-laboral, entre ellas las organizaciones sindicales y la representación de los trabajadores en las empresas, a las que se reconoce cierto papel.

Las diferencias de cultura legal se constatan en determinados instrumentos utilizados en ambos reglamentos. Por ejemplo, la interpretación del término 'información' es diferente, y el concepto de 'transparencia' es utilizado en el Derecho del trabajo con otro significado y utilidad, especialmente en lo que respecta a las cuestiones retributivas y a ciertas formas de empleo.

4. COMPONENTES REGULATORIOS DEL DERECHO ALGORÍTMICO LABORAL

4.1. DERECHO DE LA UNIÓN EUROPEA

La Unión Europea es el ámbito regulador que más ha avanzado hasta el momento en la construcción del Derecho de la IA (Sierra Benítez, 2024). Su notable activismo se ha traducido en la aprobación de importantes textos normativos tanto de *soft law* como de *hard law*[4]. La tradicional atención a las cuestiones relacionadas con la protección de datos y el impulso de la Comisión y el Parlamento Europeos explican este interés temprano y sincero. El compromiso con la tutela de los derechos fundamentales ha jugado también su papel.

Un documento no vinculante de gran interés es la Declaración Europea sobre los Derechos y Principios Digitales para la Década Digital (2023/C 23/01). Se trata de un instrumento que puede encuadrarse en Derecho Digital y que incluye entre sus contenidos algunos puntos relativos a la IA. En el ámbito laboral, la Declaración parte de la premisa de que toda persona tiene derecho a unas condiciones de trabajo equitativas, justas, saludables y seguras, así como a una protección adecuada en el entorno digital y en el puesto de trabajo físico, con independencia de su situación laboral y de la modalidad o la duración del empleo.

A partir de este posicionamiento, este instrumento incorpora varios compromisos, dos de los cuales están directamente orientados a la tutela algorítmica:

– Garantizar que el uso de la inteligencia artificial en el lugar de trabajo sea transparente y siga un enfoque basado en los riesgos, y que se adopten las medidas de prevención correspondientes para mantener un entorno de trabajo seguro y saludable.
– Garantizar que las decisiones importantes que afecten a los trabajadores cuenten con supervisión humana y que, en general, se los informe de que están interactuando con sistemas de inteligencia artificial.

Los dos reglamentos que constituyen el núcleo del Derecho digital de la UE se ocupan de la IA, el primero de manera puntual, y el segundo monográficamente. El Reglamento General de Protección de Datos contiene una regulación de las decisiones basadas en procesos automatizadas (incluida la elaboración de perfiles) que tengan consecuencias o efectos jurídicos sobre las personas. Su artículo 15 dispone que el interesado tendrá derecho a obtener del responsable del tratamiento la confirmación de se están tratando o no datos personales que le conciernen y que, en tal caso, podrá ejercer el derecho de acceso a los datos personales, que incluirá la información sobre la existencia de decisiones automatizadas, incluida la elaboración de perfiles. En estos casos, el derecho incluye el de obtener información significativa sobre la lógica aplicada, así como sobre las posibles consecuencias de ese tratamiento para el interesado.

El RGPD también determina cuándo resulta admisible la utilización de estos instrumentos, pues la regla general que establece su artículo 22.1 es la siguiente: «Todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar». Sin embargo, la gestión algorítmica de personas resultaría admisible, toda vez que uno de los supuestos exceptuados de esta prohibición general es, precisamente, que la decisión sea necesaria para la celebración o la ejecución de un contrato entre el interesado y un responsable del tratamiento. En este caso, el RGPD exige que el responsable del tratamiento adopte las medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, obligación que, como mínimo, incluirá el derecho del interesado a exigir la intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión (Todolí Signes, 2024).

El Reglamento por el que se establecen normas armonizadas en materia de inteligencia artificial, que en la jerga comunitaria se conoce como «Reglamento de Inteligencia Artificial» o incluso como «Ley de la Inteligencia Artificial», es el cuerpo normativo que regula el tratamiento de la gestión algorítmica de personas (Rodríguez-Piñero, 2024). Las contingencias que pueden afectar a las personas trabajadoras son diversas, si bien la más relevante es el posible perjuicio que pueda ocasionar la utilización de sistemas de IA por parte de los empleadores en el acceso a un empleo o el desarrollo de una relación laboral.

El Reglamento se basa en el riesgo, y clasifica los sistemas de IA en cuatro niveles, entre los que destacan, por el detalle de su regulación, dos categorías: los sistemas prohibidos y los de alto riesgo. Entre las primeras, el artículo 5 incluye prácticas que podrían generar determinados sistemas de gestión de personas, entre ellos los sistemas orientados a inferir las emociones de una persona física en los lugares de trabajo y en los centros educativos (salvo

los que estén justificados por motivos médicos o de seguridad) (Muñoz Ruiz, 2024)[5]; y los sistemas de categorización biométrica que clasifiquen individualmente a las personas físicas sobre la base de sus datos biométricos para deducir o inferir, entre otros datos, su raza, sus opiniones políticas, su afiliación sindical, sus convicciones religiosas o filosóficas, su vida sexual o su orientación sexual. La prohibición alcanza tanto a la introducción de estos sistemas en el mercado como a su puesta en servicio para este fin específico y al uso de los propios sistemas de IA.

Para calificar un sistema de IA como de alto riesgo, se tienen en cuenta, entre otros criterios, la magnitud de sus consecuencias adversas para los derechos protegidos por la Carta de Derechos Fundamentales de la Unión Europea. Entre estos derechos se incluyen el derecho a la dignidad humana, el respeto de la vida privada y familiar, la protección de datos de carácter personal, la libertad de expresión y de información, la libertad de reunión y de asociación, la no discriminación, el derecho a la educación, la protección de los consumidores, los derechos de los trabajadores, los derechos de las personas discapacitadas, la igualdad entre hombres y mujeres, el derecho de propiedad intelectual, el derecho a la tutela judicial efectiva y a un juez imparcial, a la defensa y la presunción de inocencia, y el derecho a una buena administración. De esta manera, si en la aplicación de uno de estos sistemas se detecta un impacto negativo en los derechos de los trabajadores —sin ningún límite en la delimitación de estos—, la herramienta de IA quedará encuadrada en esta categoría con todas sus consecuencias.

El Anexo III identifica una serie de sistemas de IA de alto riego que están directamente relacionados con el empleo, la gestión de los trabajadores y el acceso al autoempleo. Estos son, concretamente:

- Los destinados a ser utilizados para la contratación o la selección de personas físicas, en particular para publicar anuncios de empleo específicos, analizar y filtrar las solicitudes de empleo y evaluar a los candidatos.
- Los destinados a ser utilizados para tomar decisiones que afecten a las condiciones de las relaciones de índole laboral o a la promoción o rescisión de relaciones contractuales de índole laboral, para la asignación de tareas a partir de comportamientos individuales o rasgos o características personales o para supervisar y evaluar el rendimiento y el comportamiento de las personas en el marco de dichas relaciones

En este Anexo aparecen los dos grandes espacios de las relaciones laborales en los que se utilizan instrumentos de IA: el proceso de contratación (que incluye tareas muy comunes como el *targeting* de candidatos o el *screening process* de currículos), y todas las actividades relacionadas con la gestión de las personas (promoción, retribución, ascenso, movilidad y despido).

Los considerandos del Reglamento justifican la inclusión de estas decisiones en el listado señalando que pueden afectar de un modo considerable a las futuras perspectivas laborales, así como a los medios de subsistencia y los derechos de los trabajadores. Asimismo, indican que estos usos específicos de la IA pueden perpetuar patrones históricos de discriminación durante todo el proceso de contratación y en la evaluación, promoción

o retención de personas en las relaciones contractuales de índole laboral[6]. Por último, los considerandos afirman que los sistemas de IA empleados para controlar el rendimiento y el comportamiento laborales pueden socavar sus derechos fundamentales a la protección de los datos personales y a la intimidad.

También se clasifican como de alto riesgo los sistemas vinculados con la educación y formación profesional, catalogación que se justifica porque, cuando no se diseñan y utilizan correctamente, estos sistemas pueden invadir especialmente y violar el derecho a la educación y la formación, y el derecho a no sufrir discriminación, además de perpetuar patrones históricos de discriminación. Adicionalmente, estos sistemas pueden decidir la trayectoria formativa y profesional de una persona y, en consecuencia, afectar a su capacidad para asegurar su subsistencia.

Aunque el listado de sistemas que serán considerados *per se* de alto riesgo del Anexo III es cerrado, el artículo 7 prevé la posibilidad de que la Comisión pueda modificarlo mediante actos delegados siempre que se cumplan estas dos condiciones:

i) Que los sistemas de IA estén destinados a ser utilizados en cualquiera de los ámbitos que figuran en el Anexo III, y

ii) que los sistemas de IA conlleven el riesgo de causar un perjuicio a la salud y la seguridad o de tener repercusiones negativas en los derechos fundamentales, y que dicho riesgo sea equivalente a, o mayor que, el riesgo de perjuicio o de repercusiones negativas que conlleven los sistemas de IA de alto riesgo que ya se mencionan en el Anexo III.

Para evaluar esta última condición, la Comisión tendrá en cuenta una serie de criterios, entre los que destacan dos por su potencial impacto en las relaciones laborales. Primero, que la utilización del sistema haya tenido repercusiones negativas en los derechos fundamentales o haya dado lugar a problemas importantes relacionados la probabilidad de que se produzcan ese perjuicio o esas repercusiones negativas. Segundo, que exista un desequilibrio de poder o que las personas que podrían sufrir ese perjuicio o esas repercusiones negativas se encuentren en una posición de vulnerabilidad respecto del responsable del despliegue de un sistema de IA debido a su situación, autoridad, conocimientos, circunstancias económicas o sociales, o edad[7].

El Reglamento hace un esfuerzo para dar entrada a la dimensión colectiva de las relaciones laborales, un aspecto que, como se ha señalado arriba, hasta el momento no ha recibido la atención que sería exigible en el Derecho digital (Sepúlveda Gómez, 2023: 181). Ya se ha indicado que la libertad sindical es objeto de una tutela especial, dado que el Reglamento prohíbe los sistemas de categorización basados en datos biométricos de las personas físicas que se usen para deducir o inferir su afiliación sindical.

Por otro lado, el Reglamento incluye a los sindicatos entre los sujetos que pueden participar en la elaboración de los códigos de buenas prácticas de las empresas. Estos códigos —en cuya elaboración se prevé también la participación de las organizaciones empresariales— deberán desarrollarse de manera inclusiva para garantizar su efectividad.

En el artículo 26 se impone a los responsables del despliegue de un sistema de IA de alto riesgo que sean empleadores la obligación de informar a los representantes de los trabajadores y a los trabajadores afectados de su exposición al mismo. Esta información deberá ser facilitada antes de poner en servicio o utilizar un sistema de IA de alto riesgo en el lugar de trabajo; cuando proceda, se hará con arreglo a las normas y procedimientos establecidos en el Derecho nacional y de la Unión y conforme a las prácticas en materia de información a los trabajadores y sus representantes. El interés de esta previsión es que impone un deber de información colectiva alejado de la lógica de la transparencia y el consentimiento informado que caracteriza al Derecho de protección de datos y de la IA.

La tercera norma que constituye el tronco del Derecho algorítmico del trabajo de la UE es la Directiva para la mejora de las condiciones laborales en plataformas digitales de trabajo (Todolí Signes, 2024*b*; Rojo Torrecilla, 2024: 228; Ponce del Castillo, 2023). Se trata de una norma extensa que, al igual que ley Rider española, regula fundamentalmente dos extremos: la calificación de los servicios de los trabajadores de estas empresas y la gestión algorítmica del personal. Por motivos obvios, solo me ocuparé del segundo.

Hay que señalar que la tramitación de la Directiva fue larga y compleja, y que en las últimas fases del proceso legislativo se redujo considerablemente su alcance a instancias del Consejo, lo que generó numerosas críticas. Los cambios afectaron sobre todo a la primera parte, mientras que la segunda, la que nos interesa, mantuvo gran parte de su contenido.

La regulación de la gestión algorítmica de estas personas se encuentra en el Capítulo III, que es, a mi juicio muy extenso y técnicamente acertado, mejor desde esta perspectiva que el propio Reglamento de Inteligencia Artificial. Este capítulo limita el tratamiento de datos mediante sistemas de control o de toma de decisiones automatizados, e impone una regla de transparencia en la utilización de estos sistemas (Todolí Signes, 2024*c*). También garantiza la evaluación del impacto de la actuación de la empresa sobre la protección de datos, así como la supervisión humana de los sistemas automatizados y de las decisiones tomadas con apoyo de estos. Por último, regula el derecho de información —individualizada en el caso de los trabajadores afectados, y colectiva en el de sus representantes—.

Siguiendo la estela de directivas anteriores (entre ellas, la de condiciones de trabajo transparentes y previsibles, y la de transparencia retributiva), el Reglamento incluye un Capítulo IV sobre transparencia en el trabajo a través plataformas, que garantiza que las plataformas declaren en trabajo que se les preste a las autoridades laborales nacionales; y que estas tengan acceso a una serie de informaciones al respecto.

Existen varias normas proyectadas que, aunque se encuentran distintas fases de elaboración, pueden completar la regulación ya aprobada. Me gustaría destacar que la comisión ya está barajando la posibilidad de aprobar una directiva específica sobre sistemas algorítmicos en el trabajo, norma cuya aprobación fue reclamada por la Confederación Europea de Sindicatos en una resolución aprobada el 6 de diciembre de 2022 (ETUC, 2022).

Junto a este paquete normativo, la Unión Europea se ha dotado de un importante aparato institucional para afrontar los numerosos desafíos que plantea la expansión del uso de la IA, y en su actuación está generando un nutrido corpus de materiales que resultan

de enorme utilidad para la aplicación del Derecho algorítmico de la Unión: el Comité Europeo de Protección de Datos (CEPD), el Grupo de Expertos de Alto Nivel sobre la Inteligencia Artificial, la Agencia Europea de los Derechos Fundamentales y el extinto Grupo de Trabajo del art. 29.

4.2. DERECHO INTERNO

En España disponemos de un importante documento de *soft law*, la Carta de Derechos Digitales, cuyas previsiones están brindando soporte a sucesivas reformas legales que han venido reconociendo esta categoría en nuestro país. La categoría de los derechos digitales es ya admitida de manera general y aparece reflejada en el ordenamiento (entre otros textos, en la Ley Orgánica de Protección de Datos y de Garantía de los Derechos Digitales[8]), hasta el momento de manera unitaria. No obstante, en el ámbito del Derecho del trabajo sería adecuado distinguir los derechos de primera generación, ya reconocidos en la ley (intimidad, desconexión digital, trabajo a distancia), y los de segunda generación, en vías de construcción, que serían fundamentalmente los que tienen que ver con la IA.

El apartado XIX de la citada carta está dedicado a los «Derechos en el ámbito laboral», y se reconocen algunos «en los entornos digitales y el teletrabajo». Estos se garantizan expresamente en el caso de la utilización de herramientas de monitoreo, la analítica, los procesos de toma de decisión en materia de recursos humanos y relaciones laborales, y, en particular, la analítica de redes sociales.

El texto añade que se informará a la representación legal de las personas trabajadoras del uso de tales dispositivos o herramientas, y que esta información deberá hacer referencia a los parámetros, las reglas y las instrucciones en los que se basan los algoritmos o sistemas de inteligencia artificial que afectan a la toma de decisiones capaces de incidir en las condiciones de trabajo, el acceso y mantenimiento del empleo, incluida la elaboración de perfiles.

Igualmente, se reconoce la garantía de derechos frente al uso por parte de la entidad empleadora de procedimientos de analítica de datos, inteligencia artificial y, en particular, los previstos en la ley en relación con la adopción de decisiones automatizadas durante los procesos de selección de personal.

Todos estos derechos se reconocen a las personas trabajadoras de los sectores público o privado, y se ejercitarán «con arreglo a la normativa vigente».

En el plano del *hard law*, creo que deberíamos comenzar con la trascendente Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación, que, como es sabido, ha actualizado de manera sustancial nuestro derecho antidiscriminatorio. Esta actualización se ha producido también en la faceta tecnológica, a través de la inclusión de mandatos orientados tutelar a las personas frente a las formas de discriminación emergentes. El artículo 3, al delimitar su ámbito objetivo de aplicación, dispone que la norma se aplicará a la «[…] inteligencia Artificial y gestión masiva de datos, así como otras esferas de análoga significación» (Vicente Palacio, 2023; Duro Carrión, 2023). El artículo 23 está dedicado a la inteligencia artificial y los mecanismos de toma de decisiones automatizados.

El precepto establece que las Administraciones públicas favorecerán la puesta en marcha de mecanismos para que los algoritmos involucrados en la toma de decisiones que se utilicen en el sector público tengan en cuenta criterios de minimización de sesgos, transparencia, rendición de cuentas y factibilidad técnica. En estos mecanismos se harán constar su diseño y sus datos de entrenamiento, así como su potencial impacto discriminatorio. Ello se hará en el marco de los principios establecidos por la Estrategia Nacional de Inteligencia Artificial, de la Carta de Derechos Digitales y de las iniciativas europeas en torno a la Inteligencia Artificial. Para lograr este fin, se promoverá la realización de evaluaciones de impacto que determinen el posible sesgo discriminatorio.

Asimismo, el artículo comentado prevé que, en el marco de sus competencias relativas a los algoritmos involucrados en procesos de toma de decisiones, las Administraciones priorizarán la transparencia en el diseño y la implementación y la capacidad de interpretación de las decisiones adoptadas por los mismos. Y añade que, con carácter general, tanto las Administraciones como las empresas promoverán el uso de una inteligencia artificial ética, confiable y respetuosa con los derechos fundamentales, siguiendo especialmente las recomendaciones de la Unión Europea. En este sentido se prevé también la promoción de un sello de calidad de los algoritmos.

La segunda norma nacional a la que debemos hacer referencia es la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales, dado que transpone el Reglamento General de Protección de Datos, si bien en la norma española las decisiones automatizadas no son objeto de gran atención. El artículo 14 de la ley prohíbe las decisiones basadas únicamente en el tratamiento automatizado —incluida la elaboración de perfiles— que produzcan efectos jurídicos negativos para el interesado o que le afecten significativamente, salvo que sean autorizadas expresamente por una norma con rango de ley o por el Derecho de la Unión Europea. En caso de que exista esta habilitación, la norma que faculte para el tratamiento deberá establecer las medidas adecuadas para salvaguardar los derechos y libertades del interesado, incluyendo el derecho a requerir la intervención humana en el proceso de revisión de la decisión adoptada.

Un hito en la construcción del Derecho algorítmico laboral fue la aprobación del Real Decreto-ley 9/2021, de 11 de mayo, por el que se modifica el texto refundido de la Ley del Estatuto de los Trabajadores, aprobado por el Real Decreto Legislativo 2/2015, de 23 de octubre, para garantizar los derechos laborales de las personas dedicadas al reparto en el ámbito de plataformas digitales, posteriormente convalidado mediante la Ley 12/2021, de 28 de septiembre. Esta norma dio cumplimiento a las previsiones del acuerdo adoptado previamente por una mesa de diálogo social estatal, y trataba de dar respuesta a una necesidad urgente y coyuntural, a saber, la lucha contra el fraude en la contratación a través de plataformas digitales. No obstante, se aprovechó su tramitación para introducir un mecanismo de garantía frente al uso de los algoritmos, aplicable no solo al sector del *delivery* o a las plataformas, sino también a todas las empresas: el deber de información colectiva sobre el uso de sistemas de IA. De acuerdo con la exposición de motivos, «[…] los algoritmos merecen nuestra atención y análisis, por los cambios que están introduciendo en la gestión de los servicios y actividades empresariales, en todos los aspectos de las condiciones de tra-

bajo y, sobre todo, porque dichas alteraciones se están dando de manera ajena al esquema tradicional de participación de las personas trabajadoras en la empresa»; en consecuencia, «[…] no podemos ignorar la incidencia de las nuevas tecnologías en el ámbito laboral y la necesidad de que la legislación laboral tenga en cuenta esta repercusión tanto en los derechos colectivos e individuales de las personas trabajadoras como en la competencia entre las empresas».

Aunque las aproximaciones doctrinales a esta norma se han privilegiado el análisis de la regulación de los servicios prestados a las plataformas de *delivery* (Castellano Burguillo, 2024; Rodríguez-Piñero Royo, 2024: 13-35), lo que ha motivado que se conozca generalmente como la «ley Rider», muchos hemos defendido que su denominación debería ser «ley Algoritmo». Su artículo único contiene una modificación del texto refundido de la Ley del Estatuto de los Trabajadores; concretamente, la adición una nueva letra d) en su artículo 64.4. Este inciso recoge los derechos de información y consulta de los órganos de representación de los trabajadores en la empresa, a los que la reforma añade el derecho «Ser informado por la empresa de los parámetros, reglas e instrucciones en los que se basan los algoritmos o sistemas de inteligencia artificial que afectan a la toma de decisiones que pueden incidir en las condiciones de trabajo, el acceso y mantenimiento del empleo, incluida la elaboración de perfiles» (Fernández García, 2023: 159; Ginés i Fabrellas, 2024*b*).

El deber empresarial de informar no se limita a las empresas de plataforma, sino que afecta a todas las empresas que tienen constituidos órganos de representación legal de sus plantillas de acuerdo con el Estatuto de los Trabajadores tanto en el sector público como en el privado (aunque es preciso señalar no se ha reformado de manera paralela el Estatuto Básico del Empleado Público). Su alcance es muy amplio, dado que alcanza a todos los algoritmos incluidos en su definición, es decir, a los que afectan a la toma de decisiones que pueden incidir en los aspectos más relevantes de las relaciones laborales, que son las condiciones de trabajo y el acceso y mantenimiento del empleo. La exposición de motivos señala expresamente que una de las utilidades más comunes de los instrumentos de gestión algorítmica es la elaboración de perfiles que afectan a la toma de decisiones que pueden incidir en los aspectos más relevantes de las relaciones laborales, que son las condiciones de trabajo y el acceso y mantenimiento del empleo.

Por otro lado, el deber que se establece tiene por objeto exclusivamente la información, lo que supone la transmisión unilateral de datos de la empresa a los representantes, sin que estos deban emitir informe u opinión alguna. Se trata de un deber de «información colectiva» que, como se ha apuntado, es ejercitable solo en aquellos centros de trabajo en los que se han elegido delegados de personal o comités de empresa. No se han buscado soluciones alternativas para aquellos centros en los que no hay delegados, como sí se ha hecho en otros casos —por ejemplo, la negociación de los planes de igualdad o los procedimientos de consulta en las decisiones empresariales de efectos colectivos—. La Directiva sobre trabajo en plataformas sí prevé el deber de información individual al trabajador en el caso de que no existan representantes a quienes comunicar la información a nivel colectivo.

Otra pieza de nuestro Derecho algorítmico laboral es la Ley 3/2023, de 28 de febrero, de Empleo (Román López, 2023: 181). Su artículo 17 se ocupa de la toma de decisiones

fundamentada en el análisis de datos, las evidencias estadísticas y el análisis del mercado de trabajo. El precepto prevé la utilización de la información contenida en el Sistema Público Integrado de Información de los Servicios de Empleo para el diseño de herramientas de apoyo a las decisiones fundamentadas en el análisis de datos y las evidencias estadísticas. Estas herramientas se desarrollarán en el seno del Sistema Nacional de Empleo con la colaboración de las Comunidades Autónomas, se pondrán a disposición de las personas responsables de la tutorización y seguimiento individualizado de las personas demandantes de los servicios de empleo o de las personas responsables de la prestación de servicios a personas, empresas o entidades empleadoras usuarias para facilitar su toma de decisiones y las recomendaciones que elaboren y emitan en el ejercicio de sus funciones. Asimismo, contempla el futuro desarrollo reglamentario del conjunto de instrucciones que compondrán el algoritmo subyacente a las decisiones y recomendaciones basadas en el análisis de datos y las evidencias estadísticas. Estas herramientas no podrán utilizar datos personales pertenecientes a las categorías especiales que regula el artículo 9.1 del Reglamento General de Protección de Datos.

De la misma manera, la ley dispone que las decisiones y recomendaciones basadas en el análisis de datos y en las evidencias estadísticas siempre podrán ser revisadas o modificadas por el personal responsable de la tutorización y seguimiento individualizado de cada persona demandante de servicios. También garantiza el acceso de los distintos usuarios al contenido íntegro de estas decisiones y recomendaciones.

Ya se indicó *supra* que la IA se ha convertido en un instrumento para apoyar la actuación de los poderes públicos en el mercado de trabajo y las relaciones laborales. En nuestro Derecho, el Real Decreto-ley 2/2021, de 26 de enero, de refuerzo y consolidación de medidas sociales en defensa del empleo modificó con su disposición final cuarta la redacción del párrafo a) del artículo 53.1 del Real Decreto legislativo 5/2000, la Ley sobre Infracciones y Sanciones en el Orden Social para incluir la posibilidad de «[…] actas extendidas en el marco de actuaciones administrativas automatizadas».

En paralelo, el Real Decreto 688/2021, de 3 de agosto, modificó el Reglamento general sobre procedimientos para la imposición de sanciones por infracciones de orden social y para los expedientes liquidatorios de cuotas de la Seguridad Social, aprobado por el Real Decreto 928/1998, con el fin de introducir un nuevo capítulo IX sobre el procedimiento sancionador promovido por actuación administrativa automatizada en el ámbito de la Administración General del Estado.

En este sentido, merece la pena subrayar que el artículo 1.1 de la Ley 23/2015, de 21 de julio, Ordenadora del Sistema de Inspección de Trabajo y Seguridad Social contiene la siguiente definición legal: «El Sistema de Inspección de Trabajo y Seguridad Social está constituido por el conjunto de principios legales, normas, órganos, personal y medios materiales, incluidos los informáticos, que contribuyen al adecuado cumplimiento de la misión que tiene encomendada, según lo establecido en la presente ley», y que el texto incluye una referencia expresa al *software* y al *hardware* que no aparecía en la ley anterior.

Una norma nonata, y que, por lo que parece, ya no llegará a aprobarse se encontraba en el Anteproyecto de Ley por la que se modifica el texto refundido de la Ley del Estatuto

de los Trabajadores y otras disposiciones en materia laboral, para la transposición de la Directiva (UE) 2019/1152 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativa a unas condiciones laborales transparentes y previsibles en la Unión Europea. Las primeras versiones de este Anteproyecto contemplaban la introducción de un nuevo artículo (el 8 bis) en el Texto Refundido del Estatuto de los Trabajadores que establecería la obligación para los empleadores de entregar información adicional a la ya prevista en el artículo 8 en aquellos casos en los que se produjera la utilización de sistemas automatizados de toma de decisiones, de seguimiento o de vigilancia. La obligación de entregar información adicional prevista por el artículo que se pretendía adicionar tenía por objeto garantizar el ejercicio de las facultades de dirección, gestión y control. El precepto también preveía que reglamentariamente se determinaría la información adicional que debería ser suministrada, y disponía que esta debería trasladarse de forma transparente, en un formato estructurado y de uso común (Rodríguez-Piñero, 2024).

En el caso de que se hubiera aprobado aquel Anteproyecto, habría significado el reconocimiento de un derecho de información individual sobre el uso de algoritmos que se sumaría a los ya existentes, es decir, al de carácter colectivo, introducido en el artículo 64 del Estatuto por la ley Rider, y al individual, contemplado por el RGPD. Sin embargo, el texto finalmente aprobado por el Consejo de Ministros como proyecto de ley y enviado para su tramitación al Parlamento ha obviado esta reforma, ya que no aparece entre las medidas a adoptar.

Fuera del espacio de regulación tradicional puede destacarse el papel de la Agencia Española de Protección de Datos, que admite la utilización de estas herramientas en la selección de personal con una serie de salvaguardas: que se prevea algún mecanismo de intervención humana significativa si la persona afectada así lo solicita; que exista algún cauce para que esta persona exprese su opinión e impugne la decisión; que se garantice el derecho a recibir una explicación de la decisión tomada después; y que se realice una evaluación de impacto en el diseño e implementación del algoritmo (AEPD, 2021).

4.3. NEGOCIACIÓN COLECTIVA

Hablar de la regulación laboral es hablar de la negociación colectiva en sentido amplio, pues la regulación autónoma es una seña de identidad del sistema de relaciones laborales y del Derecho que lo regula. Resulta indispensable, por ello conocer cómo está contribuyendo a la ordenación de la gestión algorítmica (OCDE, 2023: 221).

La Declaración Europea sobre los Derechos y Principios Digitales afirma en este sentido que «[…] las organizaciones sindicales y patronales desempeñan un papel importante en la transformación digital, en particular en lo relativo a la definición de unas condiciones de trabajo justas y equitativas, también en lo que respecta al empleo de herramientas digitales en el trabajo». También recoge el compromiso de garantizar en estos entornos digitales el derecho de negociación y acción colectiva. En la práctica, se ha desarrollado un Derecho digital de origen convencional (Baylos Grau, 2019; Amaadachou Kaldur, 2023; Gómez García, 2023), que desarrolla y complementa al estatal, centrado fundamental-

mente en la garantía de los derechos digitales (especialmente, el de desconexión digital) y en el teletrabajo.

Una primera manifestación, negociada y aplicada en el ámbito de la UE, fue el Acuerdo Marco Europeo de Digitalización, del año 2021(Goerlich Peset, 2021: 49; Sepúlveda Gómez, 2023: 213). Este documento, precursor el tratamiento normativo de la IA, partía de la premisa de que las empresas de Europa se encontraban en ese momento en las primeras fases de utilización de la IA para optimizar los procesos de trabajo o crear nuevos modelos de negocio, y de que era esencial explorar las opciones de diseño para el éxito económico y las buenas condiciones de trabajo, de tal manera que no pusieran en peligro, sino que incrementaran, la participación y la capacidad de los agentes sociales en el trabajo.

Este acuerdo autónomo de los interlocutores sociales establece algunas directrices y principios sobre cómo y en qué circunstancias se introduce la IA en el mundo laboral, incluyendo elementos que luego han sido comunes en el Derecho algorítmico como el control humano, la transparencia y la prevención del daño.

Los interlocutores sociales europeos convinieron igualmente en que una IA fiable tiene tres componentes que deben cumplirse a lo largo de todo el ciclo de vida del sistema y ser respetadas en su despliegue en el mundo laboral: ser lícita, justa, transparente, segura y fiable, cumplir con todas las leyes y reglamentos aplicables, así como con los derechos fundamentales y las normas de no discriminación; seguir las normas éticas acordadas, asegurando el respeto de los derechos humanos y fundamentales de la UE, la igualdad y otros principios éticos; y, finalmente, ser robusta y sostenible tanto desde el punto de vista técnico como social.

Las organizaciones sindicales españolas eran conscientes de la gravedad del asunto, así como del papel que debían asegurarse en su ordenación. Por ello, promovieron la introducción en los convenios colectivos de reglas específicas sobre el uso de IA por las empresas. Además de hacerlo en aquellas unidades de negociación en las que les resultó posible, promovieron esta inclusión a través del instrumento más potente del que disponen para el gobierno de la negociación colectiva: el acuerdo de ordenación previsto en el artículo 83 TRET. Así, el V Acuerdo para el Empleo y la Negociación Colectiva, firmado el 10 de mayo de 2023, contiene un capítulo (el XVI) dedicado a la transición tecnológica, digital y ecológica en el que aparece un apartado monográfico sobre esta cuestión, el 3.º, titulado «Inteligencia Artificial (IA) y garantía del principio de control humano y derecho a la información sobre los algoritmos».

El punto de partida del Acuerdo es que «[…] la IA de manera progresiva tendrá un impacto significativo en el mundo laboral y, si no se hace un uso correcto y transparente, podría llevar a adoptar decisiones sesgadas o discriminatorias relativas a las relaciones laborales». En relación con esta cuestión, el texto afirma que la negociación colectiva debe desempeñar un papel fundamental en el establecimiento de criterios que garanticen el uso adecuado de la IA y el desarrollo del deber de información periódica a la representación de los trabajadores.

Las organizaciones firmantes del V Acuerdo para el Empleo y la Negociación Colectiva toman como referente el Acuerdo Marco Europeo sobre digitalización, y consideran prio-

ritario adecuarlo a cada ámbito negocial. A tal efecto, los convenios colectivos deben, entre otras cosas, promover un enfoque orientado a las personas, particularmente en el uso de sistemas de inteligencia artificial seguros y transparentes.

Por otra parte, se acuerda que el despliegue de sistemas de IA en las empresas deberá seguir el principio de control humano respecto a la IA y ser seguro y transparente. En consecuencia, las empresas deberán facilitar a los representantes legales de su plantilla información transparente y entendible sobre los procesos basados en la IA en los procedimientos de recursos humanos (contratación, evaluación, promoción y despido) y garantizarán que no existen prejuicios ni discriminaciones.

La práctica de la negociación colectiva española presenta un tratamiento todavía muy escaso de estas materias, rasgo que, por lo demás, es frecuente en los sistemas de relaciones laborales cercanos al nuestro (Rodríguez Fernández, 2024: 217; Álvarez Cuesta, 2024). Este limitado desarrollo se explica no solo por la novedad de los problemas a afrontar, sino sobre todo por el retraso crónico de nuestros convenios en adaptarse a cuestiones emergentes, a pesar de que la voluntad de los Acuerdos Interconfederales ha sido promover su tratamiento.

Con todo, y como siempre ocurre en España, podemos presumir de algún convenio particularmente avanzado en su tratamiento de la IA. Al igual que antes ocurrió con otras materias, este tipo de convenios están apareciendo en el sector bancario. Ejemplo ma de ello es el XXIV Convenio colectivo del sector de la banca 2019-2023, que contiene un importante y potente artículo 80 sobre derechos digitales cuyo un apartado 5 («Derecho ante la inteligencia artificial») reconoce a las personas trabajadoras los siguientes derechos:

- el derecho a no ser objeto de decisiones basadas única y exclusivamente en variables automatizadas, salvo en aquellos supuestos previstos por la Ley,
- el derecho a la no discriminación en relación con las decisiones y procesos cuando ambos estén basados únicamente en algoritmos, y
- el derecho a solicitar el concurso e intervención de las personas designadas a tal efecto por la Empresa en caso de discrepancia.

También se reconoce el derecho de información algorítmica de los representantes legales de los trabajadores en términos muy similares a los previstos en el TRET, derecho que alcanzando los datos que nutren los algoritmos, la lógica de funcionamiento y la evaluación de los resultados.

Una regulación similar se encuentra en el artículo 35 del Convenio colectivo para los establecimientos financieros de crédito de 2021.

Fuera del sector bancario, destaca el Convenio colectivo del sector de grandes almacenes 2021-2022, firmado el 24 de marzo de 202, que prevé la creación de un observatorio sectorial encargado, entre otras funciones, de realizar análisis conjuntos de los cambios que la digitalización en este tipo de empresas, el establecimiento de Protocolos para la transición digital y los cambios organizativos, conectados con los objetivos y funciones del Observatorio sectorial, así como la articulación de marcos innovadores de diálogo social

(*SandBox*, observatorios o laboratorios). El convenio afirma expresamente que el observatorio prestará especial atención a «[…] la utilización de algoritmos que incidan en las condiciones de trabajo»[9].

Un caso especial es el Convenio colectivo Takeaway Express Spain, S.L. (Just Eat) 2021-2023. Fue el primero acordado en el sector de las plataformas de reparto de comida a domicilio por uno de sus grandes operadores —que, a diferencia de los otros, había optado por un modelo de contratación laboral de los repartidores—. En coherencia con este modelo, la empresa y los representantes de los trabajadores procedieron a firmar este convenio, que incluye un apartado monográfico sobre el derecho de información en el caso de utilización de algoritmos y sistemas de inteligencia artificial. En este apartado se desarrolla el derecho de información previsto en el artículo 64.4.d) TRET, y se especificán que se informará sobre el su uso para la toma de decisiones en materia de personal «[…] siempre y cuando dichas decisiones puedan incidir en las condiciones de trabajo, el acceso y mantenimiento de empleo de las personas trabajadoras, incluida la elaboración de perfiles». Esta información incluirá tanto los parámetros y los datos, como de las reglas las instrucciones que los nutren. También prevé un límite a este deber: no se facilitará información que tenga la protección que le confiere la normativa vigente. El convenio dispone que la Empresa garantizará cierto grado de supervisión humana, y que no tomará en consideración datos que puedan dar lugar a la vulneración de derechos fundamentales. Finalmente, los firmantes acuerdan la creación de una «Comisión Algoritmo», órgano paritario a través de la cual se canalizará el derecho de información algorítmica.

Por último, cabe destacar que el texto convencional prevé que tanto la Comisión Algoritmo como cada persona trabajadora de la Empresa sean informados de forma adecuada de todas las herramientas digitales que existen en su entorno laboral y de los posibles riesgos para la salud asociados a las mismas; que cuando las personas trabajadoras establezcan una conversación, deberán ser informadas previamente de si están hablando con un *chatbot* o con una persona, y que las conversaciones mantenidas con un *chatbot* no podrán ser utilizadas para sancionar a una persona[10].

Otro sector en el que la negociación colectiva ha abordado las implicaciones del uso de la IA es el de la producción audiovisual. Así, la disposición adicional cuarta del III Convenio colectivo de ámbito estatal de la industria de producción audiovisual (técnicos), suscrito el 13 de diciembre de 2023, señala que sus firmantes se comprometen a «[…] realizar un uso transparente, razonado y ético de la inteligencia artificial, conforme a los criterios y reglas que se establezcan en la normativa de aplicación». Este convenio adopta un enfoque distinto en el tratamiento de esta materia, el de las declaraciones generales, que seguramente será el que sigan muchas de nuestras unidades de negociación en el futuro.

El interés de las organizaciones sindicales en esta cuestión puede constatarse también en el documento «Recomendaciones para la negociación colectiva de la IA», elaborado por el Gabinete de Estudios de la confederación sindical UGT[11]. Por su parte, CC.OO. ha publicado una *Guía para la negociación colectiva en materia de derechos digitales* que incluye una sección titulada «Algoritmos, decisiones automatizadas e inteligencia artificial»[12].

5. REFLEXIONES CONCLUSIVAS

En su «Protágoras», Platón narra que la ley tiene su origen en el robo del fuego por Prometeo, que se lo dio a los hombres para compensarles por la falta de dones que su hermano Epimeteo les había concedido en la creación. Con el fuego les ayudó a crear la civilización, pero la civilización del hombre prometeico era puramente técnica, por lo que degeneró en violencia y destrucción. La humanidad estuvo a punto de perecer por obra de sus propias obras, y el ingenio humano no encontraba una solución a este problema circular. Zeus, que quería preservar al hombre para una misión más alta, le envío un don que le salvaría de la autodestrucción, el don de la ley y la justicia (Jaeger, 2018: 56).

Este mito resulta muy sugerente para explicar en qué situación se encuentra el Derecho en general, y el que se ocupa del trabajo en particular, ante el desarrollo de la IA, que es rápido, intenso y ubicuo. La IA es un regalo para la humanidad, como lo fue el fuego de acuerdo con la mitología; un don divino capaz de cambiar las cosas y de mejorar nuestras vidas. Pero también es un regalo lleno de peligros, un obsequio envenenado si no somos capaces de controlar. La humanidad del siglo XXI no es la misma a la que ayudó al titán, dado que ahora conocemos las leyes. Está formada por personas que tratamos con fuego y con normas, que elaboramos normas para controlar el fuego, y que gracias a ello hemos sido capaces, mal que bien, de integrar la tecnología en nuestras vidas sin habernos destruido, aunque siempre parezca que estamos a punto de hacerlo.

Es el momento de las leyes, de regular esta nueva realidad para que se desarrolle de acuerdo con nuestros valores y nuestros principios, pensando en las personas. Ordenar no es incompatible con el desarrollo tecnológico, antes bien puede facilitarlo, al prevenir problemas que podrían ser obstáculos para su avance. La regulación no impide el desarrollo, sino que, como entre nosotros ha demostrado recientemente Todolí Signes (2021), lo promueve, consideración que es extensible al Derecho algorítmico laboral, en construcción aún. En el caso de la IA, tecnología *in statu nascendi* y que tiene unos potenciales impactos económicos y sociales sin precedentes, la relevancia de ambos objetivos, controlar y favorecer, es mayor aún si cabe que en otras revoluciones tecnológicas.

Como señala el Reglamento de Inteligencia Artificial, nuestro objetivo debe ser la adopción de una IA centrada en el ser humano y fiable. Por ello, para contrarrestar los eventuales efectos perjudiciales de los sistemas de IA, debemos garantizar un elevado nivel de protección de la salud, la seguridad y los derechos fundamentales, y, en particular, la democracia, el Estado de Derecho y el medioambiente. A su vez, es necesario brindar apoyo a la innovación para que esta herramienta pueda desarrollarse y lograr el impacto beneficioso que muchos esperamos. La Estrategia Española de Inteligencia Artificial plantea como uno de sus ejes el desarrollo de una IA transparente, responsable y humanística[13].

Los iuslaboralistas tenemos una visión sesgada de la realidad jurídica, puesto que en nuestra programación la tutela de la parte débil en las relaciones laborales condiciona la manera de elaborar, interpretar y aplicar el Derecho. Cualquier regulación de la IA debería, según nuestro criterio, centrarse en la garantía de los derechos, puesto que el avance en sus capacidades nos genera poco interés, cuando no abierta desconfianza. Pero tenemos

que aceptar la realidad de los cambios tecnológicos, y de la revolución que estos están suponiendo. Por ello, los sistemas de IA deben estar en el centro de nuestra atención como juristas.

Frente a la tecnología siempre se puede ser optimista y pesimista. Personalmente creo que debemos serlo, lo uno o lo otro, no tanto respecto de la IA como de las personas y organizaciones que van a utilizarla, lo que nos incluye a todos nosotros como ciudadanos, consumidores, profesionales, usuarios… En este aspecto, la revolución de la IA no se diferencia tanto de las anteriores que hemos experimentado. Es indudable que los juristas tenemos nuestra cuota de responsabilidad a la hora de lograr una integración adecuada, justa y respetuosa los derechos de la Inteligencia Artificial. O de la pseudointeligencia de la que hace ya algún tiempo hablara Stephenson.

NOTAS

1. Para un completo, bien informado y actualizado estudio, *vid*. Ginès i Fabrellas (2024), una de las grandes expertas en el tema del impacto de la IA en las relaciones laborales.

2. Los efectos de estas tecnologías disruptivas en la seguridad y salud en el trabajo han sido objeto de numerosos estudios; entre otros, Mandinaud y Ponce del Castillo (2024); Pizzi y Payá Castiblanque (2023).

3. Luque Parra (2023: 305) destaca la contradicción entre ser factor de riesgo y sistema de protección.

4. No está de más destacar que durante la última presidencia española de la UE, que se prolongó desde el 1 de julio al 31 de diciembre de 2023, hubo avances significativos en dos de las normas más relevantes en esta materia: el Reglamento de Inteligencia Artificial y la Directiva sobre trabajo en plataformas. En el caso concreto de la Directiva de trabajo en plataformas, su elaboración se inspiró directamente en la regulación española, que había sido pionera en Europa.

5. Esta misma profesora es autora de una excelente monografía sobre este particular; *vid.* Muñoz Ruiz, (2024).

6. Ejemplos de colectivos potencialmente perjudicados son las mujeres, ciertos grupos de edad, las personas con discapacidad o las personas de orígenes raciales o étnicos concretos o con una orientación sexual determinada.

7. El articulado del Reglamento realiza otras diferenciaciones entre estos sistemas. Así, dentro de los de alto riesgo distingue aquellos que, a pesar de ser conformes con el Reglamento, presentan un riesgo para la salud o la seguridad de las personas, para los derechos fundamentales de las personas o para otros aspectos de la protección del interés público, e impone obligaciones específicas para estos. También se diferencian los sistemas destinados a interactuar directamente con personas físicas; los de reconocimiento de emociones; los de categorización biométrica; los de uso general; los que generen contenido sintético de audio, imagen, vídeo o texto; los que generen o manipulen imágenes o contenidos de audio o vídeo que constituyan una ultrafalsificación; o los que presenten un riesgo.

8. Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales; BOE núm. 294, de 06/12/2018.

9. En el siguiente Convenio colectivo de este sector, firmado el 10 de marzo de 2023, esta materia ha desaparecido de las prioridades del Observatorio sectorial.

10. Curiosamente, el posterior Convenio colectivo de elaboradores de productos cocinados para su venta a domicilio, suscrito el 7 de junio de 2022 y de aplicación en el mismo sector del *delivery* no contiene previsión alguna sobre este aspecto de la gestión de personas.

11. Disponible en: <https://servicioestudiosugt.com/recomendaciones-para-la-negociacion-colectiva-de-la-ia/>. [Consulta: 12/06/2024.]

12. Disponible en: <https://www.pv.ccoo.es/f5e225d35ddc46c32b466f06634138eb000053.pdf>. [17/06/2024].

13. *Estrategia de Inteligencia Artificial 2024*. Disponible en: https://portal.mineco.gob.es/RecursosNoticia/mineco/prensa/noticias/2024/Estrategia_IA_2024.pdf.

BIBLIOGRAFÍA

AFI (2024): «Impacto de la IA en la economía española. Ocupaciones, sectores y CC.AA.» [en línea] <https://www.afi.es/2024/05/20/inteligencia-artificial-economia-sectores-ocupaciones-comunidades-autonomas/>. [Consulta: 07/02/2024.]

ALVAREZ CUESTA, Henar (2024): «Experiencias convencionales de regulación del impacto laboral de la inteligencia artificial y su uso con fines de control», *Trabajo y derecho: nueva revista de actualidad y relaciones laborales*, 19.

ASQUERINO LAMPARERO, Maria José (2022*a*): «Algoritmos y discriminación», en J.M Morales Ortega (coord.), «*Realidad social y discriminación: Estudios sobre diversidad e inclusión*», Murcia: Laborum.

— (2022*b*): «Algoritmos, procesos de selección y reputación digital: una mirada antidiscriminatoria», en AA.VV., «*Digitalización, recuperación y reformas sociales: XXXII Congreso Anual de la Asociación Española de Derecho del Trabajo y de la Seguridad Social*», Madrid: Ministerio de Trabajo y Economía.

BAIOCCO, Sara, Enrique FERNÁNDEZ-MACÍAS, Uma RAINI y Annarosa PESOLE (2022): *The Algorithmic Management of work and its implications in different contexts*, Seville: European Comission.

BAYLOS GRAU, A. (2019), «El papel de la negociación colectiva en la ley de protección de datos personales y garantía de derechos digitales en España», *Labour & Law Issues*, 5 (1), 3-14.

— (2022): «La larga marcha hacia el trabajo formal: el caso de los riders y la ley 12/2021», *Cuadernos de relaciones laborales*, 40(1), 95-113.

CALVO GALLEGO, Francisco Javier (2022): «Algoritmo y discriminación», *Revista Latinoamericana de Derecho Social*, Extra1.

DEL REY GUANTER, Salvador (2023): «Inteligencia artificial, algortimos y big data en el ámbito laboral. Consideraciones en torno a su marco regulatorio», en A. Ginés i Fabrellas (coord.), *Algoritmos, Inteligencia artificial y relación laboral*, Cizur Menor: Thomson Reuters Aranzadi.

DURO CARRIÓN, Susana (2023): «Aspectos fundamentales de la igualdad de trato y no discriminación en la educación, medios de comunicación social y publicidad, internet. y redes sociales, inteligencia artificial y mecanismos de toma de decisión automatizados», en A. Sempere Navarro y B. García Gil (dirs.), *Una visión transversal del derecho a la igualdad: Ley 15/2022, de 12 de julio*, Las Rozas: SEPIN.

FERNÁNDEZ GARCIA, Antonio (2023*a*): «Tecnología y discriminación», en S. Guindo Morales y P. Ortega Lozano (coords.), *El desafío tecnológico en el Derecho del Trabajo en la era de la cuarta revolución industrial*, Barcelona: Atelier.

— (2023*b*): «Los algoritmos y la inteligencia artificial en la Ley 12/2021, de 28 de septiembre», en J. Moreno Gené y A. Romero Burillo (coords.), *Los nuevos escenarios laborales de la innovación tecnológica*, Valencia: Tirant Lo Blanch.

GARCÍA COCA, Olga (2022): «La intermediación laboral digital como causa emergente de discriminación», en J.M. Morales Ortega (coord.), *Realidad social y discriminación: Estudios sobre diversidad e inclusión*, Murcia: Laborum.

GIL DE ALBURQUERQUE, Román (2023): «Inteligencia artificial y trabajo por cuenta ajena: algunas reflexiones a la altura de 2023», en S. Guindo Morales y P. Ortega Lozano (dirs.), *El desafío tecnológico en el Derecho del Trabajo en la era de la cuarta revolución industrial*, Barcelona: Atelier.

GINÈS I FABRELLAS, Anna (2022): «La gestión algorítmica del trabajo: nuevos retos jurídicos, tecnológicos y éticos», en AA.VV., *Digitalización, recuperación y reformas sociales: XXXII Congreso Anual de la Asociación Española de Derecho del Trabajo y de la Seguridad Social*, Madrid: Ministerio de Trabajo y Economía Social.

— (2024*a*): «Algoritmos sesgados en el trabajo. Consideraciones entorno a su tratamiento jurídico», *Trabajo y derecho: nueva revista de actualidad y relaciones laborales*, 19.

— (2024*b*): «El derecho a conocer el algoritmo: una oportunidad perdida de la «Ley Rider»», *Iuslabor*, 2, 1-5.

GOERLICH PESET, José María (2021): «El Acuerdo Marco Europeo sobre Digitalización», *Documentación Laboral*, 122, 49-57.

— (2021*b*): «Decisiones administrativas automatizadas en materia social: algoritmos en la gestión de la Seguridad Social y en el procedimiento sancionador», *Labos: Revista de Derecho del Trabajo y Protección Social*, 2(2), 22-42.

GÓMEZ GARCIA, Francisco Xabiere (2023): «La recepción de las tecnologías digitales en el diálogo social y la negociación colectiva de la Unión Europea», en S. Guindo Morales y P. Ortega Lozano (dirs.), *El desafío tecnológico en el Derecho del Trabajo en la era de la cuarta revolución industrial*, Barcelona: Atelier.

HAYS (2024): *Guía del Mercado Laboral 2024* [en línea] <https://guiasalarial.hays.es/>. [Consulta: 09/01/2024.]

LLORENS ESPADA, Julen (2024): «La inteligencia artificial para la mejora de la seguridad y salud laboral y su encaje en el marco regulatorio europeo», *Trabajo y derecho: nueva revista de actualidad y relaciones laborales*, 19.

LUQUE PARRA, Manuel (2023): «IA y seguridad laboral: la dicotomía entre ser un gran aliado productivo y un riesgo laboral emergente», en A. Ginès i Fabrellas (coord.), *Algoritmos, Inteligencia artificial y relación laboral*, Cizur menor: Thomson Reuters Aranzadi.

MERCADER UGUINA, Jesús R. (2022): «La gestión laboral a través de algoritmos», en AA.VV. (2022), *Digitalización, recuperación y reformas sociales: XXXII Congreso Anual de la Asociación Española de Derecho del Trabajo y de la Seguridad Social*, Madrid, Ministerio de Trabajo y Economía Social.

OCDE (2023): *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*, Paris: OECD Publishing.

OIT (2023): *Generative AI and jobs: A global analysis of potential effects on job quantity and quality*, Informe n.º 96.

MANDINAUD, Vicent y Aida PONCE DEL CASTILLO, A. (2024): «AI systems, risks and working conditions», en A. Ponce del Castillo (ed.), *Artificial intelligence, labour and society*, Bruselas: ETUI, 237-249.

MUÑOZ RUIZ, Ana Belén (2024): *Biometría y sistemas automatizados de reconocimiento de emociones: Implicaciones Jurídicos-Laborales*, Valencia: Tirant Lo Blanch.

PÉREZ ANAYA, Rosa María (2022): «Algoritmo, selección de personal y protección de datos», en I. Marín Aloso, M. T. Igartua Miró y C. Solís Prieto (dirs.), *Digitalización, desarrollo tecnológico y Derecho del Trabajo: nuevas perspectivas de sostenibilidad*, Cizur Menor: Thomson Reuters Aranzadi.

PIZZI, Alejandro y Raúl PAYÁ CASTIBLANQUE (2023): «El impacto de las tecnologías digitales sobre los riesgos laborales de los trabajadores en Europa», en J. Moreno Gené y A. Romero Burillo (coords.), *Los nuevos escenarios laborales de la innovación tecnológica*, Valencia: Tirant Lo Blanch.
PWC (2024): *Barómetro global de la Inteligencia Artificial en el empleo 2024* [en línea] <https://www.pwc.es/es/consultoria/inteligencia-artificial/ai-jobs-barometer.html>. [Consulta: 01/06/2024.]
RIVAS VALLEJO, Maria Pilar (ed.), (2022): *Discriminación algorítmica en el ámbito laboral*, Cizur Menor: Thomson Reuters Aranzadi.
RODRÍGUEZ FERNÁNDEZ, María Luz (2024): «Collective Bargaining and AI in Spain», en A. Ponce Del Castillo (ed.), *Artificial intelligence, labour and society*, Bruselas: ETUI, 217-227.
RODRÍGUEZ-PIÑERO ROYO, Miguel (2022a): «El nuevo Derecho Digital y los problemas de discriminación», *Noticias CIELO*, 4.
— (2022b), «Acceso al empleo, formación y contratación en el contexto en el contexto de la digitalización», en AA.VV., *Digitalización, recuperación y reformas sociales: XXXII Congreso Anual de la Asociación Española de Derecho del Trabajo y de la Seguridad Social*, Madrid: Ministerio de Trabajo y Economía Social.
— (2023): «La Ley Rider dos años después: enseñanzas de una experiencia particular», *Revista de Estudios Jurídico Laborales y de Seguridad Social (REJLSS)*, 7, 13-35.
RODRÍGUEZ SANZ DE GALDEANO, Beatriz (2024): «La responsabilidad empresarial por accidentes vinculados a la inteligencia artificial», *Trabajo y derecho: nueva revista de actualidad y relaciones laborales*, 19.
ROMÁN LÓPEZ, Jaime (2023): «La digitalización en la Ley de Empleo: empleo, colocación e intermediación», *Trabajo, Persona, Derecho, Mercado*, 181-206.
SÁNCHEZ LÓPEZ, Celia (2023): «Los perfilados estadísticos del desempleo en los servicios públicos de empleo. Un acercamiento a los casos de España y Andalucía», *Trabajo, Persona, Derecho, Mercado*, 63-85.
SANGUINETI RAYMOND, Wilfredo y Ana Belén MUÑOZ RUÍZ (2024): «Regular el uso de la inteligencia artificial para evitar la deshumanización y proteger a las personas que trabajan», *Trabajo y derecho: nueva revista de actualidad y relaciones laborales*, 19.
SÁEZ LARA, Carmen (2020): «El algoritmo como protagonista de la relación laboral. Un análisis desde la perspectiva de la prohibición de discriminación», *Temas Laborales*, 155, 41-60.
SEPÚLVEDA GÓMEZ, María (2021): «El acuerdo marco europeo sobre digitalización. El necesario protagonismo de la norma pactada», *Temas laborales: Revista andaluza de trabajo y bienestar social* 158, 213-244.
— (2023): «Los principios generales de la Inteligencia Artificial en la propuesta de Reglamento Europeo y la negociación colectiva», *Revista de derecho social*, 104, 181-208.
SIERRA BENÍTEZ, Esperanza Macarena (2024): «Apuntes sobre las nuevas leyes "históricas" en el actual marco normativo de la UE sobre la Gestión algorítmica, la Inteligencia Artificial y el impacto de los derechos humanos en las empresas», *Noticias CIELO*, 3.
SORIANO ARNANZ, Alba (2021): «Decisiones automatizadas y discriminación: aproximación y propuestas generales», *Revista General de Derecho Administrativo*, 56.
TODOLÍ SIGNES, Adrián (2021): *Regulación del trabajo y política económica. De cómo los derechos laborales mejoran la Economía*, Cizur menor: Aranzadi.
— (2023): *Algoritmos productivos y extractivos*, Cizur menor: Thomson Reuters Aranzadi.
— (2024): «El principio de transparencia algorítmica en su dimensión individual y colectiva: especial referencia a la Directiva de Plataformas Digitales y al Reglamento de IA», *Trabajo y derecho: nueva revista de actualidad y relaciones laborales*, 19.
VICENTE PALACIO, María Arántzazu (2023): «La inteligencia artificial en la ley integral para la igualdad y no discriminación», *Revista General de Derecho del Trabajo y de la Seguridad Social*, 64.

Fecha de recepción: 1 de julio de 2024.
Fecha de aceptación: 20 de octubre de 2024.

EL DILEMA DEL RAZONAMIENTO, MOTIVACIÓN Y DECISIÓN JUDICIAL ALGORÍTMICA (IA): ¿LA POSIBILIDAD DE LO IMPOSIBLE?*

THE DILEMMA OF REASONING, MOTIVATION AND ALGORITHMIC JUDICIAL DECISION-MAKING (IA): THE POSSIBILITY OF THE IMPOSSIBLE?

Silvia Barona Vilar

Catedrática de Derecho Procesal
Universitat de València

RESUMEN

Los modelos automatizados de decisión y la aparición de algunas inteligencias artificiales decisoras en sede judicial generan diversos interrogantes, entre ellos si la decisión judicial algorítmica es función judicial en sentido estricto, si es una decisión razonada, si está motivada, cuál es su grado de individualización y si el marco jurídico y las condiciones tecnológicas exigidas son suficientes para equiparar la decisión judicial humana a la decisión judicial algorítmica. Se abre un debate sobre el mantenimiento o la redefinición de numerosos principios como el de inmediación, el juez natural, la independencia judicial, la imparcialidad, la responsabilidad, cuestiones, todas ellas que, lejos de ser baladíes, comportan la necesidad de una reflexión pausada para garantizarlos adecuadamente. El objeto de este trabajo no es analizar estas cuestiones, sino centrarme en el examen de función judicial decisora algorítmica *per se*, en su significado y contenido desde la reflexión en torno a qué significa pensar como juez, razonar como juez, motivar como juez la decisión judicial, y sobre si este modo de pensar es extrapolable a un modelo computacional, una estructura tecnológica compleja o a una inteligencia artificial.

PALABRAS CLAVE

Justicia automatizada, decisión judicial algorítmica, motivación judicial.

ABSTRACT

Both automated decision-making models and the emergence of certain artificial decision-making intelligences in the judicial arena raise various questions, including wether the algorithmic judicial decision constitutes a judicial function stricto sensu, wether it is a reasoned decision, wether it is motivated, which is its degree of individualization and wether the the legal framework and the technological conditions required are sufficient to equate the human judicial decision with this new category of algorithmic judicial decision. A debate is open on the maintenance or redefinition of numerous basic procedural principles, such as immediacy, the natural judge, judicial independence, impartiality, accountability, all of which, far fronm being trivial, entail the need for careful reflection to adequately guarantee them. Our paper does not aim to deal with all these issues but, standing on them as the new scenario for justice, to focus on the algorithmic judicial decision-making function itself. Addressing its understanding and content, from the reflection about what to think «as a judge», to reason «as a judge», to motivate the judicial decision «as a judge» actually means and implies, and to what extent it can be extrapolated to a computational model, a complex technological structure or an artificial intelligence.

KEYWORDS

Automated justice, algorithmic judicial decision, judicial motivation.

DOI: https://doi.org/10.36151/TD.2024.109

---

* Este artículo se ha realizado en el marco del Proyecto «Justicia sostenible en estado de mudanza global» (CIPROM 2023-64 GVA).

EL DILEMA DEL RAZONAMIENTO, MOTIVACIÓN Y DECISIÓN JUDICIAL ALGORÍTMICA (IA): ¿LA POSIBILIDAD DE LO IMPOSIBLE?

Silvia Barona Vilar
Catedrática de Derecho Procesal
Universitat de València

Sumario: 1. Tecnohumanidad, una realidad ¿evitable? 2. Trascendencia de los algoritmos y la inteligencia artificial en la función del juez. 3. Pensamiento y razonamiento judicial ¿maquínico? ¿Puede la inteligencia artificial pensar como un juez? 3.1. Heurística, razonamiento jurídico y pensamiento judicial. 3.2. Empobrecimiento o enriquecimiento del pensamiento judicial matematizado. 3.3. El significado de «pensar como jurista y como juez» y el papel de la argumentación jurídica. La inteligencia artificial no piensa, pero sí puede ejercer la función judicial decisora. 4. De la valoración de la prueba a la motivación en la decisión judicial. Diferencias entre el discurso argumentativo humano y el maquínico. 4.1. Valoración de la prueba. Criterios o estándares. 4.2. Decisión judicial = valoración probatoria + argumentación + motivación con proporcionalidad y derechos. 4.3. Decisión judicial maquínica o juez robot, ¿es función jurisdiccional de tutela efectiva? Bibliografía.

1. TECNOHUMANIDAD, UNA REALIDAD ¿EVITABLE?

Hermann Hesse escribió que para que pueda surgir lo posible es preciso intentar una y otra vez lo imposible. Los desarrollos algorítmicos y la inteligencia artificial le han dado la razón. Esos intentos se han venido sucediendo desde hace varias décadas, si bien es pertinente recordar que ya algunos pensadores de siglos pasados centraban su interés en diseñar

una máquina capaz de hacer lo que hacemos los humanos con mayor o menor profusión. El momento actual es más que propicio para considerar que, efectivamente, la posibilidad de lo imposible comienza a ser realidad.

El hecho en sí —la existencia de máquinas en sentido amplio que sustituyen a los humanos de forma sectorial— ya está presente y ha generado una enorme preocupación —o, al menos, mucha inquietud— entre aquellos que consideran que estamos ante un verdadero «[…] reposicionamiento de tipo ontológico y también antropológico» (Sadin dixit). Por un lado, porque se está produciendo una redefinición del concepto de lo humano: los seres humanos han empezado a ser suplantados por una nueva instancia suprahumana con mayor o menor grado de éxito. Por otro lado, porque ya no es el ser humano el que asume el poder de actuar *con ayuda de su espíritu, de sus sentidos y de su propio saber, sino una fuerza interpretativa y decisional que se tiene por más eficaz* —y que, de forma legítima, tiende a eliminar al ser humano cada vez más y cada vez en más sectores de la vida (Sadin, 2019: 146).

Esta opinión, sin embargo, contrasta con la del transhumanismo, una corriente emergente —y entusiasta— que aspira a la superación del ser humano por el *cyborg*, pues considera que la criatura humana ha tocado techo. Los avances tecnológicos, la aparición de la nanotecnología y los progresos de la neurociencia han impulsado desarrollos que caminan en esa dirección, la búsqueda de la «singularidad», del «super-yo», la expansión acelerada de diversos fenómenos a los que los filósofos contemporáneos han puesto nombre (el enjambre digital, el proletariado digital, la siliconización del mundo, entre otros). Se trata, en suma, de una corriente de pensamiento sustentada en la defensa de la dependencia y la sumisión de la humanidad a la tecnología, la asunción acrítica de los valores que predominan en la sociedad global en que vivimos, el primado de la eficacia, la eficiencia, la instantaneidad y la economización de los tiempos, y, en fin, la idea de que sin tecnología no hay paraíso. El transhumanismo aboga por la minimización del Estado y su definitivo reemplazo por la esfera la economía, y sostiene que hay que dejar que el mercado fluya y se convierta en el único mecanismo de asignación de bienes. En ese fluir, el sector económico —y hoy tecnológico— ha ocupado espacios, ha marcado la agenda política y ha alentado la construcción de unas sociedades y una vidas fundamentalmente económicas. No se trata de un simple cambio, sino de una mutación sistémica de la realidad global y de la humanidad propiciada por los desarrollos que en esta nueva etapa de industrialización han impulsado la tecnología y la ciencia, transformación concretada en la digitalización, la conectividad, la automatización, la robotización y la inteligencia artificial (Barona, 2021: 58-67).

Sin necesidad de elegir una de las dos opciones de la disyuntiva fascinación *versus* derrotismo o pesimismo, no hay duda de que asistimos a un cambio de ciclo en la humanidad que tiene su eje, su núcleo, en la tecnología. La tecnológica ha ido ganando terreno desde el inicio de la Tercera Revolución Internet, nucleada en torno a la comunicación y, muy especialmente, tras la aparición de internet, un nuevo entorno que fue ganando adeptos a medida que se verificaban sus fructíferos resultados. Ese impulso, que comenzó en las últimas décadas del siglo XX, ha generado un sinnúmero de innovaciones, entre ellas las herramientas algorítmicas que poco a poco han complementado y asistido al ser humano en la ejecución de algunas tareas (antes exclusivamente humanas) y que de manera pau-

latina pero imparable han penetrado en nuestras vidas, en nuestro *modus operandi*, en nuestra manera de relacionarnos, de adquirir productos, de leer, de pasar el tiempo libre, de trabajar. Si en el plano personal este cambio de *modus vivendi* asociado a la irrupción de las nuevas tecnologías es indiscutible, en el sector productivo estas se han convertido en un elemento imprescindible que, si bien estimula la interacción entre la persona y la máquina, provoca la paulatina pérdida de relevancia de la persona y el creciente protagonismo de las máquinas, tendencia que implica la sustitución del ser humano por la tecnología en numerosos sectores. En efecto, la tecnología informática ha propiciado la automatización en ámbitos tan distintos como la agricultura, la banca, la salud, la distribución, la automoción, y también en la Justicia. Tras la consolidación de la industria 4.0, los desarrollos tecnológicos han rebasado las expectativas iniciales (Schwab, 2018: 15), generando cambios de todo orden. Por todo ello, se ha afirmado que la revolución tecnológica no solo lleva aparejada una disrupción social y personal, sino también una metamorfosis de la humanidad.

En ese contexto, desde las últimas décadas del siglo pasado se han desarrollado iniciativas tecnocientíficas que han despertado interés en el sector público y, especialmente, en el mundo judicial, y han surgido posiciones favorables y contrarias a lo que la doctrina ha denominado la «Justicia automatizada» (Dymitruk, 2019: 27, y Guzmán Fluja, 2021: 3-4), sustentada en la aplicación de la tecnología a los procesos judiciales y arbitrales, así como su aplicación extensiva a los sistemas de solución extrajudicial de conflictos, en los que los sistemas y estructuras inteligentes arrojan unos resultados más prometedores que los que se presentan anualmente en las memorias de los tribunales humanos cuando son analizados desde parámetros de eficacia y eficiencia.

Se trata, pues, de un paisaje tecnológico que hace la vida fácil, agiliza la profesión y propicia la obtención de resultados altamente eficientes. Bajo esas coordenadas, el dilema que suscita la revolución tecnológica es si queremos evitar la fagotización de lo humano por lo tecnológico y recuperar el control de los artefactos que nosotros mismos hemos creado o si, por el contrario, preferimos abandonarnos a esa realidad expansiva que nos invade y deglute, y perder nuestra imperfecta identidad humana privilegiando un proceso tecnológico que parece imparable ¿o quizás inevitable? La evitabilidad está en nuestras manos, pues podemos aprovechar las ventajas de la tecnología y limitar sus excesos y los posibles «sometimientos» que nos imponga.

2. TRASCENDENCIA DE LOS ALGORITMOS Y LA INTELIGENCIA ARTIFICIAL EN LA FUNCIÓN DEL JUEZ

Uno de los dilemas que se plantea actualmente en la Justicia es si puede algoritmizarse la tarea del juez, es decir, si es viable lo que los anglosajones denominan la *judge craft*. La aplicación de la tecnología al desempeño la función jurisdiccional ha conocido dos momentos. El primero estuvo marcado por la incorporación de la tecnología de primera generación, que de alguna manera asumió un papel más instrumental que decisorio. En el

segundo momento, la incorporación de modelos computaciones y algorítmicos permiten, facilitan y «suavizan» la tarea judicial.

En un primer nivel de avance tecnológico en el ejercicio de la función tuitiva judicial, aplaudido y aceptado, surgieron sistemas algorítmicos que colaboraban y asistían, y asisten, a los jueces, convirtiéndose en su sustento y apoyo a través de lo que la doctrina anglosajona ha denominado la *decision support* (Leith, 1998: 290-293). Se trata de herramientas algorítmicas que facilitan la tarea judicial: extracción y análisis de documentos, predicción de riesgos, evaluación de testigos —como ADVOCATE— y análisis de su fiabilidad (Andino López, 2017: 473-485), reconstrucción de hechos (como STEVIE, Nixan, 2009: 11), búsqueda de indicios a partir de otros hechos similares, etc.

Ninguna de las herramientas que realizan estas funciones sustituye al juez. Por el contrario, la herramienta se pone al servicio del juez para facilitar el desarrollo de su tarea, a la manera de un colaborador maquínico que puede acelerar la toma de decisiones o contrastar las percepciones judiciales, máxime tras la aparición de algunos programas de *software* que permiten incluso formular hipótesis sobre el modo en que sucedieron los hechos y elaboran propuestas exculpatorias o inculpatorias. De hecho, se han diseñando programas computacionales que permiten analizar documentos, compararlos con otros que se hallan en bases estadísticas, y determinar el grado de veracidad, de autenticidad, de complejidad, de contradicción, etc. Se trata de sistemas asistenciales, complementarios o integradores de las funciones judiciales que desarrollan funciones predictivas de riesgo, perfiladoras, e incluso propositivas y que condicionan que o determinan total o parcialmente las decisiones judiciales. Huelga señalar que estas herramientas son perfeccionadas día a día.

Obviamente, también surgido algunos programas y aplicaciones que no se han librado de crítica. Por ejemplo, la canadiense ASSYST (Simon y Gaes, 1989: 198), que fue concebida para ayudar a los jueces a aplicar las directrices en la sentencia, pero que, finalmente, provocaba una suerte de aplicación automática de la condena penal, efecto indeseado que está lejos del significado del *judge craft* en el mundo de la Justicia. Otra herramienta canadiense es LIST, desarrollada en la University of British Columbia, que proporciona información relevante al juez y que, a diferencia de la herramienta ASSYST, no pretende determinar el contenido de la resolución, sino únicamente facilitarla. Si bien en este segundo supuesto no estamos ante una herramienta propositiva decisora, tanto ASSYST como LIST han sido objeto de críticas que apuntan a su incapacidad de asimilar en cada caso la posible complejidad real del supuesto y del razonamiento a seguir en el dictado de la sentencia (Schild, 1998: 159). Un paso más se dio con la herramienta canadiense SIS (Sentencing Information System) que, más allá de ser colaborativa y propositiva, se diseñó con la finalidad de promover la uniformidad de las sentencias. SIS fue muy criticada por la doctrina (Greenleaf, 1991: 45-48), que consideraba que mediante su utilización era posible controlar y manipular la función judicial en la esfera más propia, íntima e independiente que desarrollan los jueces humanos, a saber, la decisión judicial.

La imparable evolución de los desarrollos e innovaciones tecnológicas evidencia que, lejos de detenerse, la hibridación del juez y la tecnología en el ejercicio de la función jurisdiccional se está consolidando: es cierto que el juez decide, pero la funcionalidad algo-

rítmica es indudable y juega un papel esencial en la conformación de su función. Desde este planteamiento, estamos asistiendo a una evolución aceptada casi sin discusión en la que, amén de las herramientas algorítmicas asistenciales o complementarias, se emplean herramientas integradoras de la función del juez humano que allanan paulatinamente el camino hacia la incorporación de sistemas computacionales «¿capaces?» de realizar el oficio del juez, de sustituir al juez humano en las decisiones propias de la función jurisdiccional, sistemas que propician la aparición del denominado «juez-robot» (Barona, 2022*a*: 103).

Las dudas que esta evolución genera respecto a la función judicial decisora, esencia nuclear del derecho a la tutela judicial efectiva reconocido en el artículo 24 CE, son cada vez más persistentes. Entre otras cuestiones, se plantea el grado de transformación decisora que supone el empleo de algoritmos e inteligencia artificial en la decisión judicial, la necesidad de capacitar a los jueces para manejar estas herramientas, pero también para validar la decisión final desde la crítica constructiva que le permitir separarse de la solución propositiva algorítmica o, por el contrario, asumirla. Asimismo, se ha abierto un debate sobre la necesidad de mantener o la pertinencia de redefinir numerosos principios, entre ellos el de inmediación, el del juez natural, la independencia judicial, la imparcialidad, la responsabilidad, cuestiones, todas ellas, que, lejos de ser sencillas, reclaman una reflexión pausada que a los efectos de garantizarlos. La finalidad de este estudio no es analizar estas cuestiones, que hemos abordado en diversos trabajos. Aquí nos centraremos únicamente en la función judicial decisora algorítmica *per se*, en su significado y su contenido.

Más allá del empleo de herramientas asistenciales o complementarias en la toma de decisiones, la senda de la automatización procesal —incluyendo la decisión judicial (robotización judicial) y la emergente interacción inteligencia artificial-juez— conforma un escenario que es ineludible analizar. Mucho se ha escrito y discutido en torno a la función que le corresponde al juez en la toma de decisiones, culminación del derecho a la tutela judicial efectiva. Los debates desarrollados en foros nacionales e internacionales han planteado el ser o no ser de la individualidad que se exige en el ejercicio de la *judge craft* a través de la disección de los componentes esenciales que integran la función de tomar decisiones. La mirada del juez en la toma de decisión va a estar presente, individualizada, subjetivada y determinada en atención a los hechos, los sujetos y las circunstancias concurrentes. Todo ello es juzgar, y esa tarea, lejos de ser simple o automática, es muy compleja. No solo requiere conocer el Derecho, sino también «[…] haber adquirido y desarrollado las competencias y habilidades propias de este arte práctico, que por sí solas también son insuficientes si no están insufladas por la virtud de la prudencia […] es por ello que el oficio de juez requiere estas habilidades y virtudes prácticas, profesionales y personales necesarias para ser un buen juez. Su contenido no se encuentra en los libros de Derecho sustantivo o procesal. Es algo que se adquiere por la experiencia, propia o ajena, y también con la reflexión sobre esa experiencia» (Sancho Gargallo, 2020: 446).

Por todo ello, merece la pena plantear en qué consiste pensar como juez, razonar como juez, motivar la decisión judicial como juez, y reflexionar sobre la cuestión de si estos cometidos pueden ser encomendados a un modelo computacional, a una estructura tecnológica compleja o a una inteligencia artificial.

3. PENSAMIENTO Y RAZONAMIENTO JUDICIAL ¿MAQUÍNICO? ¿PUEDE LA INTELIGENCIA ARTIFICIAL PENSAR COMO UN JUEZ?

Prima facie, las máquinas no piensan: no existe el pensamiento maquínico. Como mucho, puede afirmarse que la inteligencia artificial imita el pensamiento humano mediante la suma ágil, rápida y efectiva de cálculos de probabilidad basada en el tratamiento de explotación de datos (*big data*), que son los que se integran en la simbiosis «legislación, jurisprudencia, doctrina» con un resultado estadístico, digital, que no es pensamiento. Las inteligencias artificiales no son inteligentes, sino estadísticas (Cardon, 2018: 78). Las máquinas no piensan, no comprenden, sino que trabajan estadísticamente con una masa de datos que les suministran la información necesaria realizar su función; carecen de memoria perceptiva, de sensación del paso del tiempo, de recuerdos, de sensaciones ante estos, de creatividad, etc. Naturalmente, en la evolución de estos sistemas computacionales complejos, perfeccionados y cada vez más adaptados al medio se percibe una pretensión de «pensar como juristas» y, especialmente, de «pensar como juez» a través de la aplicación de la tecnología adecuada que permita reproducir el pensamiento judicial, el razonamiento jurídico humano, a la inteligencia artificial. Las máquinas pueden imitar y replicar el pensamiento judicial —integrado por el razonamiento argumentativo, la interpretación con valoración de la prueba y la motivación de la decisión—. No se trata de un pensar maquínico, sino de la traslación del pensar humano a una suma ágil, rápida y efectiva de cálculos de probabilidad basada en el tratamiento de explotación de datos relativos a decisiones del pasado.

Es especialmente significativa la idea que sostuvo el filósofo estadounidense John Rogers Searle (Searle, 1980: 417-424), uno de los pioneros de la denominada filosofía de la inteligencia artificial, a saber, que los ordenadores fabricados por humanos carecen de la capacidad de pensar y que la probabilidad de que alguna vez adquieran conciencia es igual a un cero por ciento. Searle se opuso a quienes sostenían que con la inteligencia artificial fuerte se produce un acercamiento a la consciencia, al tratarse de un modelo para realizar lo que hace el cerebro humano, y afirmó que un sistema computacional inteligente manipula diversos códigos sintácticos que no tienen que ver con la comprensión semántica de los contenidos procesados.

Pensar no es leer letras, no es alimentarse de información y traducirla, integrarla, extraer lo esencial respecto de un caso, que puede ser parte integrante de ese desarrollo intelectual que conduce a la toma de decisiones. La máquina puede desarrollar funciones que hasta el momento realizábamos como humanos e incluso puede hacerlo a una velocidad inusitada, eficientísimamente, que supera a la velocidad humana, pero la máquina no ha podido, al menos no ha podido hasta el momento, *provocar un discurso interior en el que se plasma la continuidad de la consciencia como memoria* (Lledó, 2011: 151). Ese discurso interior se refiere al proceso cognitivo consciente, perceptivo, y sus resultados son los que provocan el juicio, el razonamiento, la conceptualización, la interpretación valorativa y la argumentación que provocan la decisión.

En ese proceso decisorio también juegan un rol importante las emociones, las percepciones o intuiciones que no se perciben conscientemente, lo que puede conducir a la emisión de juicios que difieren de las conclusiones precedidas por un análisis racional

(Kahneman, 2011), y que pueden provocar angustia o duda. En ciertos casos, el exceso de subjetividad puede ser pernicioso, Por ello, algunos autores sostienen que las decisiones tomadas a través de una inteligencia artificial son más objetivas que las humanas. Cabe responder a esta afirmación señalando que también en el diseño, la programación y la aplicación de los algoritmos y la inteligencia artificial concurren componentes subjetivos, dado que quienes programan los sistemas algorítmicos son seres. Hoy por hoy, este simple hecho impide considerar que la frialdad de la máquina es *per se* más favorable a la objetividad que el juez humano, pues detrás de ella está la subjetividad de quienes la hacen realidad.

Por tanto, debemos alejarnos de esa la neutralidad algorítmica que proclaman algunos y recordar, en todo caso, que existen numerosos asuntos, prácticamente la mayoría, en los que no se puede aplicar matemáticamente el Derecho al caso concreto sino tras una identificación de componentes variables y, en muchos casos, con dosis de elementos subjetivos basados en las emociones y las percepciones, amén de los resultados de las conductas de las partes y sus representantes en el proceso. Para poner en evidencia que esa neutralidad es falsa se han presentado numerosos exponentes de sesgos algorítmicos que, ora son inoculados por sus diseñadores, ora han asimilado los parámetros discriminadores de la sociedad y entorno en el que actúan. Por ello, el falso mito de que las matemáticas y la frialdad de los datos permiten la neutralidad de los algoritmos no debe servir de argumento para desequilibrar la balanza de la función judicial más objetiva, inclinándola a favor de la decisión maquínica, en detrimento de la decisión humana.

3.1. HEURÍSTICA, RAZONAMIENTO JURÍDICO Y PENSAMIENTO JUDICIAL

Ante la existencia de conflictos y la necesidad de tomar decisiones, el pensamiento humano realiza una búsqueda de reglas de juicio que le guíen en la toma de esas decisiones. Son estrategias, reglas, silogismos y conclusiones que se han ido integrando en la denominada «heurística» (que permite seleccionar las posibles bifurcaciones ante un camino lleno de ramificaciones, restringiendo las posibles soluciones). Las reglas heurísticas no son infalibles, ni necesariamente abocan a la solución mejor o más adecuada.

Hablar de heurística implicaba hasta hace bien poco la asociación con el pensamiento humano. No obstante, en la actualidad se observa que la irrupción de la inteligencia artificial ha permitido igualmente incorporar en algunas manifestaciones de las reglas heurísticas representaciones del conocimiento, entendidas como combinaciones de estructuras de datos y de procedimientos de interpretación integradas en un programa computacional (con diferentes lenguajes para diferentes campos de aplicación, ejemplo del cual es el *software*). Así, la aplicación de las reglas heurísticas de la mente humana son las que se trasladan al modelo computacional, si bien con un modelo actuacional diverso, dado que en los sistemas computacionales la aplicación de fórmulas algorítmicas lleva a trabajar con un modelo de automatización que permite, desde la alimentación del *big data,* ofrecer respuestas argumentadas.

La aplicación de las nuevas tecnologías al mundo judicial genera el dilema no tanto de si la inteligencia artificial es aplicable a la decisión judicial —lo es, como se percibe con

los ya existentes modelos de automatización procesal y judicial y los que presumiblemente se programarán—, sino de si esa aplicación genera realmente pensamiento judicial o, lo que es más palmario, si el razonamiento humano que se despliega en la valoración de la prueba, en la interpretación, en la argumentación y en decisión motivada puede entenderse extrapolable a cualquier asunto, a cualquier tipo de proceso y a cualquier tipo de sentencia. Se aceptan procesos automatizados, algunos, en ciertos casos, reiterativos, repetitivos y estadísticos. Es el fenómeno de la matematización de la decisión judicial, que es posible, pero no siempre ni para todo, por lo que no es ni implica la sustitución de la función de ser juez por las máquinas algorítmicas e inteligencias artificiales. Lo que ofrece son respuestas, decisiones que se fundamentan en el pasado, pero no razonamiento jurídico, o no lo es en el sentido en que este ha sido concebido hasta hace poco.

Si atendemos a los filósofos del Derecho, el razonamiento jurídico permite aunar el Derecho, la moral y la política (Atienza, 2006: 12-13), de modo que ese razonar de forma jurídica significa *relacionar problemas jurídicos específicos con una vasta red de principios de derivación jurídica o de moral política*, principios que permiten la interpretación, acogiendo los valores básicos, pero pudiendo atribuirles un mayor o menor peso a través del razonamiento justificativo de los jueces (Dworkin, 2006: 56). De este modo, si bien se pretende ofrecer una aclaración de la interpretación y aplicación del Derecho desde la concurrencia de una guía o protocolo de justificación (Alexy, 1978: 19), el razonamiento jurídico permite incorporar la actividad argumentativa a los procesos de toma de decisión (Atienza, 2006: 14), o, lo que es lo mismo, comporta ese «pensar como jurista» (expresión americana de *legal reasoning).*

Lo interesante es determinar si estamos ante una facultad atribuida tan solo a cerebros humanos o si es posible que la secuencia pensamiento-razonamiento-interpretación-argumentación jurídicos, bajo la que necesariamente se va a encuadrar la teoría de la argumentación jurídica, pueda ser desempeñada mediante sistemas computacionales inteligentes que reproduzcan de forma idéntica ese pensar.

Podríamos referirnos a varios niveles de innovación tecnológica. Por un lado, es palmario que existen sistemas de expertos, asistentes o *acompañantes digitales* que nos guían, colaboran con nosotros y nos complementan tanto personal como profesionalmente hasta el punto de que llega a forjarse una suerte de *alteridad*, sin rostro ni cuerpo, pero absolutamente imprescindible, también en la profesión jurídica y, por supuesto, en sede judicial, porque ofrecen aclaraciones, simplifican y favorecen trámites, resuelven cuestiones, dudas, etc. Estas herramientas dieron lugar a la aparición en su día de la ALI (*artificial legal intelligence*), vinculadas a los primeros desarrollos de IBM Watson, IBM's Debater Program, sistema Ross, etc., que ofrecían respuestas a las cuestiones jurídicas planteadas, pero suministraban respuestas sin razonamiento jurídico alguno; respondían a modelos de lógica computacional clásica, pero no de argumentación jurídica (Ashley, 2017: 3). La ausencia de argumentación en las soluciones ofrecidas por estos sistemas computacionales constituía un obstáculo para su expansión. Permitían obtener una respuesta basada en cálculo probabilístico y algoritmos, pero sin que la misma estuviera sustentada en un modelo deductivo racionalista.

En su evolución fueron incorporando algo similar a la argumentación. Aunque no sean propiamente argumentativas, puede afirmarse que actúan de antesala del proceso argumentativo e interpretador, son herramientas instrumentales del pensar como jurista, pero no son pensamiento *per se,* sino medios que contribuyen a que este se despliegue. En cualquier caso, esto no era ni es «pensar como jurista», pero lo parece. Y esa apariencia ha ido ganando espacio a medida que lentamente se ha aceptado en ámbitos específicos y acotados la posible sustitución del paradigma deductivo racionalista por los mecanismos de cálculo probabilístico y las inferencias algorítmicas. Por otra parte, los desarrollos innovativos computacionales permiten que se desarrolle la ya citada «justicia automatizada», que incluye soluciones estadísticas, maquínicas con motivación matematizada. Se trata de una matematización del razonamiento que se basa en el pasado, en las soluciones ya dadas, y, por ende, en una respuesta mecánica en la que no concurre el sentido del razonamiento jurídico individualizado. No hay interpretación, no hay argumentación, no hay pensamiento judicial, sino aplicación de un modelo de cálculo que ofrece respuestas a situaciones objetivo-subjetivas similares. Una automatización que está presente en las plataformas digitales, que incorporan soluciones algorítmicas propias, o en los modelos de solución de conflictos de los procedimientos monitorios (ejemplo de ello es el sistema alemán de Mahngerichte: https://www.mahngerichte.de, que ha automatizado el procedimiento y su solución). En estos supuestos, las coordenadas son similares y las decisiones que se toman son también muy similares, como ocurre, por ejemplo, en las decisiones que resuelven las discrepancias en materia de consumo.

3.2. EMPOBRECIMIENTO O ENRIQUECIMIENTO DEL PENSAMIENTO JUDICIAL MATEMATIZADO

La imparable evolución tecnológica ha traído nuevos modelos computacionales capaces de seleccionar más y mejores argumentos jurídicos. Se produce una profusión de herramientas argumentativas que son, claramente, un gran sustento para la abogacía; en muchos casos, son herramientas argumentativas persuasivas que integran razonamientos jurídicos de casos fenecidos (pasados) y que incorporan los textos, los informes dogmáticos y la jurisprudencia que los sostiene. Estas herramientas ayudan al abogado en su disertación jurídica orientada a convencer al juez para alcanzar un resultado que favorezca los intereses de su cliente. Por su parte, para el juez lo importante no es persuadir, sino justificar la decisión, de ahí que su argumentación se ampare en componentes diversos, dado la finalidad también es diversa.

Han surgido simultánea y sucesivamente numerosas herramientas computacionales con capacidad imitativa cognitiva del ser humano que han abierto una posible interconexión colaborativa entre los humanos y las máquinas, facilitando y mejorando, en ciertos casos, las respuestas. Este dualismo presenta *a priori* una perspectiva previsible de mejora de la actuación de los profesionales del mundo jurídico y, especialmente, del mundo judicial. La aceptación de esta interacción máquina-juez no significa que confluyan dos pensamientos, sino que el pensar es uno, humano, pero la asistencia para el pensamiento, basada en experiencias vividas, resueltas, fundadas, es doble, de manera que la tecnología algorítmica y la

usabilidad de la posible inteligencia artificial pueden ser de gran valía para la construcción del pensamiento judicial. Esta manera de trabajar puede repercutir positivamente en la Justicia, eliminando tiempos, en ocasiones excesivos, y trabajando con el «asistente algorítmico», que puede (y, con su perfeccionamiento, mejorar) mejorar la tutela judicial otorgada.

Frente a esta idea, la aparición de modelos de justicia civil automatizada a través de la actuación de sistemas computacionales inteligentes ha sido contestada en la doctrina (Guzmán, 2017: 68; Nieva, 2018: 58), que ha argumentado que ni es más satisfactoria la tutela que se otorga ni comporta una mejora de los resultados alcanzados, comparativamente evaluados con las decisiones tomadas por el juez humano, básicamente porque son decisiones automáticas que ignoran posibles soluciones alternativas. Ello no es óbice, a mi parecer, para que estos sistemas proporcionan mayor claridad legal, amén de reducir tiempos, favorecer la eficiencia y limitar la potencial arbitrariedad que puede concurrir, entre otros, siempre, naturalmente, que se pondere la eficiencia con la garantía de los derechos, a sabiendas de que, en ciertos casos, la ductilidad de la decisión humana, su mayor capacidad de adaptarse hoy por hoy al caso concreto, es indudable. Y, sobre todo, ofrecen grandes dosis de predictibilidad.

Así, la proliferación de herramientas de expertos asistenciales en el mundo jurídico,que ofrecen «conocimiento y experiencia» (de los humanos) es indiscutible; en ciertos casos están vinculadas a la argumentación jurídica. Estos modelos de conocimiento fundados en la automatización se despliegan en tres ámbitos jurídicos: el jurisprudencial o análisis de los precedentes, favoreciendo una determinada estrategia de defensa; el legislativo, ofreciendo una función analítica del sistema jurídico, interpretando las normas y asistiendo en la redacción y la corrección; y el análisis de documentos y la posible elaboración de contratos, informes, decisiones, etc. a partir del análisis automatizado de datos. Estos modelos permiten trabajar a dos niveles: el asistencial (con análisis e interpretación) y el creativo (analítica y creativa).

3.3. EL SIGNIFICADO DE «PENSAR COMO JURISTA Y COMO JUEZ» Y EL PAPEL DE LA ARGUMENTACIÓN JURÍDICA. LA INTELIGENCIA ARTIFICIAL NO PIENSA, PERO SÍ PUEDE EJERCER FUNCIÓN JUDICIAL DECISORA

El hecho de que un programa, una máquina, un sistema computacional o un robot puedan hacer labores que antes hacían los humanos, y que son consecuencia de la reflexión, el pensamiento, la consciencia, y la decisión, no significa que tengan pensamiento, consciencia y memoria. Puede parecer que son juristas y que piensan como ellos, pero no piensan como juristas porque no lo son, al menos hoy por hoy. Ahora bien, aun cuando no piensen, sus prestaciones se integran en la función judicial decisora. Para entender esta afirmación, debemos tomar en consideración considerar dos reflexiones.

1) Pensar como jurista y como juez exige la capacidad de argumentar. Ahora bien, decidir y argumentar no es lo mismo, aun cuando tienen una conexión directa, de manera que la argumentación es la que suministra razones de la decisión que se adopta, es la consecuencia de la argumentación (Atienza, 2006: 65). En ella confluyen información y

argumentos, pero también percepción y conciencia. Así, la actividad decisora humana es el resultado de un proceso argumentativo largo, en el que paulatinamente el órgano decisor va formando su opinión.

2) En la actividad decisora algorítmica no se da el proceso argumentativo, sino que se tiene, lo que supone que su actividad no existe, sino que se arroja una respuesta matemática que, desde el pasado —alimentado por los *inputs* algorítmicos— ofrece un resultado (¿decisión?).

Esas dos premisas son dos afirmaciones que encuentran cada vez más matizaciones debido a la evolución de la realidad computacional y a los desarrollos de la inteligencia artificial. A título de ejemplo, existen plurales modelos de argumentación jurídica —descriptiva e informativa, analítica e interpretativa, e incluso creativa— y pueden desplegarse a través de modos diversos, destacando la retórica y la dialéctica (Atienza 2006: 260-261); incluso se habla del supuesto especial de argumentación práctica racional, en el que el discurso práctico racional define un superjuego que contiene todas las reglas de los otros juegos argumentativos (Alexy, 1978: 19-20).

Esto lleva a considerar que el diálogo, el debate, los juegos, las reglas de la razón o la práctica racional son conceptos propios de la mente humana, aunque poco a poco empiezan a ser factibles en el sistema computacional automatizado-inteligente, ofreciendo una suerte de realidad dialéctica desde la máquina. De hecho, ha emergido una disciplina denominada «dialéctica computacional», que intenta describir cómo el conocimiento puede surgir a través de técnicas de argumentación simbólica; comenzó a través de modelos dialécticos inspirados en casos (defendida por Ashley), y posteriormente evolucionó hacia la integración de la exactitud de la lógica simbólica con la flexibilidad que puede proporcionar la argumentación informal, de manera que generó una nueva lógica en la que se trabaja con protocolos (mucho son los autores que se han preocupado de ello: Krabbe, Loui, Prakken y Sartor, etc.). Algunos de estos modelos protocolarios permiten a las partes construir sus propios argumentos, si bien paso a paso, de forma sucesiva, mientras que otros permiten la argumentación de forma instantánea. Incluso algunos permiten crear contraargumentos para refutar los del contrario, bien *ad libitum,* o bien de forma restringida (refutar como máximo dos veces los contraargumentos de la parte contraria). No existe, en consecuencia, una única manera de actuar argumentativamente, dada la heterogeneidad y asimetría de cada asunto, de manera que habrá que individualizar en cada caso concreto cuál de los sistemas dialécticos es el adecuado. El mundo del Derecho es poliédrico, y absorbe diversidad de discursos, técnicas, diálogos y retórica en la argumentación que son adaptables al contexto, y esto es lo que genera dudas acerca de la adaptabilidad de la argumentación jurídica del actual modelo de inteligencia artificial.

En ese modelo argumentativo, el lenguaje asume un papel protagónico. Se requiere la hermenéutica y la semántica respectivamente. No se puede separar la teoría del significado lingüístico del hecho de que seamos capaces de entender palabras y frases (Gabriel, 2019: 338). Más allá de un código o vehículo de pensamiento —que puede sin lugar a dudas reproducirse en modelos computacionales algorítmicos sin ninguna dificultad—, debe atribuirse al lenguaje un valor allende este, a saber, el lenguaje es una actividad, es una acción

que se realiza a través de las herramientas, que son las palabras. Uno de los componentes que identifican a los seres humanos es su psique, caracterizada por un lenguaje, propio, adecuado, específico, transmisor de pensamientos e interpretador de los contrarios; un lenguaje en el que no solo existen palabras sino otros átomos que lo nuclean, pero que reflejan mucho más que las palabras. El «lenguaje nos habla» (Lledó, 2011: 41). Lo trascendente no es el significado de una palabra, sino su filosofía, a la que se accede analizando el lenguaje no como una proposición filosófica, sino como una actividad (Wittgenstein, 1984: 85).

Desde esta visión, hoy es difícil aceptar que, más allá de una mente humana, pueda desempeñarse esta función, toda vez que nunca será repetida ni repetitiva, sino única en cada caso. Además, hay también una cuestión interesante planteada por los filósofos del lenguaje: ¿puedo tener pensamiento sin tener lenguaje? ¿Se puede razonar sin lenguaje? ¿Puede haber pensamiento jurídico sin lenguaje? El lenguaje no es, por ende, un vehículo de transmisión, sino una *caja de herramientas* con muchas palabras en su interior (a la que se refería Wittgenstein) que podrían extrapolarse a un modelo algorítmico inteligente. Sin embargo, lo que diferencia al ser humano de la máquina, al menos hasta el momento, es que las palabras no siempre tienen la misma función y no siempre el ser humano las acciona de igual modo, variabilidad que es fruto del pensamiento mismo. La idea de la caja de herramientas (palabras, nociones jurídicas) nos permite considerar el lenguaje como una pieza más de ese pensar como jurista.

Es indudable que la combinación de humanos y máquinas puede comportar mejores resultados, en la medida en que los humanos pueden incorporar la experiencia e historia, el raciocinio, la intuición, la empatía y los afectos, emociones y sentimientos, criterios éticos y creatividad; aportan «alma», aunándose a la mayor racionalidad y la mirada analítica de las máquinas, que además pueden aportar memoria enciclopédica y habilidades computacionales. En todo caso, también las diferencias tecnológicas son evidentes, dado que una cosa son los sistemas de expertos y otra los paradigmas computacionales cognitivos: el grado de conocimiento será diverso. En el primer caso, el sistema se construye por ingenieros que incorporan el proceso de adquisición de conocimientos, mientras que en el segundo lo que se pretende es que el conocimiento no provenga del conocimiento humano incorporado, sino de la información legal que procesa el propio sistema, de modo que sea este —el sistema— el que identifique los elementos relevantes que deben tenerse en cuenta en la cuestión o problema planteado y puedan ofrecer desde el sistema la propuesta de solución más adecuada. Son, en suma, los diversos grados de manifestación de estas estructuras computacionales concurrentes.

Ahora bien, el proceso de toma de decisiones judiciales es verdaderamente complejo. Tanto, que resulta cuando menos cuestionable que pueda reducirse a modelos lógicos algorítmicos. No puede negarse que hay cierta conflictividad respecto a la posibilidad de considerar «computarizable» esa toma de decisiones (hechos, infracciones y soluciones repetitivas, así como cuando preexisten protocolos que condicionan las decisiones), siempre en atención a las características que comporta la formulación del razonamiento jurídico del juez, y, sobre todo, asumiendo que se presentan en la realidad algunos procedimientos automatizados o estandarizados: en estos, los sistemas algorítmicos pueden desempeñar un

papel importante, eficiente y económico, amén de ágil. Se trata de aquellos supuestos en los que pueden concurrir factores de simplicidad, repetitividad y frecuencia de los procedimientos en casos idénticos o bastante similares (Taruffo, 1998: 318). Son los procesos de automatización decisional en materia laboral, de consumo, en las infracciones de tráfico, etc., pero no así en materia penal o en cuestiones que requieren un análisis individualizado que justifique la decisión singular —que, por otra parte, son la mayoría de los asuntos—.

4. DE LA VALORACIÓN DE LA PRUEBA A LA MOTIVACIÓN EN LA DECISIÓN JUDICIAL. DIFERENCIAS ENTRE DISCURSO ARGUMENTATIVO HUMANO Y MAQUÍNICO

Las reflexiones acerca de la argumentación jurídica entroncan con el papel legítimo que asume el juez en el ejercicio de su función de juzgar y hacer ejecutar lo juzgado. Esa función, y el papel de los jueces para dar respuesta a los justiciables en el ejercicio de su derecho a la tutela judicial efectiva, exige racionalidad de su decisión (Barranco Avilés, 2010: 135). La racionalidad no es arbitrariedad, sino una función valorativa judicial que puede controlarse a través de la exigencia de motivación de las decisiones.

La motivación es una garantía para los justiciables, si bien lo es también para la sociedad en general y para calibrar la salud de los sistemas jurídicos democráticos, de ahí que se afirme que no se circunscribe a la funcionalidad procesal *inter partes*, sino que cumple una función extraprocesal (Taruffo, 1986: 443). Por ello, la motivación se presenta como el «[…] signo más importante y típico de la racionalización de la función judicial» (Calamandrei, 1960: 115). La motivación es de vital importancia en un modelo garantista, como lo muestra la incorporación del deber de motivar en el artículo 120.3 CE; se trata, en realidad, de la «garantía de la garantía» (Ferrajoli, 1980: 681).

En consecuencia, la adopción. de la decisión judicial comporta la existencia del razonamiento argumentativo que realizan los jueces en relación con los posibles asuntos fácticos y legales desde un modelo dialógico basado en el entendimiento, la comparación, la valoración y la elección final de la solución en cada caso concreto. En este proceso ello juega un papel decisivo la aportación dialéctica de las partes (dependiendo del modelo procesal civil y penal, sea más o menos dispositivo) y ese diálogo del juez consigo mismo (Taruffo, 1998: 315) que se exteriorizará a través de la debida motivación en la sentencia. En la función judicial, en consecuencia, se produce una actividad decisoria, por un lado, y una actividad fundamentadora de la decisión, por otro, que deberá ser consistente —los argumentos utilizados no entran en conflicto entre sí—, adecuada —cada enunciado debe estar bien fundado en buenos argumentos justificativos— y coherente —los argumentos utilizados encajan bien con la naturaleza de los asuntos decididos—.

Ahora bien, para diferenciar el discurso argumentativo humano del de la máquina debemos tomar en consideración que los humanos centran su discurso argumentativo sobre una de las posibles alternativas que se les presentan ante una situación de incertidumbre (cambio de legislación, estrategia procesal del abogado, línea de acusación, decisión ju-

dicial, etc.). No trabajan con probabilidades, como hacen los sistemas algorítmicos, sino desde una actividad cognitiva que se adecúa a las habilidades del razonamiento humano, de modo que los humanos despliegan creencias e hipótesis que emplean para sustentar su posición. En general, los humanos no son especialmente hábiles para asignar probabilidades numéricas (Sartor, 2009: 18); esta es la objeción que esgrimen recurrentemente los entusiastas del sistema computacional, dado que consideran que es un fracaso de la cognición humana la incapacidad para trabajar con probabilidades numéricas, como sí lo hacen los sistemas informáticos, que utilizan redes complejas de probabilidades, que funcionan en ciertos dominios, al manipular las probabilidades numéricas mucho más rápidamente y con mayor precisión que una persona.

Dicho de otro modo, los humanos asignan probabilidades a todas las alternativas, mientras que el enfoque probabilístico con el que trabajan las máquinas muestra algunas ventajas frente a la incertidumbre, máxime porque, cuando ofrece una cognición probable, facilita la construcción de escenarios o historias, vincula cognición epistémica con la toma de decisiones binaria, de forma lógica, lo que no es sino una consecuencia de que las máquinas no tienen inteligencia, sino que son máquinas estadísticas (Cardon, 2018: 78). Este actuar conductual maquínico difiere del actuar humano; podemos pensar que la máquina ofrece una respuesta probable, pero la mente humana es capaz de desarrollar teorías allende las asignaciones estadísticas de probabilidad razonable a las que llegan las máquinas: la creatividad humana no ha sido —al menos hasta el momento— alcanzada por la máquina.

Ahora bien, la decisión humana no comporta una vía argumentativa monolítica, no existe la unidireccionalidad ni en la interpretación de los hechos jurídicos ni en la de las normas jurídicas. Las circunstancias de los hechos pueden alterar una respuesta humana jurídica. Y, por supuesto, no existe unidireccionalidad cuando, en la dialéctica propia de un procedimiento judicial, el juez se enfrenta a dos posiciones y recibe argumentos de ambas partes que las respaldan. El juez debe tomar una decisión judicial; no le basta la aplicación matemática de la norma porque lo que está en debate es si se aplica o no la misma y con consecuencias jurídicas. Será necesario aportar al juez razonamientos, dado que, precisamente, en la decisión juegan un papel crucial los argumentos externos e internos.

Los sistemas algorítmicos y las inteligencias artificiales incorporan un modelo lógico (o una pluralidad) que coinciden en mayor o menor medida con las estructuras del conocimiento legal, los patrones de razonamiento legal y la dialéctica de la interacción legal (Sartor, 2009: 26), ingredientes de los *software argumentativos, sistemas de razonamiento legal que permiten ofrecer respuestas contrarias a posiciones legales, doctrinales, de estrategias de defensa o acusación o interpretaciones judiciales.*

4.1. VALORACIÓN DE LA PRUEBA. CRITERIOS O ESTÁNDARES

Sin entrar en detalle en temas de prueba —que quedan, cada vez más, afectados por la irrupción de tecnologías de última generación capaces de influir en la creación probatoria e incluso en la validación o no de la fiabilidad de algunos medios de prueba—, centramos la atención en la «mediación discursiva» o proceso cognoscitivo que desarrolla el juez en

relación con los hechos probatorios y el *tema probandi*; esta mediación discursiva no se produce como mera «[…] descripción aséptica de lo que existe al margen de la actividad del operador, sino como proceso de construcción del supuesto de hecho, construcción a la que el juez contribuye activamente desde dentro. Y en ella está presente toda la carga de subjetivismo que acompaña a cualquier actividad interpretativa» (Andrés Ibáñez, 1992: 265). Nos hallamos ante uno de los ejes que anudan la prueba a la decisión judicial, a saber, la valoración judicial, en la que se integran la argumentación jurídica y la decisión judicial motivada.

Surge el dilema de si es factible estandarizar los criterios de valoración probatoria y si los mismos pueden algoritmizarse. *Prima facie,* en la individualización de la decisión judicial el juez determinará qué regla, principio o estándar, criterio o canon es el más apropiado para basar la decisión adoptada. Como apunta Taruffo, esta decisión no se puede racionalizar de forma anticipada, sino tan solo *ex post,* de manera que la decisión tomada sea razonable y válida de acuerdo con sus premisas (Taruffo, 1998: 320). En la medida en que tal racionalización pueda realizarse utilizando métodos de inteligencia artificial, pertenece al área del análisis de los modos de razonamiento por argumentos más que a la de la racionalización de las evaluaciones discrecionales. Esta afirmación, empero, encuentra algunos matices cuando de la prueba penal se trata, dado que preexisten estándares o criterios de decisión como la pena mínima y máxima del delito, la edad, la condición social, la situación psíquica, los antecedentes penales, la duda, entre otros. Se afirma que con la conjunción de estos criterios o estándares previos se produce una serie de elementos que pueden favorecer la racionalización de la decisión judicial, una racionalización que no se automatiza, sino en respuesta a la capacidad decisora del juez. La decisión encuentra sustento en componentes valorativos previos, a los que se ha denominado criterios o estándares probatorios.

Podríamos considerar que cuanto mayor sea número de estándares probatorios, más nos acercaremos a la matematización de la valoración de la prueba, a su automatización. Sin embargo, la función valorativa judicial en el proceso va más allá, dado que para determinar su fiabilidad considerará necesario conjugar conceptos como «credibilidad», «autenticidad», «aceptabilidad», que quedan en la esfera de la mediación discursiva del juez. Ello se consigue mediante el razonamiento probatorio judicial, que debe operar sobre la base de «[…] métodos acreditados como racionales, tomando en consideración todos los datos en presencia, con la atención dirigida a calibrar el rendimiento concreto de cada medio probatorio en elementos de prueba, con conciencia clara de los pasos inferenciales que los mismos permiten, a sabiendas que, a mayor número de estos, mayor será el riesgo de error» (Andrés Ibáñez, 2009: 56).

La decisión judicial que se adopte en el marco del proceso es uno de los elementos esenciales en la conformación procesal de la tutela judicial efectiva que el Estado garantiza a los ciudadanos. Ya sea en un proceso civil o en un proceso penal, la función jurisdiccional encuentra su núcleo en la decisión. Y esta, a su vez, viene conformada de una u otra forma en función de si hallamos en un sistema jurídico basado en el derecho legislativo, a saber, las normas, o en uno basado en el derecho casuístico, a saber, los antecedentes. Con

todo, las diferencias entre las «familias legales», como se las ha denominado a la hora de establecer los sistemas legales, muy remarcadas en tiempos pasados, han experimentado paulatinamente una clara intersección, fruto de su permanente contacto y de los procesos de retroalimentación recíproca.

En el *Common Law* se requiere de un sistema de rastreo, comparación y evaluación de los precedentes, actividades que pueden beneficiarse de la agilidad de las herramientas algorítmicas y los modelos de *machine learning*, generando lo que ha dado en denominarse *case-based reasoning* (CBR), orientado a la labor estadística de recopilación, análisis, y planificación de los precedentes. Por su parte, el modelo jurídico continental se nuclea en torno a las ideas de estabilidad y seguridad jurídica, necesarias para garantizar la tutela efectiva desde la igualdad y la objetividad. Se entiende que el mejor medio para garantizarlos es la ley. Aun cuando la norma es esencial en la toma de la decisión, la sentencia se dicta no solo en atención a las normas o a los precedentes legales, sino que también se considerarán las prácticas sociales, la organización y modo de actuar en un momento y en un lugar concreto: todos estos factores concurren en la configuración de la sentencia (Tata, 1998: 224-225). Incluso en los modelos jurídicos continentales se acepta la interpretación integradora de la norma, y, en ciertos supuestos, la analogía.

Extrapolado a la función decisora, se considera que el primer factor que concurre es el marco normativo, ya sea el de códigos o leyes o el de los precedentes vinculantes. Pero este factor no tiene una aplicación automática, pese a ese movimiento «pro» estandarización probatoria, proveniente de un modelo como el americano, en el que los estándares probatorios son el fruto de la actividad decisora del jurado, miembros legos que carecen de las habilidades y capacitación que tienen los jueces profesionales. Esos estándares pueden alimentar herramientas algorítmicas, ofreciendo una suerte de objetivación de la decisión judicial.

¿Podrían obtenerse por una inteligencia artificial estos estándares probatorios que llevaran a considerar que la falta de verosimilitud o de fiabilidad se asienta en una enorme duda razonable, fijando a este respecto criterios de esa duda razonable? ¿Podría un juez, en el ejercicio de su pensamiento jurídico, motivar una resolución judicial en un sentido diverso a aquel al que pareciera dirigirle la aplicación de la herramienta algorítmica?

Reiteramos que, allende la objetivación de la función judicial y la fijación de unos criterios objetivables, siempre quedará la posibilidad de incorporar y valorar factores subjetivos, que, entendemos, juegan un papel equilibrador de cara a la automatización de decisiones en el proceso. Si bien estos pueden impulsar los principios de eficacia y de eficiencia, en ciertos casos también pueden desdibujar las fronteras procesales y afectar a los derechos fundamentales de las personas. Por ejemplo, la herramienta australiana SIS (Sentencing Information System) fue diseñada para promover uniformidad de las sentencias; el programa trabaja con factores y trata de proporcionar datos estadísticos, así como información adicional. Puede ser útil en ciertos asuntos, pero no de forma general.

El dilema está presente. Quizás el trabajo conjunto de la máquina con el juez humano pueda mejorar ese pensar como juez y cumplir la función jurisdiccional. Podría pensarse

en un modelo algorítmico que diseñare componentes objetivos y objetivados, que se equilibrare con los componentes subjetivos interpretados por el juez humano. Ni la neutralidad maquínica es tan neutral, ni la neutralidad humana lo es. Habrá que trabajar para evitar que los prejuicios humanos y los sesgos maquínicos alteren la función responsable e imparcial que debe ejercerse por los tribunales de justicia.

Es más, si desde el punto de vista objetivo las bondades de la aplicación de la ley (norma o precedente) al caso concreto pueden presentarse como la panacea de un modelo de Justicia moderno, los riesgos que conllevan esta manera de «hacer o producir Justicia de fabricante» son palmarios. Además, no debemos olvidar el papel que juegan las «emociones» (Kahneman, 2011), el componente subjetivo de la emotividad, la afección y la sensibilidad en la adecuación de la valoración de la prueba y en la interpretación racional que dé lugar a la individualización de la decisión judicial. Es indiscutible que la discrecionalidad judicial racional está abriendo ante el juez la posibilidad —y la necesidad— de que, frente a las varias posibilidades de solución del caso, pueda elegir aquella que más le convenza. Esa decisión no es una elección en sentido estricto, sino un verdadero compromiso decisorio del juez; un compromiso con su función judicial, con con las partes y con la sociedad, lo que supone que, con la construcción de la sentencia, está otorgando tutela judicial, amén de asumir el compromiso de responder por la decisión tomada. En estos actos hay conciencia, hay conocimientos y hay subjetividad, es indudable, pero también hay responsabilidad.

La incorporación de los algoritmos a la Justicia se presenta como una fórmula neutralizadora de la subjetividad que envuelve la sentencia. El rigor de las matemáticas y la frialdad de los datos resolverían de forma automática, se dice, muchos de nuestros problemas. Olvidamos que los algoritmos no son neutros, que arrastran sesgos que pueden generar, lejos de esa objetivación de la Justicia, unas desigualdades poco justificables por extrapolación de prejuicios machistas, sexistas, xenófobos, clasistas, etc. Los sistemas computacionales, las estructuras inteligentes y los robots-juez son simulaciones, artefactos, pero no son humanos, no tienen pensamientos ni memoria, no dudan, sino que aplican información, aplican datos y normas de forma mecánica tomando en consideración la información recibida de numerosos casos pasados, de numerosas decisiones adoptadas, y un largo etcétera. El resultado no puede ser que el sistema algorítmico piense como un juez, aunque aplique una solución al caso concreto una solución como un juez. Tampoco puede afirmarse que el algorítmico sea capaz de sustituir al juez, pues al algoritmo ha sido diseñado, preparado y alimentado por humanos, también en lo que respecta a la simulación de la función decisoria judicial (Barona, 2021: 621).

4.2. DECISIÓN JUDICIAL = VALORACIÓN PROBATORIA + ARGUMENTACIÓN + MOTIVACIÓN CON PROPORCIONALIDAD Y DERECHOS

La conformación y toma de la decisión judicial en el seno de un proceso para dar una respuesta a los ciudadanos y ciudadanas que requieren la tutela judicial efectiva es el resultado de una acción que va de la mano de la actividad judicial de valoración de la prueba que aportada al proceso (ora por las partes, ora prueba de oficio), que necesitará de una

argumentación que, traducida en la resolución judicial, motive, justifique, fundamente el resultado judicial alcanzado propio del pensamiento jurídico (Taruffo, 1998: 312). Este pensamiento no es volcado de forma «bruta» en la resolución, sino que es conformado de forma «neta» tras la ponderación de los derechos concurrentes de las partes procesales, amén de los principios específicos del proceso civil o del proceso penal, y lo hará a través de un enfoque híbrido narrativo-argumentativo del razonamiento probatorio (Bex y Verheij, 2012: 252-258). Y esa expresión vendrá de la mano de la motivación y la justificación de la valoración de la prueba y las conclusiones sobre los hechos.

El dilema está presente, y genera dudas acerca de la extrapolación de la actuación decisora judicial humana a los procesos de matematización automatizada. Cada vez se crean más herramientas algorítmicas que permiten la resolución automatizada de cuestiones que, muy probablemente, pueden reducirse a parámetros estadísticos y que, tal como se constató *supra*, van perfeccionándose en su tratamiento. Inicialmente, fueron sistemas que pretendían alcanzar una cierta uniformidad de las decisiones judiciales, y se programaron las herramientas que sucedieron a las denominadas *sentencing guidelines,* surgidas como colaboradoras, aunque fueron condicionando obligatoriamente las decisiones judiciales (Tonry, 1996: 72-75).

La búsqueda economizada de soluciones estandarizadas, uniformes, ha sido una constante en la búsqueda de agilidad y eficiencia en la toma de decisiones, objetivo que no siempre casa con la proporcionalidad y los derechos. Aquellas generan serias dudas porque constituyen una suerte de motivación tipo. La perversión de esta automatización se ha experimentado en algunos supuestos de adopción de medidas cautelares privativas de libertad o privativas de derechos en el proceso penal, máxime cuando los sistemas de predictibilidad se convierten en el soporte de adopción de las medidas a través de la incorporación de herramientas algorítmicas que evalúan criterios que pueden ser altamente discriminatorios y que generan posibles efectos segregadores. Las consecuencias son muy peligrosas.

En el modelo anglosajón, las primeras herramientas que se programaron daban continuidad a la función de las directrices o *guidelines.* Un exponente ya citado es ASSYST, herramienta concebida para ayudar a los jueces a aplicar las directrices en la sentencia. Se criticó que abocaba a una aplicación automática de la condena, lo que no siempre conveniente, adecuado y, por supuesto, justo. El desarrollo de estas herramientas no ha cesado, especialmente las que trabajan con sistemas computacionales (estadísticos) que ofrecen la valoración de la predictibilidad criminal a efectos de condicionar la posible condena y el «grado» de la misma. Basta pensar en conocidas herramientas como COMPAS, que ha generado una ingente literatura que no procede citar aquí.

La aplicación de estos sistemas asistenciales debe considerarse como tal, es decir, no deberían ser vinculantes para los jueces. Se trata, más bien, de instrumentos de colaboración que permiten trabajar con los factores o elementos que, considerados estadísticamente, pueden arrojar un pronóstico de riesgo. En suma, una herramienta de la que disponen los jueces a la hora de dictar sentencia, sino también cuando deciden cautelarmente, ya sea en la jurisdicción penal o en la civil. No debe olvidarse que toda decisión judicial afecta a las personas y sus derechos, de ahí la enorme importancia que tiene que todos los sistemas

computacionales que directa o indirectamente tengan un resultado en la tutela de las personas deben haber sido controlados, evaluados y auditados.

En ese proceso cognoscitivo, que comporta la labor deductiva e inductiva del juez humano, han ido penetrando estos sistemas algorítmicos, que indicen, influyen, alteran o corroboran las posibles decisiones judiciales. Pese a todo, aún estamos en el escalón decisorio humano, aunque venga condicionado por la algoritmización de la Justicia. Sin embargo, el dilema continúa.

4.3. DECISIÓN JUDICIAL MAQUÍNICA O JUEZ ROBOT, ¿ES FUNCIÓN JURISDICCIONAL DE TUTELA EFECTIVA?

En este devenir maquínico y en la fusión entre tecnología y ciencia emerge la imparable inteligencia artificial, que evoluciona generación tras generación incorporando poco a poco los desarrollos de la computación cuántica. Particular importancia tiene la irrupción de las denominadas «inteligencias artificiales generativas», de las que tanto informan diariamente los medios de comunicación. Asistimos a un desarrollo constante de la IA en todos los sectores, especialmente en la medicina, la economía, la banca, la agricultura, el modelo de ciudad, la movilidad, la gestión, la contabilidad jurídica, la fiscalidad o la contratación, entre muchos otros. Por supuesto, también el modelo de Justicia que ha emergido en el siglo XXI ofrece la incorporación de modelos de inteligencia artificial en las tareas de gestión y tramitación. Asimismo, han surgido, siquiera de forma discontinua, modelos automatizados de respuestas jurídicas en sectores como la solución de conflictos que se tramitan a través de un procedimiento como el monitorio, o en cuestiones reiterativas y repetitivas como el consumo. En España, la aprobación del RDL 6/2023, de 19 de diciembre, ha supuesto la incorporación de un abrigo legislativo a instrumentos tecnológicos de gestión, información, comunicación (se han diseñado la Carpeta Justicia, el PAGAJ, la Sede Judicial Electrónica, Visor Horus, el quiosco biométrico, el EVID para cita previa, etc.); pero también se ha abierto la posibilidad de incorporar actuaciones automatizadas proactivas y asistidas, siempre con el debido respeto a las leyes procesales y bajo los criterios legales objetivos y públicos, y se acepta ya el uso de herramientas algorítmicas e inteligencias artificiales que puedan asistir a los jueces, como por ejemplo VioGén en los supuestos de violencia de género o RisCANVI en la ejecución de sentencia, así como otras que permiten gestionar depósitos judiciales y las alarmas de plazos en las medidas cautelares, y que incluso podrían ser de utilidad para garantizar la localización de bienes del deudor o del responsable civil en el proceso penal. Por último, hay algunas herramientas que están empleando como proyectos piloto (búsqueda de personas desaparecidas, cancelación de antecedentes penales).

Pero también hemos empezado a aceptar la usabilidad de la inteligencia artificial más allá de los casos expuestos. En algunos países (China, Países Árabes, Estonia, etc.) han aparecido, no sin resistencias, los robots-juez y la IA-decisora. En estos supuestos los sistemas computacionales inteligentes judiciales sustituyen a los jueces, evidenciando que el fenómeno de la robotización judicial es una realidad, siquiera sectorial, local, experimental, pero una realidad. Incluso hay quien sostiene que la idea nuclear de la inteligencia artificial es «eliminar a los humanos de la ecuación» (Zavrsnik, 2020: 568).

Prima facie, cuando un sistema computacional «inteligente» otorga una solución decide, está realizando función jurisdiccional, está otorgando tutela a quien la reclama. Esta afirmación general reclama aclaraciones que operan como premisas de la reflexión ulterior:

1) Desde el punto de vista tecnológico, no todos los sistemas computacionales llamados «IA» son aptos para ejercer esta misión, o al menos no con garantías. Estos sistemas deben estar preparados tecnológica y funcionalmente para gestionar y desarrollar de forma automatizada un procedimiento y para poder ofrecer un análisis de datos (datos del pasado, en todo caso), información, documentos, legislación, doctrina, jurisprudencia, para, desde su explotación ofrecer respuestas una vez realizada la explotación de la información. Los riesgos existen y son diversos. Un marco jurídico de enorme interés en el seno de la Unión Europea es el que se ha abierto con la aprobación del Reglamento de Inteligencia Artificial el 13 de marzo de 2024 que condiciona la viabilidad de la IA-decisora judicial (como las demás IA predictivas, clasificadoras o perfiladoras, de recomendación, etc.) al debido respeto de los seis principios estructurales: intervención y vigilancia humanas, solidez y seguridad técnicas, privacidad y gobernanza de datos, transparencia, diversidad, no discriminación y equidad, y bienestar social y medioambiental; principios que responden al sentir de la Carta de los Derechos Fundamentales de la Unión Europea (CDFUE). Y, además, deberán garantizar: la transparencia en su uso, su eficiencia energética y medioambiental, la gestión de calidad y de riesgos o su inscripción en un registro europeo, además de las obligaciones específicas de los desarrolladores que se configuran el texto europeo, teniendo en cuenta el grado de riesgo que conlleva, de manera que, si se trata de un sistema de alto riesgo, tiene que considerarse aplicable su régimen específico (Barrio Andrés, 2023: 3-4).

En suma, para ofrecer garantías a los justiciables es necesario que, con carácter general, se incorporen criterios de viabilidad tecnológica: a) por un lado, es necesario introducir mecanismos que permitan realizar la validación de la herramienta y su usabilidad en el ámbito judicial, por lo que se trata de un control *ex ante*; b) por otro lado, hay que configurar medidas *ex post* que permitan realizar una suerte de auditorías que mantengan la validez y viabilidad de la IA en el contexto judicial correspondiente.

2) La validación inicial y la valoración que se realice *ex post* no lo es de forma genérica o simplemente tecnológica, sino que se circunscribe a la validación y viabilidad de una herramienta que puede gestionar y dirigir claramente un procedimiento de forma automatizada, que es capaz de analizar datos, información y documentos y explotar la información ofreciendo una suerte de justificación que simula la decisión judicial asentada en una valoración de las aportaciones de las partes. Sin embargo, no por ello —creemos— la estructura algorítmica está capacitada para tomar libre y emocionalmente decisiones a través de razonamientos no estadísticos, sino integradores de los hechos —las pruebas—, las percepciones y también las emociones. Mucho más importante: en la intervención humana existe la posibilidad de rectificar, de enmendar, de aclarar, acciones posibles que se alejan de la actuación —al menos hasta hasta la fecha— de los sistemas computacionales de inteligencia.

3) Un componente ineludible: las máquinas, a diferencia de los humanos, no tienen emociones, aunque se trabaje y avance en relación con los desarrollos neurocientíficos que permitan volcar una suerte de reacción y activación de la IA ante la maquínica percepción de las emociones. La máquina no piensa ni siente, funciones propias del ser humano, lo que no es óbice para que puedan ser preparadas con el fin de que arrojen resultados de intelecto basados en el perfeccionamiento de un conjunto de tecnologías que lo favorecen, así como para valorar e interactuar ante las reacciones de contrario.

La otra cara de la moneda de la reflexión es precisamente la que defienden quienes consideran que las percepciones, las emociones, los estados de ánimo, los sentimientos, etc. pueden ser negativos en ciertos casos a los efectos de tomar determinadas decisiones y realizar determinados actos. Este argumento lleva a defender que la frialdad de la máquina puede ser altamente positiva en determinados supuestos. Precisamente en el ámbito que nos ocupa, la toma de decisiones judiciales puede presentarse como una gran ventaja. Estos argumentos permiten mirar hacia el futuro en busca de lo que algunos filósofos han denominado el «superyó» (Sadin, 2019: 141-148), en línea con el «superhombre» nietzschiano, figura a la que se ha recurrido para afirmar que las máquinas son más objetivas que los humanos a la hora de tomar decisiones.

4) Desde esta mirada se defiende en foros nacionales e internacionales de inteligencia artificial que la irrupción de esta en la toma de decisiones judiciales supone una mejora del sistema jurídico, al delimitarse sectorialmente lo que puede derivarse a la máquina y lo que debe ser objeto de raciocinio, reflexión, interpretación, valoración y decisión motivada por el juez humano. La IA-robot se convierte, así, en un modelo judicial sin seres humanos, pero también sin prejuicios humanos. Este objetivo favorece la búsqueda de una arquitectura algorítmica, alimentada por un *big data* sin precedentes, pero ya no necesariamente en cantidad, sino en calidad, que permita, que el juez robot pueda buscar en su memoria para conformar la sentencia o dictar cualquier otra resolución en el proceso. Es una idea fascinante emerge con cada vez con más fuerza en determinados ámbitos.

Ahora bien, seguimos preguntándonos: ¿qué es la justicia aplicada al caso concreto? ¿La aplicación matemática de la norma? ¿La aplicación del precedente? ¿Puede ser siempre esta vía «perfecta» automatizada la divinidad judicial que buscamos? ¿Se trunca con ello la individualización decisora?

Nada se niega o se afirma rotundamente porque no es posible hacerlo. En la evolución de los últimos lustros hemos asistido a un imparable modelo IA-decisora judicial que no solo pretende, no solo responder a través del programa computacional adecuado a la automatización y decisión judicial, sino también caminar hacia la programación de jueces robots que actúen de forma individualizada, interpretando allende los antecedentes, las sentencias dictadas y las normas existentes, a través de mecanismos tecnológicos de valoración probatoria asentados en estándares probatorios, de manera que permitan resolver —función judicial— de una manera adecuada, específica y determinada según las circunstancias del caso y los componentes objetivos y los subjetivos específicos que concurran. Se apunta que el juez robot puede diseñarse contando con toda la información que sea absolutamente relevante y necesaria para dictar las sentencias; por otra parte, se habla de

la posibilidad de diseñar los robots jueces para que puedan considerar y aplicar la idea de proporcionalidad, toda vez que esta puede integrarse en una suerte de casuística abundante, creciente y dispar que permita al robot optar por la disyuntiva entre aplicar de forma automatizada la norma-precedente dictando una decisión de forma, en línea con lo que ha venido siendo constante, o bien adaptar la decisión a las reglas de la proporcionalidad, lo que no significa que el robot «sepa» lo que es proporcional, sino que se le ha programado para aplicar decisiones proporcionadas en función de la alimentación de un gran número de situaciones anteriores proporcionadas o desproporcionadas.

La programación de los jueces robots no es nada fácil, pero, desde luego, es una posibilidad cada vez más cercana. Se dice que en estos casos es absolutamente imprescindible trabajar con la *deep learning*, integrando la base de datos con redes neuronales para entrenar a los robots jueces en el dictado de sentencias similares a las de los jueces profesionales. Son sentencias, son decisiones, son manifestación de la función jurisdiccional en lo que se anuda con la tutela efectiva reclamada. Pese a todos estos avances y la mirada puesta en el «efectivismo» imperante, el entorno en el que nos encontramos y, muy especialmente, la aprobación del Reglamento IA de la UE ha venido a obstaculizar la viabilidad de la IA-robot independiente y sin intervención humana cuando afecta a derechos fundamentales.

5) Una de las cuestiones que obstaculizan los avances masivos de sistemas computacionales decisorios judiciales o IA judicial es la conformación política y, en la mayor parte de los sistemas jurídicos constitucionales, las garantías que la función jurisdiccional ofrece en el marco de la salvaguarda de los jueces humanos como tales, a saber, su propio estatuto judicial. Se plantea cómo garantizar la independencia, la responsabilidad y la imparcialidad del juez robot, o todo cuanto se refiere a la determinación competencial, así como el reconocimiento constitucional del juez legal o predeterminado. Todas estas cuestiones son planteadas ante la acrítica fascinación que muestran algunos, lo que no supone, empero, un rechazo frontal a su viabilidad. Al contrario, se trata, además de establecer las condiciones para su viabilidad y eficacia, de configurar un marco jurídico adaptado a estos sistemas que permita establecer las garantías jurisdiccionales también de la máquina; algunas de estas garantías deberán regularse en las normas procesales, configurando las variables conceptuales de principios como la inmediación digital, la responsabilidad maquínica, el juez legal o predeterminado, la oralidad, la independencia judicial o la motivación judicial, entre otros, sin perjuicio de la dificultad añadida de que estamos ante un sector (tecnológico) cuyos desarrollos y progresos se despliegan a una velocidad que difícilmente puede simultanearse con las modificaciones legales.

Obviamente, también en la programación debe considerarse la necesidad de configurar IA judicial que sea imparcial, independiente y responsable para garantizar el ejercicio de la función judicial que desempeñan. Pero surgen algunas cuestiones. ¿Puede garantizarse esta independencia judicial maquínica? ¿Puede programarse la misma para actuar solo en aplicación de normas, o de precedentes, y no actuar bajo presión o mandato externo? Los defensores de la incorporación del robot-judicial afirman que es más sencillo garantizar la independencia judicial y la imparcialidad con las máquinas que con los humanos. Asimismo, consideran que el sistema de diseño, programación y actuación es objetivo y objetiva-

ble —a diferencia de lo que sucede con los humanos— y que se pueden diseñar máquinas independientes e imparciales. Los menos entusiastas refutan esta afirmación argumentando que, del mismo modo que se puede programar a las máquinas para ser imparciales e independientes, se puede «educar» a humanos para ser jueces independientes e imparciales. Además, en segundo lugar, debe buscarse la IA explicada, esto es, aquella que permita conocer el porqué de sus resultados.

En todo caso, habrá que retomar la vieja idea de hace unos años de delimitar una suerte de responsabilidad maquínica por sus decisiones, y crear el estatuto de la persona electrónica para determinar, eventualmente, las responsabilidades por sus actuaciones. Ciertamente, uno de los debates complejos en materia de responsabilidad es si lo que se pretende es determinar la viabilidad de la responsabilidad directa. La doctrina ha sugerido la posible integración o inclusión de la construcción jurídica de la responsabilidad trasladable a la máquina, planteando lo que se ha denominado como el «pseudocomportamiento» de los sistemas computacionales inteligentes, semejante a los que se producen debido a la negligencia humana (Schuhr, 2014: 23-24). Este debate se ha centrado especialmente en el caso de los denominados vehículos autónomos, si bien cabría igualmente plantearlo en relación con el tema tratado en estas páginas, máxime los jueces pueden tomar decisiones por aplicación de herramientas algorítmicas y si, como ha sucedido en algunos sistemas, pueden crearse plataformas judiciales o robots propositivos o decisorios.

Persisten, sin embargo, numerosas dudas tanto en relación con la responsabilidad de las personas hibridadas como respecto de la IA judicial decisora o juez robot. Cuando se trata de decisiones «hibridadas», hay que estar a lo que se regula en relación con la responsabilidad civil y con la responsabilidad penal. Cada vez más se siente la necesidad de integrar una específica regulación de la responsabilidad de las máquinas, en la medida en que comporta el equilibrio entre la necesidad de configurar un alto nivel de protección y promocionar y facilitar con garantías la innovación. Ahora bien, debe tratarse de una responsabilidad que se integre en las diversas modalidades de responsabilidad legalmente establecidas, en las que se conjuguen riesgos e incentivos de prevención óptimos. Sin embargo, no es factible hoy por hoy justificar la responsabilidad de la máquina más allá de los requisitos legalmente establecidos, que permitan su motivación (Hanisch, 2013: 60-61). Se habla, también, de que el tipo de responsabilidad que puede exigirse a la IA decisora judicial es contractual objetiva, por incumplimiento o cumplimiento defectuoso de una obligación de hacer o de dar. Cuestión harto compleja es establecer el reparto de riesgos del incumplimiento o cumplimiento defectuoso cuando la máquina actuó como sistema experto, colaborando o auxiliando al ser humano. Sobre todo, cuando se trata de artefactos encuadrados en el mundo de la robótica y más especialmente cuando se trata de robots corpóreos, parece interesante ahondar en materia de responsabilidad en la necesidad de contratar un seguro obligatorio de responsabilidad y podríamos incluso llegar a incorporarlo como categoría de la responsabilidad extracontractual (Barona, 2022*b*: 21).

En atención a algunas de las cuestiones que afectan al estatuto de jueces y a su proyección en la función jurisdiccional y en la función tuitiva de la ciudadanía, debe concluirse que, efectivamente, la IA decisora judicial ejerce la función de tutela, resuelve, aunque

no siga los pasos de la decisión judicial humana: valoración de los hechos y su incidencia probatoria, interpretación y argumentación (razonamiento judicial) que se plasme en la motivación —exigencia constitucional— con proporcionalidad y respeto de los derechos. La decisión judicial maquínica responde a la matematización de la tutela, al menos hoy por hoy. La ausencia de asertividad propia de los humanos (o de algunos), amén del marco jurídico garantista «conquistado» en los modelos jurídicos democráticos deberá encontrar paralelo en las actuaciones judiciales maquínicas, en sus instrumentos, en sus decisiones y también en la manera de contestar esas decisiones. Se trata, indudablemente, de un reto de la sociedad en la que vivimos. Es innegable que la sofisticación tecnológica nos sorprende diariamente con nuevos modelos computacionales de inteligencia artificial de primera, segunda y tercera generación que pueden y deben favorecer un mundo mejor, también un mundo judicial. Lejos de caer en férreas posiciones tecnófobas, como el movimiento del bioconservadurismo —que defiende que el avance de la tecnología entraña peligros para la humanidad y para su supervivencia—, hay que aprovecharse de ella, y defender que estamos en un momento idóneo para reinterpretar el significado del propio ser humano o, lo que es lo mismo, para replantearnos el significado del lenguaje, la comprensión, la conciencia, las emociones y la capacidad crítica, entre otros aspectos. Es buen momento para insuflar grandes dosis de humanidad imperfecta a la evolución del propio ser humano, ese animal racional y atractivamente imperfecto (Barona, 2021: 663-664).

BIBLIOGRAFÍA

ALEXY, Robert (1978): *Theorie der juristischen Argumentation*, Frankfurt: Surkamp.

ANDINO LÓPEZ, Juan Antonio (2017): «Los criterios Daubert y su posible encaje en la valoración de la prueba pericial en España», en J. Picó Junoy (dir.), *Peritaje y prueba pericial,* Barcelona: Bosch.

ANDRÉS IBAÑEZ, Perfecto (1992): «Acerca de la motivación de los hechos en la sentencia», *Doxa*, 12.

— (2009): *Prueba y convicción judicial en el proceso penal,* Hammurabi: Buenos Aires.

ASHLEY, Kevin. D. (2017): *Artificial Intelligence and Legal Analytics,* Cambridge: Cambridge University Press.

ATIENZA RODRÍGUEZ, Manuel (2006): *El Derecho como argumentación,* Barcelona: Ariel.

BARONA VILAR, Silvia (2021): *Algoritmización del derecho y de la justicia. De la Inteligencia Artificial a la Smart Justice,* Valencia: Tirant lo Blanch.

— (2022*a*): «La digitalización y la algoritmización, claves del nuevo paradigma de justicia eficiente y sostenible», en I. Colomer Hernández (dir.), *Uso de la información y de los datos personales en los procesos: los cambios en la era digital*, Navarra: Thomson Reuters Aranzadi.

— (2022*b*): «Persona, algoritmización y posthumanismo, una ecuación hacia la "persona maquínica" y su responsabilidad», *Revista Actualidad Civil,* 10.

BARRANCO AVILÉS, M.ª del Carmen (2010): «Juzgar y hacer ejecutar lo juzgado. Sometimiento a la ley, independencia e imparcialidad en el marco del constitucionalismo», *Teoría & Derecho*, 7.

BARRIO ANDRÉS, Moisés (2023): «Novedades en la tramitación del próximo reglamento europeo de inteligencia artificial», *Real Instituto Elcano*, ARI 67/2023, 19 de julio.

BEX, Floris y VERHEIJ, Bart (2012): «Legal stories and the process of proof», *Artificial Intelligence and Law*, 21, 253-278.

CALAMANDREI, Piero (1960): *Proceso y democracia* (traducción de H. Fix Zamudio), Buenos Aires: Ejea.

Ferrajoli, Luigi (1980): *Diritto e ragione. Teoria del garantismo penale*, Roma: Laterza.
CARDON, Dominique (2018): *Con qué sueñan los algoritmos*, Madrid: Dado.
DWORKIN, Ronald (2006): *Justice in Robes*, Harvard: Harvard University Press.
DYMITRUK, Maria (2019): «The Right to a Fair Trial in Automated Civil Proceedings», en *Masaryk University Journal of Law and Technology,*13, 1, 27-44.
GABRIEL, Markus (2019): *El sentido del pensamiento*, Barcelona: Pasado&Presente.
GREENLEAF, Graham (1991): «Information technology and the law», *Australian Law Journal, 65*.
GUZMÁN FLUJA, Vicente (2017): «Sobre la aplicación de la inteligencia artificial a la solución de conflictos», en S. Barona Vilar (dir.), *Justicia civil y penal en la era global*, Valencia: Tirant lo Blanch.
— (2021): «Proceso penal y Justicia automatizada», *Revista general del Derecho Procesal,* 53.
HANISCH, Jochen (2013): «Zivilrechtliche Haftungskonzepte für Roboter», en E. Hilgendorf y JP. Günther (eds.), *Robotik und Gesetgebung,* Baden-Baden: Nomos.
KAHNEMAN, Daniel (2011): *Thinking, Fast and Slow,* Farrar: Straus and Giroux.
LEITH, Philip (1998): «The Judge and the Computer: how best "Decision Support"?», *Artificial Intelligence and Law,* 6.
LLEDÓ, Emilio (2011): *El silencio de la escritura*, Barcelona: Espasa.
NIEVA FENOLL, Jordi (2018): *Inteligencia artificial y proceso judicial*, Madrid: Marcial Pons.
NISSAN, Ephraim (2009): «Legal Evidence, Police Intelligence, Crime Analysis or Detection, Forensic Testing, and Argumentation: An Overwie of Computer Tools or Techniques», *17 Int´l J.L. &Info. Tech*, 1.
SADIN, Eric (2019): «La inteligencia artificial: el superyó del siglo XXI», *Revista Nueva Sociedad*, 279,
SANCHO GARGALLO, Ignacio (2020): «Judge craft: el oficio o arte de juzgar», *InDret,* 4.
SARTOR, Giovanni (2009): «Deafasibility in Legal Reasoning», *EUI Working Paper LAW,* Florence, 2009/02.
SCHWAB, Klaus (2018): *La cuarta Revolución Industrial*, Barcelona: Debate.
SCHILD, Uri J. (1998): «Criminal Sentencing and Intelligent Decision Support», *Artificial Intelligence and Law*, 6.
SCHUHR, Jan C. (2014): «Neuedefinition tradierter Begriffe (Pseudo-Zurechnungen an Roboter)», en E. Hilgendorf (ed.), *Robotik im Kontext von Recht und Moral,* Baden-Baden: Nomos.
SIMON, Eric y Gerry GAES (1989): «ASSYST – computer support for guideline sentencing», en *Second International Conference on Artificial Intelligence and Law (ICAIL-89*), Vancouver: ACM Press, 159-200.
TARUFFO, Michele (1986): «La fisonomia della sentenza in Italia», *Rivista Trimestrale di Diritto e Procedura Penale*.
— (1998): «Judicial Decisions and Artificial Intelligence», *Artificial Intelligence and Law,* 6.
TATA, Cyrus (1998): «The Aplication of Udicial Intelligence and "Rules" to Systems Supporting Discretionary Judicial Decision-Making», *Artificial Intelligence and Law,* 6.
TONRY, Michael H. (1996): *Sentencing Matters*. New York-Oxford: Oxford University Press.
WITTGENSTEIN, Ludwig ([6]1984): *Tractatus logico-Philosophicus*, Madrid: Alianza Universidad, versión alemana-española.
ZAVRŠNIK, Ales (2020): «Criminal justice, artificial intelligence systems, and human rights», *ERA Forum* 20[en línea] <https://doi.org/10.1007/s12027-020-00602-0>. [Consulta: 15/02/2024].

Fecha de recepción: 1 de julio de 2024.
Fecha de aceptación: 20 de octubre de 2024.

INTELIGENCIA ARTIFICIAL Y PROCESO PENAL; ESPECIAL REFERENCIA A LAS MEDIDAS CAUTELARES

ARTIFICIAL INTELLIGENCE AND CRIMINAL PROCEEDINGS; SPECIAL REFERENCE TO PRECAUTIONARY MEASURES

Andrea Planchadell Gargallo

Catedrática de Derecho Procesal
Universidad Jaume I (Castellón)

RESUMEN

El uso masivo de las nuevas tecnologías no solo influye en nuestra vida diaria, sino también en el funcionamiento de la Administración de Justicia. Los importantes cambios que está experimentando la Administración de Justicia con la entrada de la digitalización se ven superados por la irrupción de la inteligencia artificial y su influencia en la actividad de los distintos operadores jurídicos. Si bien no son pocas las aplicaciones de la inteligencia artificial en el campo de la Justicia, la posibilidad de que esta tecnología pueda auxiliar al juez en la toma de decisiones —e incluso sustituirlo— es una las mayores preocupaciones actuales. Tras unas consideraciones de carácter introductorio, en estas páginas se analiza el uso de la inteligencia artificial en el campo de las medidas cautelares penales.

PALABRAS CLAVE

Inteligencia artificial, justicia predictiva, medidas cautelares, proceso penal.

ABSTRACT

The massive use of new technologies not only influences our daily lives, but also the way the Administration of Justice works. The important changes the Administration of Justice is experiencing with the arrival of digitalization are being surpassed by the emergence of artificial intelligence and its influence on the activity of the different legal operators. Although there are many applications of artificial intelligence in the field of Justice, the possibility that it may assist the judge in decision-making —and even replace him or her— is one of the greatest concerns today. After some introductory considerations, these the use of artificial intelligence in the field of criminal precautionary measures is considered in these pages.

KEYWORDS

Artificial intelligence, predictive justice, precautionary measures, criminal procedure.

DOI: https://doi.org/10.36151/TD.2024.110

INTELIGENCIA ARTIFICIAL Y PROCESO PENAL; ESPECIAL REFERENCIA A LAS MEDIDAS CAUTELARES

Andrea Planchadell Gargallo
Catedrática de Derecho procesal
Universidad Jaume I (Castellón)

1. INTRODUCCIÓN

En los últimos años, la Justicia se ha enfrentado a una realidad cada vez más cambiante, «líquida» y, en algún sentido, incierta que invita a plantear si está preparada para hacer frente a los cambios en curso —no pocas veces drásticos— y de qué modo debería hacerlo. Entre estas transformaciones sobresale el uso masivo de las nuevas tecnologías, consecuencia de nuestra total inmersión en la denominada cuarta revolución industrial o industria 4.0 (Barona, 2019*a*: 2; Barona, 2018: 494). Uno de los retos más desconcertantes a los que la sociedad en general y la Administración de Justicia en particular deben hacer frente es la irrupción y el auge de la inteligencia artificial, y, más específicamente, la aplicación de sistemas expertos, algoritmos (puede, de hecho, hablarse de la sociedad algorítmica, entendida como aquella que se organiza mediante la adopción decisiones algorítmicas y «se alimenta» de datos; cfr. Balkin, 2017: 1219; Goodman, 2016: 9 y ss.; Barona, 2019*b*: 24; Cotino, 2017: 131-132) y modelos computacionales al asesoramiento, la predicción y la toma de decisiones judiciales.

Las novedades asociadas a la cuarta revolución industrial (Shwab, 2016; Barona, 2019*a*: 2 ss.; Corvalán, 2018: 296), al *big data* o al internet de las cosas (IoT) no pueden ser obviadas por el Derecho[1]; al contrario, el Derecho tiene que adaptarse a ellas con todas las cautelas que sea necesarias. De hecho, son muchas las voces que han señalado que el Derecho —al igual que otros ámbitos, entre ellos la medicina— es un campo muy adecuado para la aplicación de la inteligencia artificial (Rissland *et al.*, 2003: 1 ss.; Ashley, 2018: 5; Barrio, 2018: 1). A este respecto, hacemos nuestras las palabras de Silvia Barona cuando afirma que el desafío al que nos enfrentamos —y al que debemos dar respuesta— es que «[…] estos avances se desplieguen como servicio al ser humano y no al revés, o lo que es lo mismo, que se presenten como herramientas que permitan lograr un futuro más próspero y confortable, más igualitario y humano, más solidario y justo. Por ende, sirvan a la Justicia para ser más justa» (Barona 2019*a*: 7). Una consideración de tenor similar puede encontrarse en el *Libro Blanco sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza*[2].

Indudablemente, la irrupción de la inteligencia artificial en la esfera jurídica lleva aparejados cambios sustanciales en la producción, la interpretación y la aplicación del Derecho (Nieva 2018: 19; Magro, 2018: 2; Bueno, 2020; Casabona, 2020: 253). Ahora bien, esta nueva modulación de la praxis jurídica no implica que todo valga, dado que nuestro sistema se sustenta en el respeto a los derechos y garantías esenciales de los ciudadanos. Es preciso añadir que, al abordar la cuestión del modo en que el Derecho ha de adaptarse a tales avances, tenemos que ser conscientes de que las leyes no deben —o no solo deben— referirse a los robots y las máquinas, sino también a las personas que operan con ellos: somos nosotros los que diseñamos los algoritmos, los programamos con datos que organizamos y seleccionamos, y los conectamos con las bases de datos para decidir cómo usarlos, cuándo y con qué finalidad (Balikin, 2017: 1221)[3].

En estas páginas prescindiremos de realizar precisiones conceptuales sobre la inteligencia artificial, asunto cuyo abordaje no es sencillo (Guzmán, 2017: 73 y ss.; Taruffo, 1998: 311; Bujosa, 2014: 62; Corvalán, 2018: 305; Planchadell, 2021: 389; Simon, 2021: 48), así como sobre los sistemas de expertos, los algoritmos y el resto de conceptos clave que entran en juego en el estudio de esta temática, dado que el análisis conceptual excede el propósito de este texto y las capacidades de quien lo firma.

2. BREVE REFERENCIA A LOS POSIBLES ÁMBITOS DE APLICACIÓN

En este epígrafe nos limitaremos a enumerar algunos usos actuales de la IA en el ámbito jurídico y, particularmente, en el campo del Derecho procesal. Al acometer esta tarea, debemos ser conscientes de que nos encontramos ante una materia en constante desarrollo y evolución. De hecho, desde que aparecieron las primeras bases de datos jurídicas, que en su momento ya supusieron un importante progreso en el manejo de datos e información, se ha avanzado enormemente en este campo.

Hay que distinguir dos modalidades —o, si se quiere, dos momentos— en la aplicación de la inteligencia artificial al Derecho procesal (Guzmán, 2017: 69; Barona, 2019*a*: 244-245; Bonet, 2020: 16 ss.): *i)* la utilización de herramientas expertas que *ayudan* a tomar decisiones judiciales; y *ii)* el uso de herramientas expertas que *sustituyen* al ser humano en la toma de decisiones. Estos dos extremos ponen de manifiesto la necesidad de que nos planteemos hasta dónde queremos llegar en la aplicación de la inteligencia artificial; cuestión que puede ser reformulada en estos términos: ¿debemos aceptar el tránsito de un estadio en el que estas herramientas auxilian, en mayor o menor medida, a los operadores jurídicos —e incluso a los usuarios de la Administración de Justicia— a otro en el que la tecnología sustituya a los operadores jurídicos tradicionales?

Diversos autores (Magro, 2018: 3 y ss.; Nieva, 2018: 24 y ss.; Solar, 2019: 92 y ss.; Goodman, 2016: 12 y ss.; Casanova, 2010: 207 y ss.; Casanova 2013: 473 ss.; Vega, 2018:21 ss.; Cáceres, 2006: 606 y ss.; Ponce, 2018: 6 y ss.; Belloso 2013: 139 y ss.; Bonet, 2018: 73 y ss.; Cerrillo y Velasco, 2019: 293 y ss.; Planchadell, 2021: 389; Conde, 2023: 105; Delgado, 2020: 525) han señalado que actualmente la inteligencia artificial se aplica a los siguientes ámbitos de la praxis jurídica:

i) La denominada *legal research*, cuya finalidad es obtener información relevante a través de bases de datos cada vez más sofisticadas para alcanzar distintos objetivos, entre ellos dar respuesta a un problema jurídico. Se trata de uno de los ámbitos pioneros en la aplicación de la tecnología que, tras la irrupción de la inteligencia artificial, ha conocido desarrollos y perfeccionamientos destacables[4]. Como sabemos, estos sistemas almacenan una cantidad ingente de datos, pueden gestionarlos con gran rapidez y precisión, y suministran modelos de respuesta y patrones de solución para resolver una controversia jurídica (Solar, 2019: 106, hace referencia a la minería de argumentos, una fase de desarrollo de la tecnología informática que va más allá de la denominada minería de datos).

ii) Los procedimientos de revisión, análisis y seguimiento de la ejecución de los contratos (auditorías jurídicas y *legal due diligence*).

iii) Los análisis predictivos de diverso tipo y calado[5] referidos, por ejemplo, a las posibilidades de éxito de una demanda o al pronóstico sobre el modo en que un juez o tribunal resolverá determinado asunto, es decir, la jurimetría, entendida como la utilización de la inteligencia artificial con el fin de determinar la estrategia procesal más adecuada para asegurar el éxito en un caso o recabar una propuesta de resolución partiendo del análisis cognitivo de millones de decisiones judiciales (Solar, 2019: 130, matiza que, realmente, estos programas no predicen, si bien admite que tienen la capacidad para hacerlo; *vid.*, asimismo, De Hoyos, 2020: 25 ss.)[6].

Estos sistemas de búsqueda facilitan respuestas (información) basadas en la minería de datos —y de argumentos— que permiten al profesional del Derecho conocer qué decisiones han tomado los tribunales en situaciones similares. A través de este análisis «predictivo», el abogado tiene la posibilidad de planificar con precisión la estrategia procesal a seguir, pero también puede acceder a una información

valiosa que suministra parámetros para evaluar la conveniencia de ir a juicio y de pactar con la otra parte, así como del momento más adecuado para hacerlo[7].

iv) Los sistemas expertos que se aplican en los programas de cumplimiento normativo para identificar y clarificar en qué riesgos puede incurrir una empresa. Su finalidad es establecer los mecanismos de control, prevención y reacción frente a los mismos a efectos de excluir o reducir la responsabilidad penal de la persona jurídica.

v) La prevención de la corrupción a través del control de actividades, las alarmas antifraude y otras herramientas[8] (Martin, 2019: 539 y 552; Keats, 2008: 1288; De Hoyos, 2020: 17 ss.).

vi) El *eDiscovery* (o Descubrimiento electrónico) aplicado en los países pertenecientes a la tradición del *Common Law*[9]. Se trata de un tipo de investigación digital que ha evidenciado su utilidad y eficacia para gestionar el análisis de gran cantidad datos eventualmente probatorios que, en cumplimiento del principio de contradicción, deben ser comunicados a la parte contraria. Este sistema permite buscar, localizar, asegurar y seleccionar datos electrónicos con la intención de utilizarlos como prueba en el juicio[10].

vii) La elaboración automática de documentos jurídicos personalizados (contratos, testamentos, patentes, acuerdos de separación y divorcio, entre otros) a través de aplicaciones web que ajustan su contenido a las necesidades de los usuarios (*contract analytics*). La automatización de la redacción de estos documentos permite a numerosos usuarios de la Justicia acceder a ellos de forma más rápida, económica y sencilla[11] (Goodman, 2016: 91 ss.; Ruiz, 2019).

viii) La resolución de conflictos en línea (*Online Dispute Resolution*; *vid.*, entre otros, Montesinos, 2021: 507; Valbuena, 2015: 987).

ix) Los *chatbots* que brindan a los usuarios posibles soluciones a problemas jurídicos de diversa índole. Por ejemplo, DoNotPay (Reino Unido) o Aretha, que proporcionan asesoramiento legal gratuito[12].

Es evidente que estas aplicaciones han cambiado —y seguirán cambiando— el modo de actuar de los operadores jurídicos y los profesionales de la justicia (Susskind y Susskind, 2016: *passim;* Susskin, 1986; Solar, 2019: 18), así como la manera en que los ciudadanos nos relacionamos con la justicia: los Self-Help Centers, que prestan asistencia jurídica gratuita *online*, *ProBonoNet*, un sistema integrado de prestación de servicios legales, o la plataforma *W-Lex* ejemplifican claramente esta transformación (Ashley, 2018: 5; Goodman, 2016: 29).

3. INTELIGENCIA ARTIFICIAL Y EJERCICIO DE LA FUNCIÓN JURISDICCIONAL. LA ADOPCIÓN DE MEDIDAS CAUTELARES

3.1. CONSIDERACIONES PREVIAS

La aplicación más polémica de la inteligencia artificial en el campo del Derecho procesal —o, dicho de otro modo, la que más más dudas puede generar— es la que se refiere a su utilización por el órgano jurisdiccional en el ejercicio de la función que le es propia en régimen de exclusividad, esto es, la función jurisdiccional, pero también en el desarrollo del proceso penal.

Es inevitable plantearse las ventajas y los inconvenientes de esta posibilidad, que suscita inquietudes por los posibles menoscabos que puede generar en los derechos fundamentales y las garantías esenciales del proceso penal su utilización en el ámbito de la investigación de delito a través de herramientas que permiten la reconstrucción de los hechos a partir de vestigios o escenarios de casos anteriores, la identificación de sospechosos mediante reconocimiento facial, o incluso las autopsias virtuales, como Virtobot (Nieva, 2018: 26, quien cita los programas Stevie, Echo, Peirce-Igtt o Alibi; Adderley *et al.*, 2007: 312 ss.; Alcobeda, 2019; Bonet, 2020: 16 ss.; Barona, 2019*a*: 42; Estévez, 2019: 667 ss.).

También en materia probatoria podemos encontrar herramientas de inteligencia artificial que pueden ayudar al órgano jurisdiccional a valorar la prueba a partir de casos análogos o que, desde la psicología del testimonio, ofrecen al juez parámetros que le permiten valorar la credibilidad de los testigos.

Una característica común que se predica de estas herramientas es que, *a priori*, pretenden «eliminar» la subjetividad de quien deba tomar en consideración sus resultados, evitando así posibles errores derivados de la falibilidad de la memoria humana, por ejemplo, y —como ya hemos señalado— acceder a mayor cantidad de datos y «filtrarlos» de forma cada vez más precisa.

Pero la posibilidad que genera mayor inquietud es su utilización en la toma de decisiones judiciales, auxiliando o, según las hipótesis «más espectaculares y prometedoras» (Nieva Fenoll, 2018: 18; Belloso, 2022: 341; Velasco: 2022: 519)[13], sustituyendo al órgano jurisdiccional, de forma que la solución del caso sea adoptada por una máquina lo suficientemente sofisticada.

Una pregunta clave de este debate, a la que realmente no podemos dar respuesta en estas páginas, es cómo deciden los jueces, esto es, si toman sus decisiones de forma mecánica o automática o si, por el contrario, en ella entran en juego elementos no automatizables (Belloso, 2022: 347; Simón, 2021: 114). Con la prudencia necesaria para abordar estos temas, consideramos que, a la hora de tomar partido por una u otra posición (aceptar o no el uso de la IA en este contexto), cabe distinguir las decisiones intermedias, entre ellas decretar una medida cautelar personal o fijar una caución o fianza; y la decisión final, esto es, dictar la sentencia (Nieva Fenoll, 2018: 99; Kehl et al. 2017; Taruffo, 2018, 150), posibilidad, esta última, nos sitúa ante la viabilidad real de la figura del juez robot (Gómez

Colomer 2023*a*: *passim* y 2023*b*: 129; Llano, 2022: 234; Casanova, 2013: 45; De la Oliva, 2029: *passim*)[14].

3.2. JUSTICIA PREDICTIVA

En el campo de las decisiones judiciales, desde hace años asistimos a la progresiva implantación —particularmente en y desde Estados Unidos— del denominado enjuiciamiento predictivo o *predictive sentencing* (De Keisser *et al.*, 2019: 1, con referencia a la Bail Reform Act de 1984, que prevé la adopción de la medida cautelar más adecuada con base en la peligrosidad futura del sujeto, sistema avalado por la Sentencia *US vs. Salerno*; Van Ginneken, 2019: 175; Rose, 2002: 185 y 186; Andrés Pueyo y Echeburúa, 2010, 403 ss.). A través de un modo de aplicación de la inteligencia artificial que combina elementos estadísticos y factores criminógenos (Sánchez Rubio, 2018: 192 ss.), el enjuiciamiento predictivo determina el riesgo de que el sujeto cometa un delito y, en función de este riesgo, se impone la sentencia o, en lo que aquí interesa, se decreta una medida cautelar, si bien las denominadas *risk assessment tools* (herramientas de evaluación del riesgo) también permiten tomar otro tipo de decisiones, entre ellas la concesión de permisos a reos[15]. La finalidad de estas herramientas predictivas es limitar la arbitrariedad judicial, como lo hacen las conocidas *sentencing guidelines* (Israni, 2017). Lo cierto es que, sin entrar en otras objeciones, esta modalidad de adopción de decisiones supone la anteposición de la seguridad a cualquier otro valor jurídico, dado que toma en consideración no el delito cometido y las circunstancias en que se ha cometido, sino la peligrosidad o riesgo que el presunto autor supone para la sociedad (Rose, 2002:191; Romero 2018: 3 ss.; Cotino 2017: 135; Alcoceba, 2019; Miró 2018: 95 ss.; Martínez y Montes, 2018: 1; Cunningham et al., 2009: 223; De Keisser *et al.*, XXX:1), con los riesgos que ello supone para sus derechos y garantías procesales. A este factor de riego debe añadirse que el hecho de que se utilice la inteligencia artificial no convierte a a las decisiones tomadas en infalibles, ya que puede haber un alto porcentaje de «falsos positivos» y que las decisiones pueden estar basadas en sesgos relevantes (referidos, por ejemplo, a la raza, la condición social, el lugar de residencia y otros factores), una cuestión clave cuyo análisis excede en mucho el propósito de estas páginas.

El caso *State vs. Loomis*[16], en el que se cuestionó la aplicación del programa COMPAS[17], constituye un ejemplo de estos peligros (Pérez, 2017: 236 y 237; Thadaney, 2027; Miró, 2018: 122; Martínez, 2018: 448). Sin entrar en detalles, la sentencia *State vs. Loomis* consideró que el uso de la herramienta COMPAS, que se sirve de algoritmos para predecir, mediante *scoring* (tanteo), el riesgo de sujetos a enjuiciar no vulnera el derecho al debido proceso, pese a que el acusado alegó que no tuvo conocimiento de los elementos que el tribunal tomó en consideración para calificarlo como «peligroso» y, por tanto, no pudo contradecirlos. En este caso, primó la alegación por los creadores de COMPAS del riesgo que correría su derecho al secreto comercial (basándose en el derecho de patentes) si se desvelaba el funcionamiento del sistema sobre el derecho de defensa del acusado (Sánchez García, 2018: 49 ss.; Garza, 2018: 95; De Miguel y Pérez Estrada, 2019: 545; De Miguel, 2018: 4; Ortiz *et al.*, 2020:1; Freeman, 2016: 76).

Esta sentencia pone de manifiesto el denominado lado oscuro o «caja negra» de la inteligencia artificial, dado que el algoritmo no explica cómo determina una respuesta concreta. Si bien es cierto que podemos entender los datos que maneja y el resultado que ofrece, el problema estriba en que no se conoce el procedimiento a través del cual el algoritmo elige una solución específica (Corvalán, 2018: 306; Wachter *et al.*, 2017: 3), problema que se agrava por el hecho de que cada vez son más sofisticados. Ante las alegaciones que apelaban al derecho a la protección secretos comerciales o la confidencialidad, como las que fueron planteadas en el caso *Loomis*, parece más pertinente recurrir a la ponderación de intereses en juego y plantear si tienen más peso las ganancias económicas de la empresa que genera y gestiona el algoritmo o los derechos de la persona afectada por el mismo, pues en este tipo de casos no solo están en juego los derechos procesales, sino también la dignidad de la persona y la protección de datos personales (Boix, 2020: 239 ss.; Colomer: 2023*a*; Ariza: 2023: 28). Creemos que la respuesta sobre cuál de estos dos conjuntos de derechos debe pesar más es evidente, y así lo ponen de manifiesto los trabajos preparatorios y documentos que han precedido a la aprobación del Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial) de Inteligencia Artificial de la Unión Europea en mayo de 2024 (Ariza: 2023: 18).

3.3. APLICACIÓN DE LA INTELIGENCIA ARTIFICIAL EN LA ADOPCIÓN DE MEDIDAS CAUTELARES, EN PARTICULAR LA PRISIÓN PROVISIONAL

Como hemos adelantado, uno de los ámbitos del Derecho procesal —en este caso, del Derecho procesal penal— en el que puede utilizarse la inteligencia artificial desde la perspectiva predictiva es el de la adopción las medidas cautelares (Nieva Fenoll, 2018: 63; Planchadell Gargallo, 2021: 389; Guzmán, 2017: 110; Bonet, 2020: 9; Mora, 2021: 56; Delgado, 2020: 541; Simón, 2021: 114), campo en que debe atenderse a la concurrencia de unos presupuestos legalmente establecidos para decretar la medida cautelar más adecuada para «conjurar un riesgo». Precisamente por ello, la objetivación de dichos riesgos a través de algoritmos puede ser útil para el órgano jurisdiccional, dado que le permite tomar la decisión más ponderada y proporcional posible en la situación concreta sobre la que deba decidir.

Junto con el ya citado programa COMPAS, enfocado en la imposición de sanciones alternativas en la ejecución de las condenas, en Estados Unidos se utilizan también dos sistemas relacionados con las medidas cautelares y la ejecución: el sistema PATTERN, mediante el que se toman decisiones sobre la libertad condicional en el ámbito federal y Arnold Foundation's Public Safety Assessment, algoritmo utilizado para la adopción de decisiones relativas a las fianzas que parte de factores como la edad, el delito violento cometido y los cargos pendientes, los antecedentes y las posibles incomparecencias a juicios previos, entre otros (Salom, 2020: 440; Coglianesi y Lavi, 2020; González *et al.*, 2018: 29).

En España conocemos sobradamente el uso de herramientas de prevención del riesgo en el ámbito de la violencia de género a través del sistema VioGén (Salom, 2020: 441;

Salom, 2022: 394; López *et al.*, 2016:1; Montesinos: 2021: 19; Del Pozo, 2022: 159)[18], en funcionamiento desde julio de 2007. Este sistema, que ha ido perfeccionándose con los años, permite determinar el riesgo que corren las víctimas de violencia de género para adoptar la medida de protección más adecuada en cada caso. VioGén contiene 39 ítems relativos al ofensor y su interacción con la víctima, y, a modo de *check list*, va marcando los distintos ítems para arrojar como resultado el nivel de riesgo de la víctima (bajo, medio o alto), que puede ser modificado por el agente para ajustar la medida de protección concreta e individualizada[19]. Esta herramienta supone la interacción máquina-agente y permite «mejorar el trabajo diario del agente» (González *et al.*, 2028:36) a fin de brindar la protección integral a la que hace referencia la Ley 1/2004[20], pero no sustituir ni su valoración ni, desde luego, la del tribunal competente para decretar la medida que corresponda. Dada su utilidad, se ha considerado que cada vez es más factible extender su aplicación a otros ámbitos como la violencia doméstica o el acoso escolar.

Sin individualizar el análisis en un instrumento concreto, y volviendo a la posible aplicación general de estas herramientas a las medidas cautelares, tomaremos como referencia la previsión del art. 503 LECrim referida a la adopción de la prisión provisional, que, resulta preciso recordarlo, es la medida cautelar más gravosa prevista en nuestro ordenamiento jurídico, y en la que nos centramos en las páginas que siguen. Como es sobradamente conocido, la adopción de esta medida depende de la concurrencia de dos presupuestos, el *fumus boni iuris* y el *periculum in mora*, y su adopción se dirige a evitar que el encausado se sustraiga a la acción de la justicia (evitar el riesgo de fuga), que oculte, altere o destruya fuentes de prueba relevantes para su enjuiciamiento o que pueda poner en peligro la vida o patrimonio de la víctima. Esta configuración legal facilita que puedan incluirse en el algoritmo los distintos presupuestos legalmente previstos, de forma que, cuando los mismos concurran, la máquina «recomiende» la adopción de la medida cautelar personal más adecuada; de manera similar, podemos pensar en la adopción medidas cautelares reales —y en la fijación de una caución— atendiendo a la cuantía de la deuda, la existencia de impago previos, la situación patrimonial del sujeto, etc. Es, por tanto, a la hora de determinar esos posibles riesgos cuando la inteligencia artificial puede jugar un importante papel de auxilio al juez para que este los valore, dado que le permite establecer los ítems de los que derivaría el mayor o menor riesgo y, por lo tanto, la medida más adecuada a adoptar.

A título ilustrativo, y tras una selección jurisprudencial, podemos identificar el modo en que nuestros tribunales avalan recurrentemente la adopción de la prisión provisional atendiendo a la concurrencia de una serie de elementos, entre los que el más claro es el referido al peligro o riesgo de fuga. Concretamente, pueden enumerarse los siguientes:

- La gravedad del delito o de la pena prevista para su comisión (SSTC 136/2023, de 23 de octubre, 37/2020, de 25 de febrero o 22/2020, de 13 de febrero; Auto AP de Madrid (Sección 17.ª) núm. 504/2006, de 12 junio, entre otras muchas resoluciones).
- La comparecencia a llamamientos judiciales previos en la misma causa (STC 22/2020, de 13 de febrero).

– La existencia o no de arraigo en el país y los posibles vínculos con terceros países (STC 66/2008, de 29 de mayo; STS de 26 de febrero de 2024; Auto de la AP de Guipúzcoa (Sección 1.ª), núm. 42/2004, de 31 marzo; Auto de la AP de Asturias (Sección 7ª), núm. 67/2001, de 30 abril). El Auto de la AP de Barcelona núm. 218/2020, de 22 de mayo (ARP 2020\1278), declara que «[…] para apreciar el riesgo de fuga a conjurar la gravedad de las penas [podrían alcanzar los diez años o más, si atendemos a los tipos que allí se mencionan] solo tiene arraigo personal no laboral ni social acreditado, carece dice de residencia acreditada- argumento este que no compartimos pues no dudamos de que viviera viviendo donde fue detenido. Todo ello incrementa el riesgo de ilocalizacion, riesgo que ya fue valorado y confirmado por la sala de Vacaciones de la Audiencia en auto de 29 de agosto y en autos posteriores ya mencionados. Pues bien, para constatar el peligro de fuga hay que tener en cuenta conjuntamente la naturaleza, la gravedad de la pena y la situación laboral social y económica, de salud y el transcurso del tiempo».

– El tipo de imputación que pesa sobre el sujeto; a este respecto, la STC 66/2008 se pronuncia en los siguientes términos: «Se pondera el riesgo de fuga (basado en la naturaleza y gravedad de las imputaciones; la variedad de las mismas, pues el auto de procesamiento no solo le atribuye delitos contra la salud pública, sino también falsedades documentales y tenencia ilícita de armas, como destacó el Ministerio Fiscal en la comparecencia; el protagonismo o distinto papel respecto de otros imputados; que la actividad se desarrolla en el seno de una organización con conexiones internacionales; el elevado valor de la droga intervenida: 2 820 900 euros) con la posibilidad de que los encausados puedan gozar la libertad con la necesaria garantía que elimine o minimice su sustracción del proceso […]»; en sentido similar, SSTC 29/2019, de 28 de febrero, y 65/2008, de 29 de mayo).

– El protagonismo del sujeto o su papel en la comisión del delito en contraste con el de otros imputados (STC 65/2008, de 29 de mayo).

– La pertenencia a una organización con conexiones internacionales (SSTC 65/2008, de 29 de mayo; y 37/2020, de 25 febrero, en la que el Tribunal Constitucional afirma: «En relación a las circunstancias personales de los recurrentes, considera que "no puede dejar de valorarse la estructura organizativa que los apoya, ni la consistencia económica de la misma, ni los contactos internacionales del grupo, elementos que han permitido a quienes ya se han fugado, mantener una apariencia de vida normalizada fuera de España y de la residencia que hasta entonces era su domicilio habitual"». Y añade, al respecto, que «[…] aunque se mantenga la confianza que merecen los sistemas de cooperación judicial internacional instaurados en el ámbito de la Unión Europea […] la posibilidad de que los recurrentes consideren una opción atendible la huida de la acción de la Justicia española […] se incrementa, y debe ser evitada para asegurar la tramitación adecuada del proceso». Por último, y en referencia a relaciones familiares que se han puesto de manifiesto para acreditar un arraigo que excluye el riesgo de fuga, señala que «[…] aun siendo importantes, no suponen un impedimento definitivo a una fuga que vendría impulsada por otras

razones de gran peso, como las antes aludidas», dada «[…] la eficacia y la facilidad, incluso económica, que caracterizan las posibilidades de comunicación y de traslado físico de personas de unos a otros lugares de la Unión, y del resto de Europa»)[21].

- El valor de los bienes y efectos decomisados, en su caso (SSTC 65/2008, de 29 de mayo; 29/2019, de 28 de febrero).
- La situación familiar, laboral y económica del sujeto; por ejemplo, aun ante la falta de arraigo, la existencia de un contrato de trabajo (SSTC 128/1995, de 26 de julio; 29/2019, de 28 de febrero; Auto de la Audiencia Provincial de Madrid de (Sección 17.ª) núm. 504/2006, de 12 junio; Auto de la AP de Guipúzcoa (Sección 1.ª), núm. 42/2004 de 31 marzo; Auto de la AP de Asturias (Sección 7.ª), núm. 67/2001 de 30 abril).
- La situación procedimental de la causa (STC 36/2007, de 12 de febrero, que indica, como elementos a valorar, «[…] la naturaleza de los hechos, la gravedad de la pena que pueda imponerse a los procesados y la situación procedimental de la causa que hace presumir la pronta celebración del juicio, tal riesgo está latente, sin que dicho riesgo quede conjurado por las circunstancias familiares, personales o económicas de los procesados a las que aluden sus defensas […]»; en sentido similar, SSTC 150/2007, de 18 de junio; y 146/2001, de 18 de julio).
- El hecho de que la persona sometida al proceso tenga o no un medio de vida conocido (sentencia de la AP Barcelona núm.186/2020, de 8 de mayo).
- Si se trata de una persona que está enferma y necesita algún tipo de tratamiento o control médico en el país (Auto de la AP de Barcelona núm. 164/2020, de 21 de abril).
- La existencia o no de empadronamiento en el país (Auto de la Audiencia Provincial de Barcelona núm. 164/2020, de 21 de abril).
- Si se trata de una persona que tiene o no trabajo, domicilio o familia a su cargo (Auto de la Audiencia Provincial de Pontevedra, núm. 260/2009, de 22 de mayo).

Aunque de forma menos prolija, los elementos que permiten entender que concurre el peligro de reiteración delictiva se refieren, por ejemplo, a:

- La conducta concreta del imputado y su participación en el hecho delictivo (SSTC 37/2020, de 25 de febrero; 22/2016de 13 de febrero; y 23/2020, de 13 de febrero).
- Los datos objetivos que constan en la causa (STC 4/2020, de 25 de enero).
- La continuidad de las actuaciones (STC 29/2019, de 28 de febrero, o Auto de la AP de Asturias núm. 561, 2022, de 2 de noviembre).

Sobre la objetivación de estos elementos, la STC 36/2007, de 12 de febrero afirma que «[…] estos datos objetivos que pueden ser tenidos en cuenta inicialmente no pueden operar como único criterio —de aplicación objetiva y mecánica— sin ser puestos en relación con otros datos relativos a las circunstancias personales del procesado —el arraigo familiar, profesional o social, las conexiones en otros países, los medios económicos de que dispo-

ne…— una vez que el tiempo transcurre (SSTC 128/1995 [RTC 1995, 128]; 47/2000 [RTC 2000, 47]; 23/2002 [RTC 2002, 23]). En este sentido subraya la demanda que las SSTC 146/1997 [RTC 1997, 146] y 66/1997 [RTC 1997, 66] consideraron que ni la gravedad de la pena ni la inminencia para dictar sentencia son datos suficientes justificativos de la prisión provisional, pues el último es ambivalente y, en todo caso, sería necesario individualizar las circunstancias personales del imputado para no caer en una sospecha de riesgo genérico de fuga». La objetivación es también destacada por las SSTC 146/2001, de 18 de julio; 22/2020, de 13 de febrero; y 23/2020, de la misma fecha. En relación con la reiteración delictiva, la STC 4/2020, de 15 de enero, afirma: «Al igual que hemos señalado en la STC 155/2019, de 28 de noviembre [RTC 2019, 155], FJ 13, hemos de reiterar ahora que tal apreciación no solo se funda en datos objetivos, sino que lo determinante del peligro de reiteración delictiva que se aprecia no es que el recurrente mantenga su aspiración de que Cataluña alcance la independencia, sino la perseverancia en el modo ilegal con el que ese designio se pretende lograr, con las consecuencias que ello comportaría para el orden constitucional».

Así, junto con la necesaria concurrencia de la apariencia de buen derecho, es decir, la «[…] existencia de indicios atendibles de perpetración culpable de un delito sancionado» en el sentido a que se refiere el art. 503 LECrim», como dice la STC 37/2020, de 25 de febrero, que exige la ineludible valoración judicial, «[…] debe recurrirse a una serie de elementos externos que permitan construir un juicio razonable de pronóstico, y no solo respecto de su voluntad presente, sino de la eventualidad de que esta pueda modificarse con ocasión del propio desarrollo de la causa»; precisamente, la construcción de esos elementos externos de carácter objetivo es la tarea para la que el uso de los algoritmos podría ser útil (Simón 2021:120 y 149; Planchadell Gargallo, 2021: 289); los algoritmos auxilian —pero no sustituyen— al juez en la toma de la decisión facilitándole el mayor número de datos posibles que le permiten constatar si concurre o no un riesgo real y adoptar la medida más adecuada en función del tipo de peligro que debe ser conjurado.

4. REFLEXIÓN FINAL

Como hemos indicado en las páginas precedentes, uno de los mayores retos a los que se enfrentan los profesionales de la Administración de Justicia en general es la irrupción de la inteligencia artificial. Las herramientas de inteligencia artificial ofrecen un amplio abanico de aplicaciones, pero al utilizarlas debemos ser conscientes de los límites marcados por el respeto al debido proceso y a los derechos y garantías de los justiciables.

Dada la evidente utilidad de esta tecnología, debemos plantearnos si estamos dispuestos a aceptar que las herramientas de IA sustituyan al razonamiento judicial en la toma de decisiones. Esta riesgosa posibilidad es una realidad desde la perspectiva de la justicia predictiva, y uno de sus ámbitos de aplicación privilegiados es la adopción de las decisiones sobre las medidas cautelares que deben ser adoptadas. La objetivación de los presupuestos que llevan adoptar las medidas cautelares permite su sistematización a través de la inteli-

gencia artificial, de forma que, combinando algoritmos y atendiendo, entre otros factores, a criterios estadísticos y criminológicos, proporcione al juez un pronóstico relativo a la posibilidad de que un investigado se fugue o no pague una multa y, con base al mismo, el operador dicte una u otra medida cautelar si así fuera recomendado. Siendo esta una posibilidad factible, lo que debemos cuestionarnos es si la aceptamos y, de no ser evitable, con qué cautelas y garantías lo hacemos.

NOTAS

1. Son muchas las noticias que encontramos en prensa al respecto. Entre otras: https://confilegal.com/20191013-china-y-estonia-desarrollan-jueces-virtuales-basados-en-inteligencia-artificial-para-resolver-demandas-de-cantidad/; https://www.elperiodico.com/es/economia/20191019/inteligencia-artificial-microsoft-jordi-ribas-7689870; https://www.diariodesevilla.es/juzgado_de_guardia/actualidad/drones-criminales-guerra-jueces-futuro-robots_0_1835818203.html (acceso junio 2025).

Lo que Isaac Asimov vaticinó en *Yo robot* (1950) no es tan ficcional como pensamos cuando leímos por primera vez la novela.

2. Bruselas 19.2.2020 COM (2020) 65 – final. Disponible en: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf.

3. Para poner de manifiesto que el problema no es la máquina sino el humano, el autor explica las lecturas que se realizan el primer día de clase de la asignatura Law and Robots que se imparte en la Universidad de Yale. Por ejemplo, la leyenda del Rabino y el Gólem, así como la falacia del homúnculo.

4. Pensemos en las cada vez más sofisticadas bases de datos que utilizamos en nuestra vida cotidiana. Lexis fue diseñada en 1973 y Westlaw en 1975, y, desde entonces, no han dejado de evolucionar y mejorar tanto en la cantidad de información que contienen como en su organización y forma de acceder a ella. Se ha llegado a afirmar que, hoy en día, la maquina Watson —que ganó el concurso Jeopardy en 2011— aprobaría el examen de acceso a la abogacía en la mayoría de los estados de EE. UU.

5. Como afirmó Holmes en su obra *La senda del Derecho*, «[…] las profecías acerca de lo que los tribunales harán realmente, y mucho más pretencioso que eso, es lo que yo entiendo por Derecho». Es decir, gran parte del trabajo de un abogado consiste en preguntarse qué hará el juez en el caso que el letrado le presente y según cómo lo presente: la clave es el análisis de las decisiones previas sobre la materia (Holmes, 2012: 60).

6. Por ejemplo, el programa *LexMachina* diseñado en la Universidad de Stanford y que ha sido adquirido por *Lexis Nexis*, https://lexmachina.com/media/press/lexisnexis-acquires-lex-machina/. La editorial Wolters & Kluwer ofrece ya a los despachos de abogados el producto Jurimetría, al que define como la «[…] herramienta de analítica jurisprudencial más innovadora del mercado que permite definir la estrategia procesal más idónea para el éxito del caso, a través de indicadores gráficos interactivos, basados en el análisis cognitivo de millones de decisiones judiciales». Disponible en: https://jurimetria.wolterskluwer.es/content/Inicio.aspx. En este enlace puede verse un vídeo muy ilustrativo, distinguiendo entre jurimetría del caso, del magistrado, del abogado, de la empresa, del tribunal y del organismo público.

7. Existen, en este sentido, dos experiencias. Una tuvo lugar en el Tribunal Supremo de Estados Unidos y otra en el Tribunal Europeo de Derechos Humanos (Alatras *et al.*, 2016) A un nivel más modes-

to, la experiencia también se ha llevado a cabo en tribunales de Australia y Nueva Zelanda o India, entre otros países (Simester y Brodie, 1993: 49).

8. Por ejemplo, la Ley 22/2018, de 6 de noviembre, de la Generalitat, de Inspección General de Servicios y del sistema de alertas para la prevención de malas prácticas en la Administración de la Generalitat y su sector público instrumental. Disponible en: http://www.transparencia.gva.es/es/anteproyecto-de-ley-de-la-generalitat-de-inspeccion-general-de-servicios-y-del-sistema-de-alertas-para-la-prevencion-de-malas-practicas-en-la-administracion-de-la-generalitat-y-su-sector-publico-instrumental [Consulta: 29/12/ 2020.]

9. La sentencia *Estados Unidos Da Silva Moore vs. Predius Group & MSL Group* en 2012 ya admitió la codificación de la información como un medio aceptable de *e.Discovery* basándose en el valor del precedente en dicho país y en los llamados Sedona Principles.

10. Aplicaciones como ROSS, que combina el *machine learning* y el procesamiento de lenguaje natural, cambian completamente el modo en que trabajan los grandes despachos de abogados en Estados Unidos (https://rossintelligence.com) (Goodman, 2016: 41 ss.; Solar, 2018:77 ss.; Lawton *et al*., 2014).

11. El Center for Access to Justice & Technology del Chicago Kent College of Law y el Center for Computer-Assisted Legal instruction de la misma Universidad son ejemplos de cómo puede desarrollarse esta función.

12. Disponible en: https://elderecho.com/aretha-chatbot-inteligencia-artificial-asesoramiento-legal-gratuito-frente-al-covid-19-ontier-lefebvre.

13. Nieva Fenoll (2018) hace referencia a «IBM's Watson. What's next Debating», video de youtube en que se nos muestra cómo, una vez planteado un tema de debate a la máquina, en muy poco tiempo elabora una lista de argumentos a favor o en contra con soporte documental. Disponible en: https://www.youtube.com/watch?v=sEf0GLvrP9U, junto con otros similares de IBM [Consulta: 29/12/2020.]

14. China y Estonia han desarrollado también «jueces virtuales» para resolver demandas de cantidad. En China se ha creado el Tribunal de Internet de Pekín, centro en línea en que se resuelven litigios simples, si bien en una primera fase esta corte actuará apoyando a los jueces. En Estonia, se prevé que, jueces robots resuelvan reclamaciones de cantidad de hasta 7000 euros si bien parece que el proyecto ha sido "suspendido", y para asuntos en los que para su solución únicamente deba aplicarse normas o cláusulas contractuales, por lo que basta subir a la plataforma la documentación a utilizar al efecto. La resolución de la IA sí se recurrirá ante un juez «humano». El Consejo General del Poder Judicial español ha manifestado su intención de desarrollar una inteligencia artificial propia que permita predecir el resultado del procedimiento.

15. Además de Compas, pueden citarse Savry, para jóvenes, SVR-20, en el contexto de la violencia sexual, Spousal Assault Risk Assessment (Sara), HCR-20, para la violencia, o Hart, entre otras.

16. *State vs. Loomis* 881 N.W. 2d 749 (Wiss, 2016), donde se reconocen las limitaciones del sistema y el hecho de que COMPAS no pueda ser utilizado para determinar si el sujeto debe ser encarcelado, o para tomar en consideración agravantes o atenuantes, aplicándose tanto a presos preventivos como definitivos. Sobre el particular, *vid.* Harvard Law Review (2017).

17. Lo cierto es que desde hace algunos años un número importante de estados —por ejemplo, Oregon, Oklahoma, Texas, California, Nebraska, Colorado, Indiana, Ohio, Delaware, Washington, Arizona, Virginia o Alabama— utilizan herramientas de predicción; COMPAS es la más conocida. Mediante su uso, se pretende evaluar el presente y pasado del sujeto, comparándolo con un grupo concreto de origen y características similares. Los datos que permiten la comparación se obtienen del historial del sujeto, partiendo de un documento detallado de sus antecedentes y de un cuestionario de

137 preguntas que se le realizan, cuyas preguntas versan sobre si alguno de los progenitores ha estado en prisión, si alguno de sus amigos toma drogas, etc. Nieva Fenoll (2018: 68) califica de «variopintos» los datos que este programa toma en consideración; además del número de detenciones previas o la pertenencia del sujeto a una organización criminal, también se toman en consideración el drogas y alcohol, los antecedentes penales de su familia y entorno o la situación domiciliara (por ejemplo, si vive en una casa de su propiedad o no, la comparte o similares), entre muchos otros.

18. La información oficial está disponible en: http://www.interior.gob.es/web/servicios-al-ciudadano/violencia-contra-la-mujer/sistema-viogen.

19. Consultadas la Guías de procedimiento VPR5.0 y VPER4.1, encontramos, por ejemplo, que los elementos que se toman en consideración son los siguientes: si en los últimos seis meses el agresor muestra celos exagerados o sospechas de infidelidad; si realiza conductas de control o de acoso; si ha agredido a personas o animales o si presenta algún tipo de adicción. A la hora de valorar la vulnerabilidad de la víctima se tienen en cuenta, entre otros, estos por ejemplo a si tiene algún tipo de discapacidad, si ha presentado alguna idea de suicidio, si cuenta con apoyo familiar o social.

20. A estas consideraciones o criterios para adoptar la orden de protección se refieren, entre otros, los Autos de la AP de Barcelona (Sección 20.ª), núm. 822/2008 de 13 agosto, de la AP de Madrid (Sección 27.ª), núm. 55/2020 de 13 enero; de la AP de Madrid (Sección 26.ª) Auto núm. 198/2008, de 6 febrero; y de la AP de Madrid (Sección 27.ª), núm. 1558/2020, de 29 octubre.

21. La STS de 25 de enero de 2019 hace expresa referencia a la doctrina del Tribunal Europeo de Derechos Humanos dictada en el caso *Demirtas c. Turquía*, alegada por la defensa de los Sres. Santos, Pablo y Ramón como modelo de inspiración para la legitimidad democrática de una medida cautelar que ofrece un elemento de especial valor interpretativo. En efecto, esta resolución incluye entre los factores que el juez tener en cuenta para evaluar el riesgo de fuga de las personas privadas preventivamente de libertad antes del juicio, «los contactos internacionales». Así consta en el parágrafo 186 de la sentencia citada, con referencia, a su vez, a otras resoluciones dictadas por el TEDH que evocan una doctrina plenamente consolidada, entre ellas, la sentencia de fecha 26 de enero de 1993, recaída en el caso *W. c. Switzerland*. En relación con el riesgo de fuga, allí puede leerse que el Tribunal Europeo ha mantenido que ese riesgo ha de ser valorado atendiendo también a otros factores tales como la personalidad del acusado, sus bienes, su proximidad a la jurisdicción y los contactos internacionales (en la version inglesa oficial: «[…] as regards the risk of flight, the Court has held that it must be assessed with reference to a number of other factors, such as the accuseds character, morals, assets, links with the jurisdiction and international contacts»).

BIBLIOGRAFÍA

ADDERLEY, Richard, Jonh W. BOND y Michael K. TOWNSLEY (2007): «Predicting Crime Scene Attendance», *International Journal of Police Science & Management*, 9(4), 312-323.

ALCOBEDA GIL, J.M. (2019): «Potencialidades y riesgos del Big Data, la minería de datos y *machine learning* como herramienta predictiva de la criminalidad», *Comunicación presentada en el II Congreso de la Asociación de Profesores de Derecho Procesal de las Universidades españolas*, Valencia.

ALETRAS, Nikolaos, Dimitrios TSARAPATSANIS y Daniel PREOȚIUC-PIETRO (2016): «Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective», *PeerJ Computer science*, 2.

ANDRÉS-PUEYO, Antonio y Enrique ECHEBURÚA (2010): «Valoración del riesgo de violencia: Instrumentos disponibles e indicaciones de aplicación», *Psicothema*, 22(3), 403-409.

ARIZA COLMENAREJO, Maria Jesús (2023) «Fuentes de datos al servicio de sistemas de inteligencia artificial dirigidas a la toma de decisiones», en R. Castillejo Manzanares y M.L Noya Ferreiro (dirs.), *Inteligencia artificial y proceso penal: Un reto para la Justicia*, Cizur menor: Aranzadi.
ASHLEY, Kevin D. ([3]2018): *Artifical Intelligence and Legal Analytics. New tools for Law practice in the Digital Age*, Cambridge: Cambridge Univesity Press.
BALIKIN, Jack M. (2017): «The three laws of robotics in the age of Big Data», *Ohio State Law Journal*, 78(5).
BARONA VILAR, Silvia (2018): «El título VI de la Constitución "Poder Judicial", desde una mirada en el siglo XXI», *Corts: Anuario de Derecho parlamentario*, n.º extra, 31, 475-496.
— (2019): «Cuarta revolución industrial (4.0) o ciberindustria en el proceso penal: Revolución digital, Inteligencia Artificial y el camino hacia la globalización de la justicia», *Revista Jurídica Digital UNIANDES*, 1(3), 1-17.
— (2019b): «Inteligencia artificial o la algoritmización de la vida y de la Justicia: ¿Solución o problema?», *Revista Boliviana de Derecho*, 28, 18-49.
BARRIO ANDRÉS, Moisés (2018): «Robótica, Inteligencia Artificial y Derecho», *ARI* 103.
BELLOSO MARTÍN, Nuria (2013): «Entre la ciencia y la técnica del derecho: ¿Hacia una hermenéutica telemática?» *Anales de la Cátedra Francisco Suárez*, 47, 139-161.
— (2022): «Inteligencia artificial en la teoría de la decisión judicial ¿Del juez robot al asistente robot del juez?», en A. Peralta, L. Torre, y F. Herrera (coords.), *El derecho y la inteligencia artificial*, Granada: Universidad de Granada.
BOIX PALOP, Andrés (2020): «Los algoritmos son reglamentos: La necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la Administración para la adopción de decisiones», *Revista de Derecho Público: Teoría y método*, 1(1), 223-269.
BONET NAVARRO, José (2018): «La tutela judicial de los derechos no humanos. De la tramitación electrónica al proceso con robots autónomos», *Revista CEFLEGAL*, 208, 55-92.
— (2020): «Algunas reflexiones sobre la viabilidad de la inteligencia artificial en el proceso penal», *Revista Inmexius*, 40, 16-24.
BUENO MATA, Federico (2020): «Macrodatos, inteligencia artificial y proceso: luces y sombras», *Revista General de Derecho Procesal*, 51.
BUJOSA VADELL, Lorenzo M. (2014): «De iudicio. Variaciones en torno a la potestad y acto de juzgar», *Ars Iuris Samanticensis*, 2(1), 41-71.
CÁCERES, Enrique (2006): «Inteligencia Artificial, Derecho y E-Justicia (El proyecto Conacyt)», *Boletín Mexicano de Derecho comparado*, 116, 593-611.
CASABONA ROMERO, Carlos (2020): «Inteligencia artificial, proceso y derechos fundamentales», *Comunicaciones en propiedad industrial y derecho de la competencia*, 89, 253-271.
CASANOVA, Pompeu (2010): «Inteligencia Artificial y Derecho. A vuelapluma», *Teoría y derecho: revista de pensamiento jurídico*,7, 203-222.
— (2013): «Tecnología, Inteligencia Artificaial y Web semántica. Un mundo para todos y cada uno» en J.l Fabra Zamora y A. Núñez Vaquero (coords.), *Manual de Filosofía y Teoría del Derecho*, vol. 1, México: UNAM.
CERRILLO MARTÍNEZ, Agustí y Carlos VELASCO RICO (2019): «Jurisdicción, algoritmos e inteligencia artificial» en F. López Ramón y J. Valero Torrijos (coords.), *20 años de la Ley de los Contencioso-administrativo*: *actas del XIV Congreso de la Asociación Española de Profesores de Derecho Administrativo*, Murcia: INAP, 291-303.
COGLIANESE, Cary y Lavi M. BEN DOR (2021):«AI and adjudication and Administrations», *All Faculty Scholarship*, 86(3), 791-838.
CONDE FUENTE, Jesús (2023): «La irrupción de la Inteligencia Artifical en el ámbito jurisdiccional: especial referencia a la toma de decisiones judiciales», en R. Castillejo Manzanares y M.L Noya Ferreiro (dirs.) *Inteligencia artificial y proceso penal: Un reto para la Justicia*, Cizur menor: Aranzadi.

CORVALÁN, Juan Gustavo (2018): «Inteligencia Artificial. Retos, desafíos y aportaciones – Prometea: La primera Inteligencia Artificial latinoamericana al servicio de la Justicia», *Revista de Investigaçoes Constitucionais*, 5(1), 295-316.

COTINO HUESO, Lorenzo (2017): «Big data e Inteligencia Artificial. Una aproximación a su tratamiento jurídico desde los Derechos Fundamentales», *Dilema*, 24, 131-150.

CUNNINGHAM, Mark D., Jon R. SORENSEN y Thomas J. REIDY (2009): «Capital Jury decision making. The limitations of predictions of future violence», *Psychology, Public Policy and Law*, 15(4), 223-256.

DE HOYOS SANCHO, Monserrat (2020): «El libro Blanco sobre la inteligencia artificial de la Comisión Europea: Reflexiones desde las garantías del proceso penal como sector de riesgo», *Revista Española de Derecho Europeo*,76, 9-43.

DE KEISSER, Jan, Julian V. ROBERTS y Jesper RYBERG (2019): *Predictive Sentencing. Normative and empirical perspectives*, Oxford: Hart Publishing.

DE LA OLIVA SANTOS, Andrés (2019): «Justicia predictiva. Interpretación matemática de las normas, sentencias robóticas y la vieja historia del *Justizklavier*», *El cronista del Estado Social y Democrático de Derecho*, 80, 30-37.

DE MIGUEL BERIAIN, Íñigo (2019): «¿Does the use of risk assessment in sentences respect the right to due process? A critical analysis of the Wisconsin v. Loomis Ruling», *Law, Probability and Risk*, 17.

DE MIGUEL BERIAIN, Íñigo y Josune PÉREZ ESTRADA (2019): «La inteligencia artificial en el proceso penal español. Un análisis de su admisibilidad sobre la base de los derechos fundamentales implicados», *Revista de Derecho UNED*, 25, 531-561.

DEL POZO PÉREZ, Marta (2022): «¿Los algoritmos sirven para proteger a las víctimas de violencia de género?», en F. Bueno Mata (dir.), *El impacto de las tecnologías disruptivas en el derecho procesal*, Cizur Menor: Aranzadi, 155-181.

DELGADO MARTÍN, Joaquin (2020): *Judicial-Tech, el proceso digital y la transformación tecnológica de la Justicia. Obtención, tratamiento y protección de datos en Justicia*, Madrid: Wolters Kluwer.

ESTÉVEZ MENDOZA, Lucana (2019): «Algoritmos policiales basadas en IA y derechos fundamentales a la luz de HART y VALCRI»: Garantía versus eficacia», en F. Jiménez Conde y R. Bellido Penadés (dirs.), *Justicia: ¿Garantía versus eficiencia?*, Valencia: Tirant lo Blanch, 665-675.

FERNÁNDEZ HERNÁNDEZ, Carlos y Pierre BOULAT (2019): «Inteligencia Artificial y Derecho. Problemas y perspectivas», *Noticias Jurídicas* [en línea], <https://noticias.juridicas.com/conocimiento/articulos-doctrinales/9441-inteligencia-artificial-y-derecho-problemas-y-perspectivas/>. [Consulta: 4/03/2024.]

FERNÁNDEZ, Carlos B. (2020): «Primera sentencia europea que declara ilegal un algoritmo de evaluación de características personales de los ciudadanos», *Diario La Ley* [en línea], <https://diariolaley.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAAEAMtMSbH1czUwMDAyNDa3NDJUK0stKs7Mz7M1MjACC6rl5aekhrg425bmpaSmZealpoCUZKZVuuQnh1QWpNqmJeYUp6qlJuXnZ6OYFA8zAQCfSdkrYwAAAA==WKE>. [Consulta: 18/03/2024.]

FREEMAN, Katherine (2016): «Algorithmic injustice: How the Wisconsin Supreme Court failed to protect due process rights in State v. Loomis», *North Carolina Journal of Law & Technology*, 75.

GARZA BARBOSA, Roberto (2018): «Nanotecnología, derecho y propiedad intelectual. La evolución de una sinergia poco probable», *Informática y Derecho. Revista Iberoamericana de Derecho Informático*, 5, 95-105.

GÓMEZ COLOMER, Juan Luis (2023*a*): *El juez robot. La independencia judicial en peligro*, Valencia: Tirant lo Blanch.

— (2023*b*): «Algunas consideraciones de tipo orgánico sobre la posibilidad de que un día nos juzgue una máquina (¿Serán necesarias normas sobre la organización judicial?)», en R. Castillejo Manzanares y M. L. Noya Ferreiro (dirs.), *Inteligencia artificial y proceso penal: Un reto para la Justicia*, Cizur menor: Aranzadi.

GONZÁLEZ ÁLVAREZ, José Luis, Juan José LÓPEZ OSSORIO, Carlos URRUELA CORTÉS y Marina RODRÍGUEZ DÍAZ (2018): «Integral monitoring system in cases of Gender Violence VioGén system», *Behaviour & Law Journal*, 4(1), 29-40.
GOODMAN, Joanna (2016): *Robots in Law. How artificial Intelligence is transforming legal services?,* London: ARK Group.
GUZMÁN FLUJA, Vicente Carlos (2017): «Sobre la aplicación de la Inteligencia Artificial a la solución de conflictos (Reflexiones acerca de una transformación tan apasionante como compleja», en S. Barona Vilar (coord.), *Justicia civil y penal en la era global*, Valencia: Tirant lo Blanch, 67-122.
HOLMES, Oliver Wendell (2012): *La senda del Derecho*, Madrid: Marcial Pons.
ISRANI, Ellora (2017): «Algorithmic Due Process: Mistaken Accountability and Atribution in *State vs. Loomis*», *Jolt Digest* [en línea] <https://jolt.law.harvard.edu/digest/algorithmic-due-process-mistaken-accountability-and-attribution-in-state-v-loomis-1>. [Consulta: 23/01/2024.]
KEATS CITRON, Danielle (2008): «Technological Due Process», *Washington University Law Review*, 85, 1249-1313.
KEHL, Danielle, Priscilla GUO y Samuel KESSLER (2017):«Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing», *Responsive Communities Initiative*, [en línea] <https://dash.harvard.edu/handle/1/33746041>. [Consulta: 15/02/2024].
LAWTON, D. R. STACEY y G. DODD (2014): E-discovery in digital forensic investigation», *CAST Publication*, 32/14, Reino Unido: Home Office.
LÓPEZ OSORIO Juan J., José GONZÁLEZ-ALVAREZ y Antonio ANDRÉS-PUEYO (2016): «Eficacia predictiva de la valoración policial del riesgo de la violencia de género», *Psychological Intervention*, 25, 1-7.
LLANO ALONSO, Fernando Higinio (2023): «Justicia digital, algoritmos y derecho: De la predictibilidad del big data al mito del juez robot», en J. I Solar Cayón y M.O Sánchez Martínez (coords.), *El impacto de la inteligencia artificial en la teoría y la práctica jurídica*, Madrid: La Ley, 219-244.
MAGRO SERVET, Vicente (2018): «La aplicación de la Inteligencia Artificial en la Administración de Justicia», *La Ley Digital*, 9268.
MARTIN DIZ, Fernando (2009): «Aplicación de IA en procesos penales relacionados con la corrupción», en N. Rodriguez García, A. Carrizo González-Castell y F. Rodríguez-López (eds.), *Corrupción: compliance, represión y recuperación de activos*, Valencia: Tirant lo Blanch.
MARTÍNEZ GARAY, Lucía (2018): «Peligrosidad, algoritmos y due process: El caso State v. Loomis», *Revista de Derecho Penal y Criminología*, 20, 482-502.
MARTINEZ GARAY, Lucía y Francisco MONTES SUAY (2024): «El uso de valoraciones del riesgo de violencia en el Derecho Penal: Algunas cautelas necesarias», *InDret*, 2, 2-47.
MIRÓ LINARES, Fernando (2018): «Inteligencia artificial y justicia penal: Más allá de los resultados lesivos causado por robots», *Revista de Derecho Penal y Criminología*, 20, 87-130.
MONTESINOS GARCÍA, Ana (2021*a*): «Los algoritmos que valoran el riesgo de reincidencia. En especial, el sistema Viogen», *Revista de Derecho y Proceso Penal*, 2021, 64, 19-55.
— (2021*b*): «Inteligencia artificial y ODR», en S. Barona Vilar (ed.), *Justicia algorítmica y neuroderecho: una mirada multidisciplinar*, Valencia: Tirant lo Blanch, 507-531.
MORA SÁNCHEZ, Jefry José (2021): «Predictibilidad conductual y proceso penal: Algunos apuntes sobre el fundamento epistémico de los hechos futuros en las medidas cautelares», *Quaestiofacti*, 2, 53-86.
NIEVA FENOLL, Jordi (2018): *Inteligencia Artificial y proceso judicial*, Barcelona: Marcial Pons.
ORTIZ HERNÁNDEZ, Susana, Inma GARRÓS FONT y M.ª Nuria ROMERA SANTIAGO (2020): «Hacia una implantación de la inteligencia artificial en nuestro sistema judicial», *Revista Aranzadi Doctrinal*, 3.
PÉREZ ESTRADA, Miren Josune (2017): «El uso de algoritmos en el proceso penal y el derecho a un proceso con todas las garantías», en S. Barona Vilar (dir.), *Claves de la Justicia Penal. Feminización, Inteligencia Artificial, Supranacionalidad y Seguridad*, Valencia: Tirant lo Blanch, 253-254.

PLANCHADELL GARGALLO, Andrea (2021): «Inteligencia artificial y medidas cautelares», en S. Barona Vilar (ed.), *Justicia algorítmica y neuroderecho: una mirada multidisciplinar*, Valencia: Tirant lo Blanch, 389-419.

PONCE SOLÉ, Juli (2018): «La prevención de riesgo de mala administración y corrupción, la Inteligencia Artificial y el derecho a una buena administración», *Revista Internacional Transparencia e Integridad*, 6, 1-19.

RIOS LÓPEZ, YOLANDA (2019): «La tutela del consumidor en la contratación inteligente. Los "smart contracts" y los "blockchain" como paradigma de la Cuarta Revolución industrial», *Revista de Consumo y empresa*, 9, 1-12.

RISSLAND Edwina L., Kevin ASHLEY y Ronald LOUI, (2003): «AI and Law: A fruitful sinergy», *Artificial Intelligence*,150, 1-15.

ROMERO CASABONA, Carlos María (2018): «Riesgo, procedimientos actuales basados en Inteligencia Artificial y medidas de seguridad», *Revista de Derecho, Empresa y Sociedad (REDS),* 13, 39-55.

ROSE, William (2002): «Crimes of colour, Profiling and the contemporary racializacion of social control», *International Journal of Politics, Culture and Society*, 16(2), 179-205.

SALOM LUCAS, Amparo (2020): «Artificial Intelligence and the future of human rights», en I. Florek, A. Koronziová y J.L Zamora Manzano (eds.), *Crisis as a challenge for human rights*, Bratislava: Publishers Comenius University in Bratislava.

— (2022): «Inteligencia artificial y proceso», en A. Peralta, F. Herrera Triguero, A. Peralta Gutiérrez y L. Salvador Torres López (coords.), *El derecho y la inteligencia artificial*, Granada: Ed. Universidad de Granada.

SÁNCHEZ GARCÍA, Luz (2018): «Las invenciones generadas con Inteligencia Artificial y sus implicaciones para el derecho de patentes», *Informática y Derecho. Revista Iberoamericana de Derecho Informático*, 5, 49-84.

SÁNCHEZ RUBIO, Ana (2018): «Los peligros de la probabilidad y la estadística como herramienta para la valoración jurídico-probatoria», *Rev. Brasileira de Direitto Processual Penal*, 4(1), 183-214.

SCHWAB, Klaus (2016): *La cuarta revolución industrial*, Barcelona: Penguin Random House.

SIMESTER, Ducan I. y Roderick J. BRODIE (1993): «Forecasting criminal sentencing decisions», *International Journal of Foreca*sting, 9, 49-60.

SOLAR CAYÓN, José Ignacio (2018): «La codificación predictiva: Inteligencia Artificial en la averiguación procesal de los hechos relevantes», *Anuario de la Facultad de Derecho de Alcalá*, 11, 75-105.

— (2019): *La Inteligencia Artificial jurídica. El impacto de la innovación tecnología en la práctica del Derecho y el mercado de servicios*, Cizur Menor: Aranzadi.

SUSSKIND, Richard y Daniel SUSSKIND (2016): *El futuro de las profesiones. Cómo la tecnología transformará el trabajo de los expertos humanos*, Zaragoza: Teell.

SUSSKIND, Richard.E. (1986): «Expert systems in Law: A jurisprudential approach to Artificial Intelligence and legal reasoning», *The Modern Law Review*, 49(2), 168-194.

TARUFFO Michele (1998): «Judicial decissions and Artificial Intelligence», *Artifical Intelligence and law*, 6, 311-314.

— (2018): «Statistics notes about statistical evidence», AAVV, *O processo civil entre a tecnica processual e a tutela dos direitos. Estudios Homenagem a Luiz Guilherme Marinoni*, Ed. Thomson Reuters, São Paulo 2018.

VALBUENA GONZÁLEZ, Felix (2015): «La plataforma europea de resolución de litigios en línea (ODR) en materia de consumo», *Revista de Derecho Comunitario Europeo*, 52, 986-1016.

VAN GINNEKEN, E.F.J.C (2019): «The use of risk assessment in Sentencing», en J.V. Roberts y J. Ryberg (eds.), *Predictive Sentencing. Normative and empirical perspectives*, Oxford: Hart Publishing.

VEGA IRACEHY, Jorge J. (2018): «Inteligencia Artificial y Derecho: Principios y propuestas para una gobernanza eficaz», *Informática y Derecho. Revista Iberoamericana de Derecho Informático*, 5, 13-18.

VELASCO NUÑEZ, Eloy (2022): «Inteligencia artificial: Aspectos penales y procesales», en F. Herrera, A. Peralta y L. Torres (coords.), *El derecho y la inteligencia artificial*, Granada: Universidad de Granada.

WACHTER, Sandra, Brent MITTELSTADT y Luciano FLORIDI (2017): «Why a right to explanation of automated decision-making does not exist in the general data protection regulation», *International Data Privacy Law*, 0, 1-24.

Fecha de recepción: 15 de julio de 2024.
Fecha de aceptación: 20 de octubre de 2024.

ESTUDIOS

PASIÓN DE LOS FUERTES.
EL DERECHO COMO TRATADO DE PASIONES

PASSION OF THE STRONG.
LAW AS A TREATY OF PASSIONS

Javier de Lucas

Catedrático de Filosofía del Derecho
Profesor emérito
Universitat de València

RESUMEN

Tanto la función social de ese artefacto cultural que llamamos Derecho como el trabajo de los juristas pueden ser interpretados a la luz de su relación con las pasiones. Este trabajo discute las hipótesis complementarias del Derecho como pasión de los fuertes o como recurso de los débiles. Examina, en segundo lugar, algunas pasiones con las que se vincula al Derecho, entre ellas el poder, el odio y el resentimiento o la venganza y discute el modelo del jurista puro. Finalmente, propone el análisis de dos pasiones contradictorias: la pasión por el Derecho, que incluye el tipo extremo del afán de la justicia a todo trance, y la pasión por superar el Derecho como instrumento de mediación cial.

PALABRAS CLAVE

Derecho, pasiones, poder, odio, justicia, mediación social.

ABSTRACT

Both the social function of the cultural artefact that we call Law and the work of jurists, can be interpreted in term of passions. This paper discusses the complementary hypotheses of Law as a passion of the strong or as a resource of the weak. Secondly, it examines some passions which are inseparable from the Law, such as power, hate, resentment or revenge. Finally, the article proposes the analysis of two contradictory passions: the passion for the Law, which includes the extreme desire of justice at all costs, and the passion to overcome Law as a social mediation tool.

KEYWORDS

Law, passions, power, hate, justice, social mediation.

DOI: https://doi.org/10.36151/TD.2024.111

PASIÓN DE LOS FUERTES. EL DERECHO, COMO TRATADO DE PASIONES

Javier de Lucas
Catedrático de Filosofía del Derecho
Profesor emérito
Universitat de València

Sumario: Introducción. 1. Dos hipótesis no necesariamente contradictorias: las pasiones, motor del Derecho, o el Derecho, señor de las pasiones. 2. Pasión por y contra el Derecho. Notas. Bibliografía.

INTRODUCCIÓN

Como no se le escapará a cualquier lector mínimamente interesado en el cine, el título de este artículo tiene una inspiración inequívocamente fordiana. En efecto, *Pasión de los fuertes*, filme dirigido por el maestro John Ford en 1946[1], es uno de los más aclamados westerns de la historia. Pues bien, *pasión de los fuertes* me parece un lema particularmente adecuado para encabezar estas páginas, que versan sobre la relación entre las pasiones y el Derecho.

Para sentar las bases del discurso, debo aclarar que utilizaré el término «pasiones» en su acepción común en el discurso filosófico, psicológico y antropológico, significado que, desde Platón y, sobre todo, tras la sistematización que ofrece la *Ética* de Spinoza —preferible, a mi juicio, a la tipología que propone Descartes[2]—, sugiere la contraposición de los conceptos 'pasión' y 'razón', y añade al primero una dimensión emocional fuerte que conlleva, a su vez, las nociones de deseo, sentimiento o afección[3]. Fue Spinoza quien, además de su clasificación de las pasiones primarias (el deseo, alegría y tristeza) y secundarias (las dos pasiones clave: el odio y el amor), dejó sentado aquello de que «[…] nosotros no intentamos, queremos, apetecemos ni deseamos algo porque lo juzguemos bueno, sino que, al contrario, juzgamos que algo es bueno porque lo intentamos, queremos, apetecemos y

deseamos». Ante todo, el objetivo de estas páginas es plantear una mirada sobre la relación entre el Derecho y las pasiones que pone de manifiesto la existencia de argumentos sólidos para sostener que el Derecho puede ser explicado como un *tratado de pasiones*[4]. A tal efecto, dividiré mi exposición en dos partes.

En la primera (1) abordaré la paradoja básica que permea la relación entre las pasiones y el Derecho. Si bien el tópico nos dice que el Derecho es una de las creaciones culturales que sirven para moderar o dominar las pasiones (desde luego, las que se consideran malas e incluso *bajas pasiones* porque causan *daño*, entre ellas el odio, la ira, el resentimiento, la envidia, la codicia, los celos, etc.), lo cierto es que ese tópico es desmentido en gran medida por la experiencia: en realidad, y con frecuencia, el Derecho y el trabajo de los juristas parecen más bien guiados por esas pasiones. Algunos dirán que semejante hipótesis alude a casos especiales de la patología jurídica, a ciertas disfunciones excepcionales tanto del Derecho como de su sistematización, estudio e interpretación. Yo creo que, en todo caso, habría que hablar de funciones negativas y reconocer su carácter recurrente: no son hechos aislados. En resumen: las pasiones pueden ser entendidas como el motor, la yesca del Derecho, del mismo modo que el Derecho puede ser entendido como un instrumento racional para domeñar las pasiones[5].

La tesis de acuerdo con la cual las pasiones son la fuerza motriz del Derecho puede rastrearse ya en Sófocles, cuyo *Edipo*, al decir de Foucault, es el primer texto que aborda la relación entre verdad, justicia, poder y pasiones. La literatura ha sabido mostrar esta realidad desde sus orígenes: las pasiones mueven el recurso al Derecho, como saben todos los juristas y también los que frecuentan la literatura o el cine jurídico; hablo de pasiones como el resentimiento[6], la ira, el odio, la envidia, los celos o la codicia, entre otras.

Entre las pasiones relevantes para el Derecho hay que mencionar, ante todo, la pasión por el poder, que parece guiar la vida jurídica. Una *pasión por el poder*, entendido este en el sentido amplio de dominación de los otros y de lo otro que no se limita al ámbito político e institucional, dado que también impera en las relaciones entre particulares y es esencial para entender el origen y la evolución del Derecho privado. Lo cierto es que ese protoderecho que es la propiedad, la manifestación por excelencia del dominio, de la dominación, que sirvió como paradigma para la construcción teórica de los derechos públicos subjetivos, tiene mucho que ver con el recurso a lo que llamamos Derecho.

Prestaré atención, asimismo, a otra pasión jurídicamente significativa, la que enciende el sentimiento jurídico de lo injusto y puede derivar en una verdadera patología, la del *justiciero*, que dio lugar a algunos de los brocardos latinos más populares entre los juristas (*summum ius, suma iniuria*; *fiat iustitia et pereat mundus*) y ha sido tantas veces recreada en la literatura y el cine. Finalmente, haré una referencia sucinta al modo en que, se supone, deben afrontar las pasiones los diferentes tipos de juristas, entre ellos, desde luego, el juez, cuestión que obliga a abordar el tópico de la neutralidad u objetividad pura, así como la prevención con la que los juristas perciben la pasión de la empatía.

En la segunda parte (2), más breve, me interesa abordar el análisis, no muy frecuente, de la disyuntiva entre dos formas de pasión sobre el Derecho que, en mi opinión, actualmente

se enseñorean del mundo jurídico: por un lado, la *pasión patológica por el Derecho* —que lleva aparejada la pretensión de juridificarlo todo— tanto en su versión hiper-reglamentista como en la del querulante. Por otro, lo que podríamos llamar *la pasión patológica contra el Derecho*, inclinación de acuerdo con la cual es preciso orillar el Derecho porque es un instrumento propio de sociedades menos evolucionadas. De suyo, esta segunda manifestación de la pasión sobre el Derecho no es una novedad, al menos desde Hume, que vincula Derecho y escasez y que, por ello, pronostica que la utilidad del Derecho decaerá cuando se consiga superar las carencias materiales, tesis que será reiterada por Saint-Simon y Comte, y que alcanza su formulación más acabada en Marx, que pretende sustituir la dominación de las personas por la administración de los recursos una vez instaurado un sistema social en el que cada uno aporta según su capacidad y recibe según sus necesidades. Según esta aproximación, llegará un día en que el Derecho y los juristas serán superfluos, dado que resolveremos nuestros problemas de convivencia mediante la economía, la administración, las nuevas tecnologías y sus algoritmos, la estadística, la sociología y la biotecnología. En su versión más vulgar y populista, esta forma de pasión patológica que pretende sustituir al Derecho por otros dispositivos de regulación de las conductas y de mediación en los conflictos porque lo considera una herramienta obsoleta tiene diversas manifestaciones, entre ellas la a mi juicio descabellada aspiración de *desjudicializar la política*, como si la política fuera una realidad autónoma, ajena al Derecho, la moral o la ideología. Se trata de una pretensión que, a mi modo de ver, revela en el fondo una ideología no solo mal compadecida con la democracia, sino también ignara: frente a aquellos que pretenden contraponer la legitimidad democrática y la soberanía de la ley y el Estado de Derecho, considero imprescindible tener siempre presente que el Estado de Derecho y la soberanía de la ley son condiciones *sine qua non* de la democracia.

1. DOS HIPÓTESIS NO NECESARIAMENTE CONTRADICTORIAS: LAS PASIONES, MOTOR DEL DERECHO, O EL DERECHO, SEÑOR DE LAS PASIONES

Como decía, si estimo idónea la referencia a la *pasión de los fuertes* para hablar de la relación entre las pasiones y el Derecho es, ante todo, porque el Derecho puede ser visto —de hecho, así es percibido por amplios sectores de la población— como un instrumento particularmente útil, eficaz, al servicio de las pasiones y, en particular, de una pasión, la *pasión por el poder*, por la dominación, si preferimos la versión weberiana, lo que es lo mismo que decir la pasión de quien alcanza el poder y mantiene solo porque tiene más fuerza.

Esa versión de la *pasión de los fuertes* es el argumento del que podemos considerar el texto fundacional de nuestra tradición cultural, la *Ilíada* de Homero[7]. Es digno de señalar que la piedra fundacional de nuestra cultura sea un poema sobre la épica de la más extrema forma de despliegue de la fuerza, la guerra, cuyo primer verso evoca el formidable poder de una pasión, *la ira, la cólera de Aquiles*: μῆνιν ἄειδε θεὰ Πηληϊάδεω Ἀχιλῆος.

En esa tradición se inscribe el hilo argumental que presenta al Derecho como un ejercicio de la fuerza más allá de la razón, o con independencia de ella, impuesto por quien tiene más fuerza. Esta visión dio pie a una discusión filosófica de gran raigambre desde que Platón la pusiera en boca de los sofistas Trasímaco y Calicles, aunque quizá entre los juristas la referencia más conocida al vínculo entre fuerza, pasión de poder y Derecho sea la fórmula de Juvenal: *Sic volo, sic iubeo: stet pro ratione voluntas* (*Satirae,* 6, 223), esto es, el Derecho como manifestación del *imperium*.

Bien es cierto que, frente a la caracterización del Derecho como pasión (o voluntad) de los fuertes, cabe dar la vuelta al argumento, como hiciera ver Nietzsche, y sostener que el Derecho es todo lo contrario: una herramienta producto del resentimiento, *un recurso desesperado de los débiles, del rebaño*, para hacer frente al poder de los fuertes. Resulta incluso posible dar una vuelta de tuerca más a esa mirada nietzscheana y sostener, como hace Ferrajoli, que el Derecho entendido como fuerza encuentra su justificación, su legitimidad, precisamente cuando encarna la «ley del más débil» [8].

Es sabido que la discusión sobre Derecho, razón y fuerza enfrenta a diversas concepciones del Derecho, entre ellas, particularmente, el iusnaturalismo y el positivismo jurídico —y, dentro de este último, el normativismo y el realismo—. Creo que es conveniente alejarse de los estereotipos con los que muy frecuentemente se caricaturizan esas concepciones. Por ejemplo, el realismo jurídico sostenido por el juez Oliver Wendell Holmes. El *justice* Holmes fue un agudo jurista, quizá el más interesante precedente de la concepción propia del realismo jurídico[9]. A él se atribuyen frases como «el Derecho no es lógica, sino experiencia» o «el Derecho no es otra cosa que las profecías acerca de lo que los tribunales harán en concreto: nada más y nada menos», y de él se cuenta la famosa anécdota según la cual, interpelado para que materializara el cometido de un juez, incluso de un juez del Tribunal Supremo de EE. UU., es decir, «hacer justicia» respondió que su función no era hacer justicia, propósito, dijo Holmes, metafísico e incluso religioso, sino aplicar el Derecho al caso concreto: «Yo no me dedico a hacer justicia; yo aplico el Derecho»[10]. Holmes es, también, uno de los más característicos representantes del «creacionismo judicial», concepción que en su día se consideró exclusiva de la tradición del *Common Law*, pero que, como muestran los hodiernos representantes del *Richterstaat*, penetró en nuestra tradición jurídica y constituye en nuestro presente inmediato una de las paradojas más agudas de la evolución del Estado de Derecho en relación con la separación de poderes, paradoja que subyace a la caracterización misma del poder judicial en Montesquieu, para quien ese poder es en realidad un *poder invisible y nulo* porque el juez es *boca muda de la ley*, precisamente para que no usurpe la soberanía que viene representada por el poder legislativo.

Insisto en recordar que el realismo jurídico es una concepción del Derecho compleja, plural y próxima a cierto sociologismo jurídico. Prueba de la pluralidad de esta concepción es el hecho de que dentro del realismo hay representantes más extremos (Llewelyn y, sobre todo, Jerome Frank, el primer impulsor de las clínicas jurídicas[11]), autores a los que, con la mira puesta en su reacción frente a los excesos metafísicos de la concepción formalista del Derecho, han llegado tachados de teóricos de la *digestive jurisprudence* a propósito de la razón de ser del Derecho que supuestamente crean los jueces, sin que sea posible establecer

con certidumbre sus fallos debido a multitud de factores que influyen en su proceso de decisión. Hay otros realistas más sofisticados, por ejemplo, Alf Ross[12], recordado por su conocida polémica con Kelsen en torno al concepto de validez jurídica (debate que, a mi juicio, parece no haber entendido bien Habermas en su monumental obra de filosofía del Derecho, *Facticidad y validez*).

Lo que resulta innegable es que a esa visión *realista* yace el *leitmotiv* del *stat pro ratione voluntas*, que fácilmente propicia la interpretación del Derecho como instrumento de la pasión del más fuerte —inclinación, ya se ha señalado, que tiene hitos de prestigio en la historia de nuestra civilización—. No digamos nada de esa manifestación suprema del ejercicio de la fuerza, ese presunto *secundum iu*s que es el *ius ad bellum* y que conduce a la justificación de la guerra por medio del Derecho, es decir, al oxímoron de la *guerra justa*.

En todo caso, la triple tensión entre fuerza, razón y Derecho quizá encuentra su mejor solución, o la menos mala, en el famoso aserto de Radbruch: «Macht, ohne Recht, gilt nichts in dieser Erden; Recht, ohne Macht, kann niemals Sieger werden» («El poder, sin el Derecho, no vale nada en este mundo; el Derecho, sin el poder, no puede prevalecer jamás»)[13]. El Derecho, o al menos el *mejor Derecho*, no es tal sin el ejercicio de la fuerza, sí, pero este ejercicio debe ser proporcionado, ajustado a razón y no simplemente expresión de la voluntad del poder.

Precisamente es lo que trata de señalar otra visión del Derecho como *pasión de los fuertes*, la que insiste en que el verdaderamente fuerte acepta el control del Derecho porque en él radica la clave de la durabilidad del poder —quid que no es tanto la *auctoritas* cuanto la aceptación de cierto grado de autorrestricción—, es la que supone la aparición de la noción de Estado de Derecho, que desarrolla la crítica al poder *solutus a(b) legibus,* propio del *Ancien Régime*. Es la paradoja que consiste en entender que solo seré verdaderamente fuerte y solo podré asegurar mi pasión de poder si recurro al Derecho, si pongo al Derecho por encima del poder. Una *pasión razonable*, por tanto.

Este es, sin duda, el enfoque políticamente correcto, el más habitual si tratamos de examinar la relación entre el Derecho y las pasiones. Se trata, como es obvio, de sostener que el Derecho nace precisamente para domeñar esos motores individuales y sociales que son las pasiones porque producen *daño*. Es el *dictum* que nos legara nuestro Juan Luis Vives, que dejó escrito que el Derecho puede alcanzar poco más que «sujetar las manos y la ira». Algo que, bien pensado, no es poco: no se trata de eliminar las pasiones —tarea, por lo demás, imposible—, sino de someterlas a la razón mediante decisiones, normas y hábitos virtuosos. Pero, desde luego, las más de las veces esos instrumentos virtuosos no conseguirán su propósito si no es recurriendo a la fuerza: *no hay Derecho sin espada*, como bien muestra la iconología clásica.

El problema es que la tesis que concibe el Derecho como *señor razonable* que modera las pasiones lleva en no pocos casos al extremo de pretender que el Derecho y los juristas —particularmente, quienes administran la justicia— sean (deban ser) fríos, neutrales, objetivos, desprovistos de las pasiones que mueven al común de las gentes.

En cualquier caso, como he señalado, parece difícil negar que las pasiones, más incluso que las emociones, son la yesca del Derecho. De hecho, no sería arriesgado sugerir que cabe una mirada sobre el Derecho que nos lo muestre como *tratado de las pasiones*. A poco que se detenga a pensarlo, cualquier jurista, especialmente quienes tienen experiencia en la *práctica del Derecho*, caerá en la cuenta de que, en el día a día, nosotros —los juristas— nos dedicamos a enfrentarnos con las pasiones.

Por supuesto, lo acredita ese auténtico *tratado de pasiones* que es el Código Penal, como revela la mera lectura del índice del Libro segundo, pero también las pasiones son esenciales para apreciar la concurrencia de circunstancias modificativas de la responsabilidad criminal —agravantes y atenuantes—. Sucede que, además de ese lugar propio de las pasiones que es el Derecho penal, aquellas también están presentes en el núcleo fundacional del Derecho —el Derecho privado— es decir, en el Derecho civil y el Derecho mercantil: los códigos civil y mercantil son también tratados de pasiones: basta leer, por ejemplo, a Dickens, para hacerse una idea de las pasiones que mueven ese núcleo del Derecho[14]. Asimismo, y en no poca medida, el Derecho laboral es un tratado de pasiones, por no mencionar el Derecho constitucional y la ciencia política, que tienen a la *pasión de poder* como *leitmotiv*. Incluso lo es el Derecho internacional, que está recorrido por esa pasión de poder y por la furia, como ha mostrado la profesora Ramón Chornet en un reciente libro dedicado a analizar la llamada guerra contra el terrorismo[15]. Esa relación explica la peculiar vinculación entre la retórica y el Derecho. Hablo de retórica, entendida en un sentido más amplio que argumentación, para recordar el *motto* de Aristón, que define la retórica como **el** arte de engañar al pueblo mediante la manipulación de las pasiones del auditorio al que se dirige el discurso con el fin de dominarlo. En definitiva, como enseñan los demagogos, para así conquistar y permanecer en el poder. Eso explica la acerba crítica de Tácito: «La retórica es un instrumento inventado para manejar y agitar la turba y al pueblo desordenado, un instrumento que sólo se emplea en los Estados enfermos».

Como decía, esta presencia de las pasiones como motor del Derecho -la ira, el odio, la venganza o la codicia- o, si se prefiere, en un sentido negativo, esto es, esa pasión que es el sentimiento de lo injusto, la rabia ante lo que es percibido como injusticia (el sentido original del «¡no hay Derecho!»), es una realidad que ha sabido mostrar la literatura desde momentos muy iniciales: es imposible dejar de mencionar la pasión por la justicia que encarna *Antígona,* la primera gran protagonista de la lucha por el Derecho entendida como una exigencia que va más allá de la obediencia a la legalidad. Particularmente interesante es el reflejo en la literatura de la experiencia de la *pasión por la justicia* como motor del Derecho.

La pasión por la justicia y su patología

Probablemente sea Rudolf Ihering quien explica mejor esa patología de la pasión extrema por el Derecho (como respuesta a un sentimiento de rabia ante la injusticia) cuando recurre a *Michael Kohlhaas*, la novela corta de Heinrich von Kleist, para ilustrar su tesis de *la lucha por el Derecho*[16]. Es este un hilo narrativo que recorre la literatura clásica, desde Sófocles y Aristófanes a Shakespeare y Cervantes, de Dostoievski a Kafka hasta esos

contemporáneos descendientes del Sherlock Holmes de Arthur Conan Doyle que son los relatos jurídicos de von Schirach[17]. No en vano, la sabiduría jurídica clásica condensó los riesgos de esa pasión extrema por el Derecho (entendido como justicia o, si se quiere, como la *pasión del justiciero*) en dos brocardos: *summum ius, suma iniuria*, y *fiat iustitia et pereat mundus.*

En la nómina de esos *justicieros* que viven su pasión por la justicia hasta el desborde de las riendas de la razón[18] se inscriben el Shylock de Shakespeare y, sobre todo, el justiciero por excelencia, Don Quijote[19], personaje que se autodefine en una famosa cita que muchos llegamos a aprender de memoria: «Y es mi oficio y ejercicio andar por el mundo enderezando entuertos y desfaciendo agravios».

Con todo, la pasión por la justicia no comporta necesariamente una patología, como ilustra el debate en torno al controvertido episodio de los galeotes en *El Quijote*, en el que contrasta la actitud más legalista que encarna Sancho —propia de un formalismo jurídico apegado a la práctica más común—, frente a la pasión —compasión— encarnada en nuestro caballero andante. Porque aquí aparece el sentido de la compasión entendida no como la simple conmiseración paternalista, sino como la capacidad de compartir la pasión del otro y, muy frecuentemente, su sufrimiento. Dicho de otro modo, nos encontraríamos aquí ante una *pasión positiva*, la que nos lleva a empatizar con la injusticia sufrida por los otros, incluso más allá de la justicia. A mi juicio, y coincido en ello con Manuel Atienza, la *compasión* que muestra Don Quijote en ese pasaje no debe confundirse con la equidad, con la tolerancia o con la *pietas* romana, asunto que, por su trascendencia jurídica, me parece merecedor de una consideración particular. Veámoslo.

Ante todo, Me parece claro ante todo que la compasión, entendida como capacidad de empatía que sobrepasa la justicia, tal y como propone Arteta en coincidencia con Nussbaum, no es identificable con la tolerancia. En todo caso, la compasión, para ser cabal y, desde luego, para no desaguar en un paternalismo hiriente, la compasión requiere no solo la capacidad empática, el ponerse en los zapatos del otro —de acuerdo con el famoso alegato de otro personaje literario modelo de juristas, el Atticus Finch de *Matar a un ruiseñor*—, sino algo más que le falta a nuestro caballero cuando mira a los galeotes, a saber, la capacidad de sentir lo mismo que el otro, lo que guarda relación con el hecho de que haber vivido su experiencia dolorosa o una muy similar. Así lo he tratado de mostrar, echando mano de la lección que recibe la hija de Atticus por parte de su aya, Calpurnia, en uno de los diálogos de *Ve y pon un centinela* —primer borrador de la antes citada obra de Harper Lee— al tratar de explicar por qué el personaje de Atticus Finch resume los rasgos del *buen jurista*[20].

Lo que trato de señalar es que advertir la condición de miserables, la *vulnerabilidad* de los galeotes, requiere un plus que no es exigible ni probablemente está al alcance del Derecho, dado que lo supera. Es, en sentido profundo, esa capacidad de *padecer con* el otro que supo conceptualizar la gran filósofa del siglo XX, Simone Weil[21], a la que Camus —el responsable de la difusión de las obras de Weil desde Gallimard— calificó como «el único gran espíritu de nuestro tiempo».

Es cierto que, en un sentido, por así decirlo, funcional, la relación entre Derecho y compasión tiene que ver con esa intuición original a la que ya aludí, la capacidad del Derecho para embridar las pasiones, tarea que se espera del jurista, en particular del juez, que se ve, a su vez, ante la exigencia de lograr el difícil equilibrio entre la cordura y la pasión en el ejercicio de la función de juzgar. Los juristas, en particular los jueces, deben ser formados en el terreno de las virtudes.

Sí. El Derecho, como producto humano, histórico, es una creación cultural que no se explica sin la existencia de virtudes y pasiones: desde la venganza y el despecho hasta la codicia o los celos, desde la envidia a la ira hasta el afán de gloria, al igual que sucede, sin ir más lejos, con la economía, como nos explicaron aquellos escoceses dieciochescos (Hume, Smith, Mandeville) que, además de economistas, eran filósofos dedicados al estudio de los sentimientos y las pasiones morales.

Ahora bien, al mismo tiempo que encontramos ese *pathos* inescindible de la vida jurídica real, de la experiencia del Derecho, es preciso reconocer la complejidad de esa experiencia: el Derecho trata de *sujetar las pasiones* y, por eso, encuentra su mejor sentido en el marco del proyecto civilizatorio que hallamos en Platón, Aristóteles y los estoicos: *no se trata de eliminar las pasiones —tarea, por lo demás, imposible— sino de someterlas a la razón mediante el cultivo de hábitos virtuosos*. Como ha escrito Atienza: «La razón de ser del Derecho, al fin y al cabo, no puede ser otra que la de encontrar algún remedio al conflicto social, esto es, a un tipo de mal que causa dolor, por más que los remedios jurídicos no suelen ser precisamente indoloros (por lo menos para una de las partes del conflicto); pero, precisamente por eso, por esa conciencia de que los males sociales que el Derecho trata de combatir (dejemos de lado los casos en los que el Derecho es la causa de ellos) no tienen una perfecta curación, quienes hacen uso de la maquinaria jurídica y particularmente quienes controlan los resortes que pueden causar un mayor daño (no son sólo los jueces), tendrían que mostrar un comportamiento compasivo, sin entrar por el momento en cómo haya que entender la compasión» (Atienza 2020*a*: 20-21). Desde luego, esos hábitos virtuosos muchas veces no se conseguirán si no es con el recurso a la fuerza: ya he señalado arriba que, como bien muestra la iconología clásica, *no hay Derecho sin espada*.

El proyecto de una convivencia pacífica de contrarios, recuerdo, encuentra distintas vías civilizatorias. La de la *paideia*, la *buena educación* consiste en la transmisión de las virtudes, su aprendizaje y su aprehensión por el sujeto. La vía del Derecho, más *realista*, consiste en configurarlas como normas, esto es, como mandatos reforzados por la amenaza de la sanción, la coacción o, más sencillamente, la fuerza, que es lo que permite imponerlas. El peso de la fuerza que acompaña inexorablemente al Derecho disminuirá en la medida en que el proceso civilizatorio alcance lo que llamamos legitimidad democrática del Derecho: es decir, en la medida en que lo que se propone como pautas de comportamiento (hábitos virtuosos) se haya convertido en virtudes cívicas exigibles y aceptables, y lo son, deben serlo, en primer lugar, entre los juristas, en particular entre los jueces. Eso sucede cuando se proponen como pauta para seguir comportamientos que la mayoría acepte racionalmente como virtuosos —en el sentido de imprescindibles— e incluso como deseables para la

consecución de los objetivos de convivencia que se han decidido por común (mayoritario) acuerdo.

Sobre el tipo ideal del jurista sin pasiones

Estas exigencias refuerzan la tesis del tipo ideal del *jurista ecuánime* que sabe rastrear las pasiones sin contagiarse de ellas. Es el modelo del jurisconsulto romano que nos ha legado un propósito de objetividad (muchas veces confundido con la neutralidad y, peor aún, con la asepsia valorativa) propio de una cierta cultura jurídica, la del positivismo legalista, impulsada por el anhelo de evitar que los operadores jurídicos por excelencia, los jueces, pudieran poner palos en la rueda legal de la revolución que acabó con el Antiguo Régimen y que se concretó en el brocardo de Montesquieu que nos presenta a los jueces como boca muda de ley.

Pues bien, a mi juicio, ese modelo debe ser matizado no solo porque resulta irreal pensar en un sujeto jurídico desprovisto de pasiones o de ideología, sino también porque no me parece deseable ese modelo de operador jurídico del que es botón de muestra la concepción de juez como mera *boca de la ley*, como aplicador mecánico. Creo que el tipo de operador jurídico —desde luego, de juez— que debemos tratar de formar es aquel que, conociendo las pasiones e intereses y siendo él mismo sujeto de esas pasiones e intereses, las somete a control para saber realizar su función de mediación en los conflictos. Es cierto que eso no será posible si, además de la observancia de la ley que le vincula y le confiere la legitimidad en su tarea de mediación, el juez no lleva a la práctica esas virtudes que equilibran las pasiones.

Pero la relación entre las pasiones y el Derecho no desaparece, e incluso puede retornar con mayor acuidad, haciéndonos sentir que ese Derecho es solo pasión desbocada, violencia. No olvidemos que, frente a quienes describen de manera caricaturesca el Derecho y representan a los juristas —en particular, a los jueces— como un objeto y unos sujetos anclados a un mundo frío, ajenos a los sentimientos de la gente común y corriente, alejados de lo que preocupa en la calle, cabría apuntar que, en no pocos casos, podría ser más bien que las normas —y, sobre todo, las interpretaciones de los juristas— que nos chocan fueran el fruto de intereses, pasiones y sentimientos propios, ajenos a los de la mayoría.

2. PASIÓN POR Y CONTRA EL DERECHO

En esta segunda parte quisiera abordar una contraposición que, me parece, está particularmente presente en nuestro contexto. Creo que, aunque parezca paradójico, asistimos hoy a la convivencia de dos miradas apasionadas y contradictorias sobre el Derecho. En este sentido, puede hablarse tanto de una pasión *por* el Derecho como de una pasión *contra* el Derecho.

Sin duda, la experiencia del *terribile diritto*, por utilizar la expresión acuñada por Rodotà (2013), genera pasiones negativas frente al Derecho (miedo, desconfianza, e incluso ira antijurídica), aunque en no pocas ocasiones de forma contradictoria, como creo que sucede hoy, tal y como se ejemplifica, sobre todo, en las redes sociales, pero también en no pocos medios de comunicación tradicionales: prensa, radio, televisión.

Por un lado, hay que estar ciego para no detectar hoy el crecimiento exponencial que actualmente experimenta de ese *anhelo del Derecho* que ejemplificaba Shylock. Es difícil no ver cómo crece sin medida la pasión litigante, cómo florece la pasión legiferante, reglamentista, sobre los aspectos más nimios de la realidad, hasta qué punto bordeamos esa otra pasión de monopolio del Derecho que lleva al extremo del *RichterStaat*, un gobierno de quienes en puridad no deben ser gobernantes, sino guardianes —los jueces o, en fin, cómo la pasión vindicativa propia del justiciero, al que da alas el populismo penal, impulsa y alienta una marea prohibicionista que, a 50 años de mayo del 68 y de su *prohibido prohibir*, parece querer prohibir y castigar sin descanso[22].

Por otro lado, y simultáneamente, asistimos a una aparente desconfianza generalizada hacia el Derecho, que generalmente se presenta como *miedo* ante la fuerza del Derecho, pero que, a veces, se manifiesta a través de otra pasión: la *furia contra el Derecho*, al menos *contra quienes nos dicen qué es Derecho*, particularmente los abogados y jueces, figuras respecto a las que nuestro refranero nos previene («pleitos tengas y los ganes»). Esta pasión aparece claramente ejemplificada en ese pasaje de *Enrique IV* en el que Shakespeare pone en boca del carnicero Dick esta frase: «Let's skill all the lawyers».

Pues bien, sin abandonar el campo de esta percepción negativa, me parece que hoy abunda una pasión que, a falta de mayores precisiones, llamaría el *menosprecio* por el Derecho. Esta pasión es fomentada y estimulada desde trincheras ideológicas que no son las habituales —es decir, las clásicas tesis anarquistas, comunistas o libertarias que rechazan de plano el Derecho porque lo consideran un simple instrumento de dominación y explotación, pura violencia institucionalizada—, sino muy otras: por ejemplo, ciertas versiones del nacionalismo y ciertas versiones del feminismo. El menosprecio por el Derecho también se expresa desde las alturas —o tal vez los abismos— de la ciencia y, particularmente, desde el estrado en el que están instalados ciertos científicos sociales al alza (mediáticos, digámoslo así) que no dejan de salmodiar sobre la importancia de «la política» y sobre la necesidad de superar el torpe recurso al Derecho y a sus instrumentos, algo secundario, claro, una actitud displicente a la que no son ajenos numerosos periodistas y comunicadores.

Hablo, por ejemplo, de esos escenarios dominados por escribidores y locutores (me cuesta llamarles periodistas) que jalean el linchamiento de jueces machistas, prevaricadores, corruptos y demás despreciable ralea, y que nos explican —desde su contacto privilegiado con la realidad y, al parecer, su dominio sobre los más recónditos arcanos del Derecho— cuándo tal o cual comportamiento es ilícito, cuándo es justa o abominable una u otra sentencia —que no acostumbran a leer, no digo ya a estudiar, dado que vierten su furibunda crítica o su juicio encomiástico en el preciso momento en que se anuncia la publicación de la resolución judicial— adornando su verborrea con insólitos conocimientos procesales que deben, sobre todo, a *gargantas profundas* de los pasillos de tribunales y no a aburridas y

nada *glamurosas* horas de estudio. Lo hacen porque dicen que saben lo que la calle concibe y quiere como justo, que sería algo muy distinto a lo que han secuestrado como justo los *clérigos* que administran (usurpan) el (verdadero) Derecho.

Todavía más preocupante me parece el caso de respetados politólogos que, desde la tribuna de la ciencia política —que muchas veces parece más bien púlpito de un predicador—, nos aleccionan sobre cuándo hay un delito de rebelión, cuándo de sedición o cuándo, simplemente, una manifestación cívica mientras imprimen alguna inflexión gamberra a su discurso, nutrido de una sabiduría sobre el Código Penal adquirida a base de lecturas de Wikipedia, como si el Derecho no mereciera mayor atención. Lo hemos visto en algunos artículos que argumentan sobre la relevancia jurídica y política de las actuaciones del Govern nacionalista de la Generalitat de Catalunya y de su Parlament, a las que restan importancia mientras proclaman que la verdadera amenaza real para la democracia es el monstruo del nacionalismo español, todo ello sin haber leído aparentemente una página de las que Kelsen dedica a los *coup d'Etat* jurídicos en su *Teoría pura del Derecho*. Parece como si quisieran instruirnos con esta consigna: dejemos esto del Derecho, que al fin y al cabo lo podemos cambiar cuando queramos, y vayamos a lo importante.

No me resisto a apuntar, por cierto, que todavía estamos esperando que esos gurús nos expliquen cómo se puede *hacer política*, no ya excelsa, sino simplemente civilizada —es decir, algo mejor que la nuda imposición de la voluntad del que más puede— sin recurrir al Derecho, y que nos aclaren dónde quedarían los intereses del común —no digamos de los más vulnerables— si todo fuera negociación («pónganse a hablar», conminan esos iluminados), olvidando que, si se trata de negociar sin más, como ellos pregonan, es decir, de «superar» las tediosas normas, las estériles instituciones, los inútiles procedimientos y las amenazantes y siempre inconvenientes sanciones del artefacto jurídico, la palabra quedaría como atributo exclusivo de los que están *de facto* en condiciones de hacer o dictar el negocio. El monopolio de una élite que ya no está conformada por reyezuelos perezosos y viciosos ni juristas entogados, sino por elegantes CEO y ejecutivos con más desprecio e ignorancia por las necesidades y preocupaciones del común de los mortales que la que exhibían aquellos déspotas con los que aún quieren asustarnos.

Obviamente, la negociación adquiere un tinte distinto si se trata de negociar bajo el imperio del Derecho (hablo del Estado de Derecho), algo que, por cierto, no tiene nada que ver con la pretensión —a mi juicio, inaceptable— de negociar «sin condiciones previas» tantas veces se ha defendida en nuestro país (por ejemplo, aunque no solo, desde el catalanismo independentista). A mi juicio, semejante pretensión no es sino una incitación al enfrentamiento de pasiones, una instigación a mantener una justa en la que lo único que se dirime es quién resiste y puede más reafirmándose en las suyas.

Al cabo, lo que podemos aprender de este sumario recorrido es que la ambición de que el Derecho represente el dominio de la razón sobre las pasiones hasta alcanzar el objetivo de erradicarlas es una pretensión vana. Como he señalado arriba recordando las palabras maestro Luis Vives, el Derecho puede aspirar, como mucho, a «sujetar las manos y la ira». Pero insistiré: ese no es un resultado tan magro, dado que se trata de racionalizar y sujetar

las pasiones a través de la mediación de las normas e instituciones que permiten alcanzar acuerdos respetables.

Sea como fuere, las pasiones no desaparecen: siguen ahí, presentes en todos los ciudadanos y son más difíciles de someter o incluso de regular y controlar cuando se trata de quienes tienen poder. También, evidentemente, en los propios juristas, por más que a ellos les exigimos un plus, que está implícito en esa iconografía de la justicia a la que ya me he referido: además de la espada, la balanza y el equilibrio, nos habla de esa racionalización de las pasiones, al igual que la venda que cubre los ojos de la justicia. En caso contrario, la espada con la que se adorna nos parecería una exacción y, como dejó escrito San Agustín, finalmente no habría diferencia alguna entre el mandato del Derecho y el de una banda de ladrones.

La consecuencia es clara: hay que tratar de formar a los juristas en el cultivo de ciertas virtudes que ayudan a racionalizar las pasiones, pero, sobre todo, hay que vigilar con la mayor atención las pasiones de quienes tienen el poder de decidir sobre nosotros desde el Derecho.

Vuelvo al luminoso ensayo de Simone Weil que he citado anteriormente: el texto de Homero pone de manifiesto la acción de la fuerza que somete tanto a vencedores como a vencidos. Entre los resquicios del imperio de la fuerza, sin embargo, aparece de forma casi milagrosa la gracia. Esa es la enseñanza más importante del poema homérico: la lección última de la *Ilíada* es la transformación de la cólera de Aquiles, gracias a la compasión hacia Príamo, la *pietas*, la piedad por los muertos, la piedad por las familias de los muertos, la piedad que vence a la crueldad. La esperanza está en la pasión común nacida de la común convicción sobre la fragilidad humana, la condición común de humanidad, una piedad que nos hace capaces de no sucumbir a la fascinación de la fuerza.

NOTAS

1. Lo cierto es que el título original de esa película es otro: *My Darling Clementine.* Seguro que muchos conocen e incluso han tarareado esa canción, que ilustra momentos clave de la película. Lo que quizá sea menos conocido es lo que apunta el conocido hispanista Gerald Brenan, quien sostiene que el origen de esta canción se remonta a una balada gascona, *Aqueras Montanhas*, basada a su vez en el romance español *Dónde vas, buen caballero.*

2. Aunque es posible encontrar una aproximación filosófica a las pasiones en Platón y en el *Del Alma* de Aristóteles, no existe en ellos un tratado de las pasiones. No obstante, la *Ética* de Aristóteles distingue entre facultades, modos de ser y pasiones («llamo pasiones a lo que sigue: indignación, temor, vergüenza, apetito y en general todo lo que va acompañado de placer o dolor». Tampoco hay un tratado de las pasiones en San Agustín, que las considera siempre desde la perspectiva del amor —para él, la emoción básica para él– ni en Santo Tomás de Aquino, que considera las pasiones como *movimientos* del alma y del cuerpo y propicia la teoría de las malas pasiones, que da lugar a la concepción teologal de los vicios capitales. El tratamiento sistemático de las pasiones nos remite a Descartes y Spinoza. Como se recordará, en su última obra (Descartes, 1648), dedicada a la reina Cristina de Suecia, distingue seis pasiones primigenias (admiración, amor, odio, deseo, alegría y tristeza). En buena medida, Spinoza coincide con

ese elenco en su teoría de las pasiones, que arranca de un principio formulado en la proposición III.6 de su *Ética*1(679): «Cada cosa se esfuerza, cuanto está a su alcance, por perseverar en su ser». En la proposición III, 9 explica que ese *conatus* en el alma se llama voluntad y que, cuando está en el alma y el cuerpo, se llama apetito. El apetito es, por lo tanto, la esencia del hombre entendido como conjunto de alma y de cuerpo. En este marco teórico, Spinoza introduce su famosa cita: «Nosotros no intentamos, queremos, apetecemos ni deseamos algo porque lo juzguemos bueno, sino que, al contrario, juzgamos que algo es bueno porque lo intentamos, queremos, apetecemos y deseamos». Spinoza distingue pasiones primarias (el deseo, y junto a él, la alegría y la tristeza) y secundarias, entre las que considera claves el amor y el odio. La primacía del deseo es la clave.

3. Son las acepciones 5.ª, 6.ª y 7ª que recoge el Diccionario de la Real Academia (DRAE). Así, la 5.ª: «Perturbación o afecto desordenado del ánimo» (cuyos sinónimos son emoción, arrebato, frenesí…); la 6.ª: «Inclinación o preferencia muy vivas de alguien a otra persona (sus sinónimos serían preferencia, inclinación, predilección…), y la 7.ª: «Apetito de algo o afición vehemente por ello» (sinónimos: deseo, entusiasmo, vehemencia).

4. Sobre lo que en otro lugar he llamado —con un término a la moda, pero a mi juicio no improcedente— «narrativa de las pasiones y el Derecho», hay que recurrir a la obra de François Ost, a mi juicio el autor que más y mejor ha escrito sobre esta temática. Ost es un especialista en la relación entre Derecho y literatura (Ost 2004; 2011; 2018 y 2021) e incluso ha escrito obras de teatro. Es Ost quien señala que un personaje de Aristófanes, el juez Filocleón, es el primer *justiciero*, en el sentido de *obsesionado por la justicia*, junto a los personajes antagónicos creados por Sófocles, Edipo y Antígona. Recomiendo en particular el libro subtitulado shakesperianamente *I crave the Law,* que es el comienzo del conocido alegato de Shylock en *El mercader de Venecia*: «My deeds upon my head! I crave the law, the penalty and forfeit of my bond!»).

5. Es interesante señalar la convergencia entre las tesis de la escuela escocesa de economía política y la teoría de Jhering sobre el interés como clave de comprensión del Derecho: en realidad, unos y otro proponen la noción de *interés como eje sobre el que puede bascular el control racional de las pasiones. Como se recordará, fue Albert Hirschmann quien subrayó ese potencial racionalizador propio de la lógica del capitalismo —no atribuido al Derecho, sino a la ratio oeconomica—, en su* ensayo de 1977 *The Passions and the Interests: Political Arguments For Capitalism Before Its Triumph*, en el que, frente a las tesis de Marx y Weber, propone una interpretación que es una ironía histórica: el capitalismo logró exactamente lo que había denunciado como su peor característica, la represión de las pasiones en favor de los *inofensivos* intereses de la vida comercial.

6. Junto a la teorización de Nietzsche sobre el papel central del resentimiento, hay que tener presente la obra capital de Scheler (1993).

7. Así lo supo explicar magníficamente la filósofa Simone Weil en un breve pero profundo ensayo, cuya lectura me recomendaron en su día mis compañeros Emilia Bea y Massimo La Torre y que lleva por título *Iliade ou le poéme de la force.*

8. Es una tesis defendida por Ferrajoli en diversos ensayos, particularmente en Ferrajoli (2002)

9. Sigue siendo muy recomendable la lectura de su ensayo *The Path of the Law* (1881). Hay una buena edición en castellano de J.I Solar, *La senda del Derecho*.

10. Así lo cuenta, entre otros, Dworkin: «Siendo Oliver Wendell Holmes magistrado del Tribunal Supremo, en una ocasión de camino al Tribunal llevó a un joven Learned Hand en su carruaje. Al llegar a su destino, Hand se bajó, saludó en dirección al carruaje que se alejaba y dijo alegremente: «*¡Haga justicia, magistrado!*». Holmes paró el carruaje, hizo que el conductor girara, se dirigió hacia el asombrado Hand y, sacando la cabeza por la ventana le dijo: «**¡Ese no es mi trabajo!**» A continuación, el

carruaje dio la vuelta y se marchó, llevándose a Holmes a su trabajo, supuestamente consistente en no hacer justicia». Eso de «hacer justicia», pertenecería a Dios, pero no a los hombres; el Juez sólo debe limitarse a aplicar la ley al caso concreto, respetando las reglas establecidas para cada procedimiento y punto. Hacer justicia no es trabajo de los jueces. La cita se encuentra en (Dworkin, 2007).

11. En este sentido, puede consultarse Frank (1930); póstumamente, su hija, Barbara Frank Kristein, publicó *A Man's Reach: The Selected Writings of Judge Jerome Frank*.

12. Sus dos obras de referencia son Ross (1953) y Ross (1968). Inicialmente, en su trabajo de tesis doctoral (1929) sostuvo una concepción próxima a Kelsen. Pronto pasó a una posición más crítica, enunciada en (Ross, 1946).

13. Radbruch acuñó lo que Richard Lange denominó «fórmula de Radbruch» sobre las leyes injustas y la relación entre Derecho, justicia y poder en Radburch (1946). La revisó con posterioridad en Radbruch (1959) y en su ya clásico manual *(Radbruch,1990)*.

14. No puedo dejar de señalar aquí la obra capital (Menger,1898). Se puede consultar la edición española de Comares, que conserva la traducción de Adolfo Posada, y el estudio introductorio de José Luis Monereo que explica a Menger desde el prisma del reformismo social y el socialismo jurídico.

15. Me refiero a Ramón Chornet (2021).

16. Ihering fue probablemente el primero en recurrir al nexo entre Derecho y literatura a través de esa novela breve de Kleist, para explicar el motor de la lucha por el Derecho (Ihering, 1872). La obra Fue traducida al castellano con el título *La lucha por el Derecho*, por Adolfo Posada y se publicó en 1921 por Vitorino Suárez, con prólogo de Leopoldo Alas, *Clarín,* que, según el propio Losada, fue el impulsor de esta edición en castellano. Conviene recordar que, bajo el motto *ridendo dicere verum*, Ihering había publicado antes su *Scherz und Ernst in der Jurisprudenz. Eine Weihnachtsgabe für das juristische Publikum* (1848)*,* una inolvidable sátira de la dogmática jurídica de la época, un modelo de estudio del Derecho situado en un paraíso conceptual ajeno a las pasiones e intereses. La primera versión en castellano se tituló *Jurisprudencia en broma y en serio,* y fue traducida por Román Riaza para la editorial Revista de Derecho Privado en 1933. El original conoció diferentes ediciones ampliadas y la que se considera hoy como versión castellana más completa, publicada por Civitas en 1987, se basa en la decimotercera edición alemana (1924). La edición de Civitas, publicada con el título *Bromas y veras en la ciencia jurídica: un presente navideño para los lectores de obras jurídicas*, reproduce la versión publicada en Buenos Aires en 1974, que traduce la edición alemana de 1924 (la traducción es de Tomás A Banzhaf).

17. De Ferdinand von Schirach recomiendo casi todos sus relatos, que han sido traducidos al castellano: *Crimen* (2009), *La culpa* (2010); *El caso Collini* (2011); *Castigo* (2018), llevada a la televisión en 2022, y *Café y cigarrillos* (2019).

18. En el elenco de personajes desbordados por una pasión por la justicia más allá de la razón, además del Michael Kohlhaas de Kleist (que Mario Losano encuentra actualizado en la película de Zhang YiMou, *QiuJu, la historia de una mujer china*, León de Oro en el Festival de Venecia de 1992), hay que tener presentes los patológicos y obsesivos, representados en el extremo por el Raskolnikov de Dostoievski y por dos de los grandes personajes de Kafka que sufren el dolor que causa la maquinaria de la justicia: K. y Peter el rojo.

19. Son muy numerosos los ensayos sobre el Derecho en la gran novela de Cervantes. Suele mencionarse el ensayo de Alcalá Zamora (1947). Entre los más recientes, *vid.* Aguilera Barchet (2005); Prats Westerlindh (2006); Botero Bernal (2009), Fernández Montalvo (2016); y Atienza (2020).

20. Sobre el modelo deontológico o lo que podría llamarse el «patrón Atticus Finch» he escrito en De Lucas (2020: 19-69).

21. Para su reconocimiento como tal (algo en lo que coincido con otro colega común, Juan Ramón Capella), baste leer Weil (1996 o 2014). En castellano, *vid.* los ensayos de Bea (1992 y 2010), la mejor especialista sobre su obra que conozco.

22. Algunos ensayistas llegan a hablar del riesgo de «juristocracia», expresión que bascula entre la denigración y la denuncia, y que creen haber inventado. Lo cierto es que, como tal, el término fue acuñado por el constitucionalista y magistrado Bela Pokol (2021). Creo, en todo caso, que se trata de una versión actualizada del mencionado riesgo de *RichterStaat.*

BIBLIOGRAFÍA

AGUILERA BARCHET, Bruno (2005): «El Derecho en el Quijote. Notas para una inmersión jurídica en la España del Siglo de Oro», *Anuario de Historia del Derech*o, LXXVII.
ARTETA, Aurelio (1996): *La compasión: apología de una virtud bajo sospecha,* Barcelona: Paidós.
ATIENZA, Manuel (1998): «Virtudes judiciales. Selección y formación de los jueces en el Estado de Derecho», *Claves de Razón Práctica*, 86, 32-42.
— (2020): *Una apología del Derecho y otros ensayos*, Madrid: Trotta.
BEA, Emilia (1992): *Simone Weil. La memoria de los oprimidos*, Madrid: Encuentro.
— (coord.) (2010): S*imone Weil, la conciencia del dolor y de la belleza*, Trotta.
DE LUCAS, Javier (1996): «Tolerancia y Derecho», *Isegoría,* 14, 152-163.
— (2016): «Déficits y falacias de la democracia liberal ante la gestión de la diversidad cultural: el caso de las políticas migratorias y de asilo», *Deusto Journal of Human Rights*, 1, pp. 15-37.
— (2020): *Nosotros que quisimos tanto a Atticus Finch*, Valencia: Tirant lo Blanch.
HIRSCHMANN, Albert (2013): *The Passions and the Interests: Political Arguments for Capitalism Before its Triumph*, Princeton: Princeton University Press.
HOLMES, Oliver Wendell (2015 [1881]): *The Path of the Law*, New Orleans: Quid Pro Books.
IHERING, Rudolf (1921): *La lucha por el Derecho* (traducción de A. Posada con prólogo de L. Alas), Madrid: Librería General Victoriano Suárez.
— (1987): *Bromas y veras en la ciencia jurídica: un presente navideño para los lectores de obras jurídicas*, Madrid: Civitas.
MENGER, Anton (1998): *El derecho civil y los pobres*, Madrid: Librería General Victoriano Suárez.
OST, François (2012): *Shakespeare. La Comédie de la Loi*, París: Michalon.
— (2018): *Le droit, objet des Passions?,*Bruselas: Académie Royale de Belgique.
— (2021a): *Le droit ou l'empire du tiers*, París: Dalloz.
— (2021b): *Nouveaux Contes juridiques*, París: Dalloz.
RAMÓN CHORNET, Consuelo (2021): *La guerra contra el terrorismo, veinte años después. Zero Dark Thirty,* Valencia: Tirant lo Blanch.
RODOTÀ, Stefano (2013): *Il terribile diritto. Studi sulla propietà privata e i beni comuni,* Bolonia: Il Mulino.
SCHELER, Max (1993): *El resentimiento en la mora*l, Madrid: Caparrós editores.
WEIL, Simone (1996): *Echar raíces* (traducción J. C. González y J. R. Capella), Madrid: Trotta
— (2014): *La condición obrera* (Introducción de R. Chenavier), Madrid: Trotta.
— (2023): *La Ilíada o el poema de la fuerza*, Madrid: Trotta.

Fecha de recepción: 12 de julio de 2024.
Fecha de aceptación: 20 de octubre de 2024.

ANTISUBORDINACIÓN Y DISCRIMINACIÓN ALGORÍTMICA
ANTI-SUBORDINATION AND ALGORITHMIC DISCRIMINATION

Anna Capellà i Ricart
Investigadora Postdoctoral
Institut de Dret i Tecnologia
Universitat Autònoma de Barcelona

RESUMEN

La perspectiva antisubordinatoria del derecho antidiscriminatorio se basa en la noción de igualdad como no exclusión. Tal como opera actualmente en distintos contextos, la inteligencia artificial contribuye a la perpetuación de sistemas sociales de opresión/subordinación que generan exclusión. Este artículo analiza la dicotomía entre las corrientes anticlasificación y antisubordinación y el impacto de ambas perspectivas en el combate contra la discriminación algorítmica. A tal efecto, se estudian cuatro elementos clave para entender la perspectiva antisubordinatoria (los estereotipos, la no neutralidad, la descontextualización y la monocultura) que muestran que la misma crítica que se hace desde la perspectiva antisubordinatoria al Derecho puede dirigirse a la inteligencia artificial, dado que en estos dos ámbitos se verifica una concentración de poder. Finalmente, proponemos algunas medidas antisubordinatorias orientadas a hacer frente a la discriminación algorítmica.

PALABRAS CLAVE

Antisubordinación, inteligencia artificial, Derecho, opresión, algoritmo.

ABSTRACT

The anti-subordination perspective of anti-discrimination law is based on the notion of equality as non-exclusion. As it currently operates in different contexts, artificial intelligence contributes to the perpetuation of social systems of oppression/subordination that cause exclusion. This article analyzes the dichotomy between the anti-classification and anti-subordination perspectives and their impact in the fight against algorithmic discrimination. To this end, four key elements are studied to understand the anti-subordination perspective (stereotypes, non-neutrality, decontextualization and monoculture) that show that the same criticism made from the anti-subordination perspective to the Law can be directed to artificial intelligence (as both are areas in which power is concentrated). Finally, we propose some anti-subordination measures aimed at confronting algorithmic discrimination.

KEYWORDS

Anti-subordination, artificial intelligence, Law, oppression, algorithm.

DOI: https://doi.org/10.36151/TD.2024.112

ANTISUBORDINACIÓN Y DISCRIMINACIÓN ALGORÍTMICA

Anna Capellà i Ricart
Investigadora Postdoctoral
Institut de Dret i Tecnologia
Universitat Autònoma de Barcelona

1. INTRODUCCIÓN

El análisis masivo de datos —y, específicamente, la elaboración de perfiles— se basa en el descubrimiento de correlaciones entre datos alojados en bases y su finalidad es identificar, representar y clasificar a personas como miembros de grupos o categorías. A cada persona clasificada en un grupo o una categoría se le puede aplicar un perfil de grupo (es decir, ciertas características que el algoritmo considera probable que sean comunes en las personas que lo integran). Dado que la disponibilidad de datos relevantes crece de manera exponencial, cuanto más sofisticado sea el perfil del grupo, tanto más se inclinará el algoritmo a modelar un perfil personalizado y más sutilmente discriminará entre las personas que forman parte del grupo y las que no pertenecen a él (Schreurs *et al.,* 2008: 242).

Si el análisis masivo de datos es el conjunto de técnicas y tecnologías que permiten explorar y explotar grandes cantidades de datos de forma automática y semiautomática con el objetivo de encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los sujetos en un determinado contexto (Rivas Vallejo, 2020: 103), siempre que los datos estén ligados a comportamientos sociales (basados el estructuras de opresión, subordinación y exclusión) los resultados que se obtengan a partir de ellos serán discriminatorios. Las tareas que realizan los sistemas algorítmicos suponen la aplicación de estos patrones repetitivos, tendencias o reglas a conjuntos de datos para extraer resultados.

En la medida en que estos patrones, tendencias o reglas tomen como referencia contextos sociales determinados, los sistemas algorítmicos tienderán a reproducirlos. Por lo tanto, es muy probable que el uso de sistemas algorítmicos predictivos en un contexto social determinado implique la obtención de resultados discriminatorios, dado que los resultados simplemente reflejarán una sociedad desigual y discriminatoria que oprime a determinados grupos de población que comparten ciertos atributos protegidos.

Para hacer frente a la discriminación que genera la utilización de sistemas algorítmicos, se pretende «reparar» estos sistemas mediante métricas predefinidas de «equidad algorítmica», término técnico utilizado en la comunidad de aprendizaje automático para describir investigaciones y aplicaciones que tratan de abordar la discriminación en sistemas algorítmicos (West, 2020: 4).

En este trabajo ponemos en cuestión la capacidad de las herramientas de equidad algorítmica presentadas hasta ahora —mayoritariamente basadas en la equidad formal— para hacer frente a la discriminación. En la sección 2 presentamos la perspectiva anticlasificatoria (o formal) del Derecho antidiscriminatorio y la perspectiva antisubordinatoria (o sustantiva) del Derecho antidiscriminatorio. Seguidamente, la (sección 3), analizamos cuatro elementos para entender la perspectiva antisubordinatoria: los estereotipos, la no neutralidad, la descontextualización y la monocultura; del estudio de cada uno de estos elementos extraemos la crítica que la perspectiva antisubordinatoria ha formulado al abordaje de la discriminación por parte del Derecho y la extrapolamos al tratamiento de la discriminación generada por la inteligencia artificial (en adelante, IA). En la sección 4 proponemos algunas medidas orientadas a entender la discriminación producida por los sistemas de la IA desde la perspectiva antisubordinatoria. Finalmente, en la sección 5 presentamos unas reflexiones conclusivas.

2. ANTICLASIFICACIÓN Y ANTISUBORDINACIÓN

La cláusula específica de no discriminación incorporada a los textos constitucionales y a los instrumentos internacionales de derechos humanos pretende proteger a determinados grupos o colectivos que, debido a la minusvaloración, la subordinación y la postergación social que han padecido a lo largo de la historia, no han participado, de hecho, en los procesos de formación de las normas (Añón Roig, 2013: 131). Sin embargo, la interpretación dominante del principio de igualdad en el en el campo del Derecho reenvía a la idea de «igualdad de todos», es decir, a la atribución del mismo estatus jurídico a todas las personas, y no reconoce la existencia de grupos subordinados. Semejante aproximación a la igualdad jurídica y política está basada en un falso universalismo, dado que en el ámbito jurídico la abstracción jurídica del «individuo» o «sujeto» del que se predica la igualdad ha sido edificada y desarrollada tomando como referencia un tipo *específico* de sujeto (hombre, blanco, propietario…) (Barrère Unzueta, 2008: 54). El principio normativo de acuerdo con el cual todas las personas tienen el mismo estatus jurídico (y, por tanto, deben ser consideradas

iguales por el Derecho) presume la neutralidad del Estado y, por ello, no considera las diferencias entre las personas en la ley ni en la aplicación de la ley. Desde el momento en que las personas ya tienen reconocidos los mismos derechos, los mecanismos de protección previstos en el ordenamiento jurídico se limitan a evitar que el ejercicio abusivo (pero individualizado) de los derechos de unas no interfiera en los derechos (individuales) de otras. En definitiva, la idea de justicia se traduce en la igualdad del estatus jurídico individual (o, si se prefiere, en la indiferenciación tanto en la ley como en la aplicación formal de la ley) (Barrère Unzueta, 2008: 54-55).

La función del Derecho antidiscriminatorio se ha interpretado desde dos perspectivas distintas. Por un lado, la corriente anticlasificación sostiene que el Derecho antidiscriminatorio debe garantizar la igualdad formal entre los individuos, considerando que cualquier trato desigual solo resulta admisible si es razozable y no arbitrario poner la atención en que cualquier trato es igual cuando es razonable y no arbitrario, rechazando cualquier trato diferenciado entre personas que no guarde relación funcional con la finalidad perseguida. Desde este enfoque, la discriminación se asocia como ruptura de la igualdad de trato en términos básicamente individuales, focalizándose en la igualdad formal y asumiendo una visión litigante (Añón Roig, 2022: 41).

Por ejemplo, desde el punto de vista de la privacidad, la anticlasificación prohíbe cualquier diferenciación basada en un rasgo o motivo protegido. El razonamiento que subyace a esta concepción es que, si puede prevenirse el acceso a cierta información especialmente protegida, en el futuro se evitará la discriminación basada esta información protegida (Roberts, 2015: 2123-2124). Sin embargo, algunos autores han puesto de manifiesto que la perspectiva anticlasificación puede no funcionar en los sistemas algorítmicos (Barocas *et al.,* 2019: 199). Para evitar la clasificación de las personas a partir de su pertenencia a uno o varios grupos en los procesos de tratamiento masivo de dados, el modelo algorítmico debería ignorar o suprimir intencionadamente todos los atributos protegidos y todos los factores correlacionados con estos. Esta acción generaría un modelo con resultados mucho menos precisos (Barocas y Selbst, 2016: 727). Además, cabe poner en cuestión la posibilidad de desarrollar este enfoque, dado que en muchas ocasiones no se puede determinar qué datos correlacionados acabarán suministrando información que la legislación considera protegida.

Desde un punto de vista crítico, se ha señalado que la perspectiva anticlasificación ignora la historia de minusvaloración y opresión que han sufrido ciertos grupos de población (Barrère Unzueta, 2018: 13). Esta aproximación «ciega» es la que se sirve de base para hacer frente a la discriminación algorítmica. Como señala Green (2022: 4), la lucha contra la discriminación en el ámbito de los sistemas algorítmicos restringe el margen de apreciación de la información a momentos decisorios concretos —datos de entrada y salida de la decisión en cuestión—, es decir, desvincula los datos del contexto en el que han sido extraídos y se limita a asegurar que el resultado del tratamiento no sea formalmente discriminatorio, pero no tiene en cuenta por qué razón los datos son unos y no otros; esta desatención, asociada a la definición de la equidad como un mero atributo técnico de los algoritmos, permite ignorar abiertamente la historia de minusvaloración y opresión que han sufrido ciertos grupos

de población (Green, 2022: 12). Por su parte, West (2020: 59) observa que, más allá de que los discursos son cada vez más interdisciplinares, para hacer frente a la discriminación algorítmica se ha consolidado un corriente dominante centrada en los marcos computacionales del problema y en el desarrollo de herramientas diseñadas para identificar y mitigar sesgos en los sistemas algorítmicos —por ejemplo, corregir el sesgo estadístico en conjuntos de datos para que se alcancen tasas de precisión en todas las categorías demográficas—. Este enfoque dominante parte de la premisa de que hay problemas en el funcionamiento de los sistemas de IA que pueden ser solucionados e incluso regulados con el objetivo de legitimar la venta de los bienes y servicios de la IA, pero omite señalar que este tipo de problemas de funcionamiento viene causado por un conflicto estructural y dinámico que no puede solventarse tecnológicamente. Como acabamos de sugerir, el reconocimiento de este déficit podría dificultar la comercialización de bienes y servicios de IA para su utilización en contextos donde pueden producirse situaciones de discriminación. Adicionalmente, este planteamiento se aleja de la concepción según la cual, si bien los algoritmos pueden remediar problemas sociales, en muchos casos son herramientas innecesarias o incluso perjudiciales (Green, 2022: 18).

Por su parte, la corriente antisubordinación sostiene que la principal finalidad del Derecho antidiscriminatorio es hacer frente a la exclusión social, la opresión y la subordinación de ciertos grupos de población y que, en este sentido, debe articularse en torno a la noción de igualdad entendida como no sometimiento o no exclusión. Esta perspectiva considera especialmente relevante la existencia de personas y grupos de personas que, al haber sido histórica y sistemáticamente excluidas de distintos ámbitos de la vida social, y haberse visto subordinadas y oprimidas por determinadas estructuras sociales de poder y dominación, rompen el esquema basado en la presunta igualdad entre todas las personas (Añón Roig, 2022: 41 y 42), y propone expandir el ámbito de análisis a consideraciones relacionales y estructurales que trascienden la toma de decisiones en casos específicos (Green, 2022: 8 y 16). La ruptura de aquel esquema y la falta de atención a la dimensión estructural de la discriminación tiene implicaciones no menores, dado que los grupos subordinados ven limitadas sus oportunidades vitales y su acceso a los bienes y servicios básicos, y se encuentran, por ello, en una situación de desequilibrio en el desarrollo de sus relaciones sociales (Álvarez del Cuvillo, 2022: 189). La perspectiva antisubordinación centra su atención en la dimensión material de la igualdad y aboga por la eliminación de las desventajas que afectan los grupos oprimidos a través de la incorporación al ordenamiento jurídico de medidas de trato diferencial, la puesta en marcha de estrategias de protección de la diversidad y la implementación políticas que integren las particularidades de cada grupo en los instrumentos reguladores generales con el objetivo de reconducir y subsanar la historia de subordinación que han padecido aquellos grupos (Soriano Arnanz, 2021: 10 y 11). Se trata, pues, de un planteamiento crítico informado por una finalidad transformadora de las estructuras que reproducen la desigualdad y la discriminación cuyo arsenal argumentativo se orienta a poner en evidencia la insuficiencia y aun el agotamiento del modelo formalista/anticlasificatorio para abordar la discriminación (Morondo Taramundi, 2023: 713 y 714). La clave del estudio de la perspectiva antisubordinatoria radica en el análisis del modo en que las sociedades dan respuesta a las diferencias en los atributos y capacidades de las personas

pertenecientes a determinados colectivos (por ejemplo, el color de la piel, el sexo o la capacidad física), y su objetivo último es luchar para que todas sean tratadas con igual respeto con independencia de aquellos rasgos diferenciadores (Green, 2022: 8).ç

Esta concepción, el Derecho antidiscriminatorio no puede prescindir de la crítica a determinados planteamientos sustentados en una concepción del poder que se ha denominado grupal, sistémica y estructural, dado que, como hemos dicho, responde a la idea de opresión o «poder sobre» personas excluidas o subordinadas que, en su condición de tales, adquieren conciencia e identidad política (Barrère, 2018: 20). La perspectiva antisubordinación entiende, por tanto, que la discriminación no solo resulta perjudicial cuando afecta a una persona específica, sino también (o, sobre todo) cuando el sujeto pasivo es todo el colectivo social (o los colectivos sociales) a los que aquella pertenece, puesto que en tal caso el bien lesionado es la posición del colectivo en la comunidad (Barocas *et al.*, 2019: 79). Es otros términos, si bien los actos discriminatorios concretos son individualizables, en realidad solo son una consecuencia de opresiones sociales subyacentes (Añón Roig, 2021: 36). En el proceso de individualización —que inevitablemente comporta la ignorancia o el desprecio de la importancia del colectivo— la discriminación es abordada como un problema de las personas excluidas y no como un ejemplo que ilustra la existencia de opresiones no abordadas (Bodelón, 2010: 88). Los mecanismos reparadores centrados en situaciones concretas de discriminación perpetradas contra individuos fallan cuando pretenden asegurar la igualdad de tratamiento de los propios individuos en riesgo de discriminación sistémica, y también fracasan a la hora de proteger a otras personas que pueden sufrir daños colaterales derivados de las estructuras sociales que discriminan, de modo que están mal equipados para hacer frente al daño masivo y las prácticas sistémicas ilegales (Farkas, 2014: 26). El rol fundamental que corresponde al grupo en el Derecho antidiscriminatorio interpela a las estructuras jurídicas basadas en una concepción de la igualdad de corte liberal y basada en el individuo y pone de manifiesto sus limitaciones (Barrere Unzueta, 2001: 146; Morondo Taramundi, 2023: 715).

El enfoque de la perspectiva antisubordinación está orientado a identificar el modo en que los sistemas jurídicos plasman las desigualdades que tienen su razón de ser en las estructuras de poder capaces de ordenar las relaciones sociales para atribuir o asignar estatus (subordinados o inferiores, privilegiados o superiores) y, por tanto, para robustecer las dinámicas e inercias que reproducen estas relaciones de subordinación (Barrère, 2018: 32; Green, 2022: 9 y 10). Barocas *et al.,* afirman que la ley es un instrumento que puede emplearse tanto para crear las condiciones que propicien el florecimiento de la discriminación como para contrarrestarla. En este sentido, el Derecho es (puede ser) tanto una herramienta de subyugación como una herramienta de liberación (Barocas *et al.*, 2019: 143). Esta disyuntiva es aplicable a la inteligencia artificial: en efecto, la utilización de la IA puede perpetuar la discriminación, pero también puede contribuir a luchar contra ella. Ahora bien, lo relevante no son los usos (buenos) que hipotéticamente se le puedan dar tanto a los sistemas jurídicos como a la IA (que, por cierto, en el caso de la AI legitiman su utilización como herramientas de diferenciación para la igualdad con vistas a materializar un cambio altamente improbable que solo se producirá a través de transformaciones socia-

les mucho más profundas). La cuestión central es quién detenta el poder y cómo se están empleando actualmente estos instrumentos: la utilización de la IA en diversos contextos sociales evidencia una vocación clara de perpetuación del *statu quo* opresivo que aqueja a ciertos grupos de población.

A este respecto, West plantea que, desde sus inicios, la cultura de la IA ha estado dominada por los hombres, y que, al moldear el comportamiento relacional a través de las máquinas, estos tomaron como referencia una concepción particular de la inteligencia humana basada en la masculinidad adulta, blanca y de clase alta, de modo que la discriminación está se integró en el campo de la IA en el mismo momento de su fundación. A título ilustrativo, cabe hacer referencia al uso recurrente de personajes femeninos virtuales en los sistemas de asistencia telemática que replican los estereotipos de género e imprimen a su ejecutoria la docilidad y la disposición servil tradicionalmente atribuida a las mujeres (West, 2020: 9). En otro orden de cosas, la IA también promueve la servidumbre en lugar de la autonomía y la independencia.

En la actualidad, el Derecho antidiscriminatorio no se ha comprometido de manera clara con el principio de antisubordinación (y, por consiguiente, este principio tampoco se ha trasladado al ámbito de la toma de decisiones a través sistemas algorítmicos). Por ello, sigue dominando la perspectiva que concibe el principio de igualdad de trato como indiferenciación o como trato razonable no arbitrario (Añón Roig, 2022: 43-44).

3. ANTISUBORDINACIÓN E IA. ¿QUÉ ELEMENTOS HAY QUE TENER EN CUENTA?

En esta sección analizaremos el impacto de los estereotipos (3.1), la presunción de neutralidad (3.2), la descontextualización (3.3) y la promoción de una monocultura (3.4) en la discriminación derivada del uso de sistemas algorítmicos en contextos sociales. Además, pondremos en relación el impacto de estos elementos en los resultados que presenta la IA con la crítica al modo en que el Derecho antidiscriminatorio aborda habitualmente esta cuestión.

3.1. ESTEREOTIPOS

Un elemento clave para entender la discriminación son los estereotipos. Los estereotipos son visiones generalizadoras o preconcepciones de los atributos, características o roles que, se supone, deben cumplir los miembros de un colectivo social concreto y que contribuyen a reforzar la presunción de que cualquier persona que se integre en el colectivo social actuará de conformidad con aquellas visiones generales o preconcepciones (por ejemplo, las mujeres son cuidadoras por naturaleza) (Cook y Cusack, 2010: 11). Se trata de representaciones mentales (mecanismos cognitivos neutros normalmente de carácter inmutable) que nos ayudan a reducir la complejidad de la realidad y son interseccionales,

es decir, están constituidos simultáneamente por diferentes ejes, entre ellos el género, la raza, la clase social o la orientación sexual (Ghidoni y Morondo Taramundi, 2022: 62). A pesar de que se consideran neutros, su función, su uso o su explotación por parte de los sistemas de opresión puede tener impactos positivos o negativos. Esta caracterización de los estereotipos permite analizarlos como elementos sistémicos y difusos, así como alejar el análisis de la discriminación de algunos elementos clave, entre ellos los efectos que producen (como preeminentemente sucede en los casos de discriminación algorítmica) (Ghidoni y Morondo Taramundi, 2022: 41-43).

Los estereotipos forman parte del proceso de subordinación, que construye las identidades y el estatus de los grupos desaventajados y dota de significado al comportamiento social sin que exista ningún mecanismo especifico que aísle su influencia del razonamiento jurídico (Morondo Taramundi, 2023: 710, 718 y 726) o, en relación con la cuestión que nos ocupa, del desarrollo de la IA. Los estereotipos tienen un origen social (están estructurados socialmente como convenciones). Los sistemas jurídicos los absorben y los refuerzan a través de la producción, la interpretación y la aplicación del Derecho (Añón Roig, 2021: 45), y justifican, naturalizan e invisibilizan las estructuras de poder (por ejemplo, en las normas que prevén los permisos para madres –más largos y obligatorios– y para los padres —optativos y más cortos—) (Ghidoni y Morondo Taramundi, 2022: 46 y 49). Los estereotipos desempeñan un rol fundamental en la perpetuación de la discriminación y son difíciles de identificar y neutralizar a través de las categorías jurídicas (Morondo Taramundi, 2023: 722).

Los estereotipos tienen un impacto sobre la persona —dado que presumen en ella ciertos atributos, características y roles que la despersonalizan, impiden su libre individualización e influyen en la construcción de su identidad—, pero también sobre el grupo (o los grupos) al cual pertenece —pues los rasgos o roles atribuidos por el estereotipo al colectivo (ya sean positivos o negativos) determinarán su posición en la jerarquía social: a este respecto la cuestión clave es quién generaliza o puede generalizar y para qué generaliza— (Ghidoni y Morondo Taramundi, 2022: 57-60).

Añón Roig subraya que los estereotipos perpetúan relaciones de poder opresivas, excluyen a algunas personas del acceso a bienes básicos, y reducen la autonomía y el autorrespeto porque repercuten en dimensiones que afectan centralmente al control sobre la propia vida o a las decisiones básicas relacionadas con ella (Añón Roig, 2021: 48).

Centrándonos en la IA, cabe apuntar que, a pesar de que se ha presentado como una tecnología potente, omnipresente y casi infalible para resolver problemas (decisionales y de eficiencia), en esencia se limita a encontrar patrones en enormes cantidades de datos (Powles y Nissenbaum, 2018). Si bien algunos patrones que se encuentran en los datos de entrenamiento representan conocimientos adquiridos (fumar está asociado al cáncer), otros patrones representan estereotipos que operan socialmente (a las niñas les gusta el rosa y a los niños les gusta el azul). Los algoritmos de aprendizaje automático no disponen de herramientas capaces de distinguir estos dos tipos de patrones, de manera que extraerán estereotipos de la misma manera en que extraen conocimiento (Barocas *et al.,* 2019: 10). Según Ghidoni y Morondo Taramundi, los «sesgos sin prejuicio» de las máquinas han lle-

vado también a replantearse cómo ingresan y cómo operan los estereotipos en los procesos de decisión en general (Ghidoni y Morondo Taramundi, 2022: 40).

Los estereotipos también están presentes en la preparación previa para que el sistema de aprendizaje automático pueda llegar a detectar patrones. Por ejemplo, el etiquetaje de datos puede plantear conflictos cuando las personas que llevan a cabo esta tarea deben realizar elecciones difíciles para determinar qué etiquetes son más adecuadas (es decir, cada ejemplo puede cumplir ciertos criterios, pero no todos, para ser asignado a una etiqueta concreta). También puede ocurrir que el etiquetaje no sea lo suficientemente preciso como para capturar diferencias significativas entre los casos, circunstancia que motiva que las personas que hacen el etiquetaje realicen juicios de valor que influyen en la toma de decisiones posterior (Barocas y Selbst, 2016: 681). Un caso representativo puede ser el de ImageNet, una base de datos que funciona con la sistemática de clasificación de imágenes según nueve categorías (que después se desglosan en subcategorías): planta, formación geológica, objeto natural, deporte, artefacto, hongo, persona, animal y variado. A pesar de que algunos conceptos —por ejemplo, 'manzana'— puedan ser fáciles de catalogar, hay otros que plantean una mayor complejidad. Cuando nos referimos a «persona» se integran etiquetas como «persona sin habilidades» o «mulato» que integran imágenes permeadas por estereotipos de raza, genero, edad y habilidad. No hay, pues, categorías neutras en ImageNet, dado que la selección de las imágenes siempre interactúa con el significado de las palabras (Crawford, 2021: 138-147). Los creadores de ImageNet recurrieron a la plataforma Mechanical Turk (MTurk), de Amazon, un mercado laboral en línea que permite a los individuos y las corporaciones contratar trabajadores bajo demanda para realizar tareas simples (Barocas *et al.*, 2019: 238). A los trabajadores de MTurk se les daba la tarea de clasificar cincuenta imágenes (extraídas de navegadores web) por minuto en las categorías propuestas. En los resultados recogidos en la base de datos pueden encontrarse estereotipos, errores y contrasentidos absurdos, entre ellos que una persona que se encuentra en la playa sea etiquetada como «cleptómana» o que una persona adolescente que lleva una sudadera deportiva sea clasificada como «perdedora». A pesar de que se intenten suprimir o matizar los peores ejemplos, la perspectiva esta fundamentalmente construida con base en la separación entre los datos, las personas, los sitios y los contextos de donde provienen a fin de transmitir una visión del mundo técnica que pretende infundir una suerte de objetividad a materiales culturales complejos y de procedencia heterogénea (Crawford, 2021: 138-147).

También se ha alertado de que los sistemas de búsqueda y recomendación satisfacen las necesidades de algunos usuarios mejor que las de otros porque los perfiles de consumidores de los primeros son similares a los de las personas que han creado los sistemas. Las personas que crean los sistemas pueden privilegiar ciertos contenidos sobre otros y generar daños representacionales mediante la amplificación y la perpetuación de estereotipos culturales, dado que influyen en la selección de los mensajes más difundidos (Barocas *et al.*, 2019: 20 y 192). Adicionalmente, se ha observado que el lenguaje y las imágenes capturan tal variedad de estereotipos culturales que los métodos de eliminación de los sesgos algorítmicos no suprimen, en realidad, el sesgo, sino que simplemente lo esconden (Gonen y Goldberg, 2019: 5).

Los sistemas sociales opresivos (machismo, racismo, clasismo) generan estereotipos opresivos (por ejemplo, la madre abnegada, el inmigrante gorrón, el pobre holgazán), de manera que los estereotipos deben ser entendidos como evidencias de los mecanismos que pretenden invisibilizar (Ghidoni y Morondo Taramundi, 2022: 54). En este sentido, el hecho que los sistemas de IA aplicados a contextos sociales simplifiquen la realidad, la presenten desde el punto de vista de la cultura mayoritaria (cfr. apartado 3.4) y amplifiquen sus resultados puede facilitar la detección de los estereotipos que pretenden reforzar y perpetuar los sistemas de opresión (Wachter, 2021: 372).

Sobre todo, hay que tener en cuenta que el diseño de métodos para hacer frente a la discriminación algorítmica sin tener en cuenta los estereotipos que tiñen la sociedad no solo implica la inefectividad de aquellos dispositivos, sino también la participación de la IA en la perpetuación de la discriminación. En la construcción del concepto de «discriminación» ya se ha recorrido este camino, que ha llevado a fomentar la confusión sobre su significado y a propiciar que personas que no han sufrido una historia de opresión puedan considerarse discriminadas. Por ejemplo, como plantea Lousada Arochena, las leyes de igualdad de género clásicas se caracterizan (entre otras cosas) por el hecho de que se han construido tomando únicamente en consideración el concepto de sexo, es decir, como si los estereotipos de género no existieran. El efecto colateral de este enfoque legislativo ha sido la extensión de la protección contra la discriminación a los dos sexos y la creación del espejismo del varón discriminado (Lousada Arochena, 2022: 4).

3.2. ¿NEUTRALIDAD?

En este apartado pretendemos mostrar que las críticas a la supuesta neutralidad del Derecho son las mismas que pueden formularse a la supuesta neutralidad de la IA.

En línea con la teoría crítica del Derecho, Lousada Arochena sostiene que el Derecho no representa la razón universal, sino la razón de los hombres en cuanto a detentadores del poder. De esta manera, la perspectiva social y cultural de los hombres es la que ha servido de base para construido la totalidad de un ordenamiento jurídico que pretende aparentar neutralidad (Lousada Arochena, 2022: 6 y 7). Uno de los objetivos del análisis antisubordinatorio es identifica las diversas vías a través de la que se plasman las desigualdades en los sistemas jurídicos y los modos en que estos ayudan a asignar estatus y establecer dinámicas e inercias que reproducen las relaciones de subordinación. En otros términos, uno de los objetivos del Derecho antisubordinatorio es poner en entredicho la supuesta neutralidad del derecho (Barrère, 2018: 32).

En un principio, las decisiones tomadas por sistemas de IA se consideraban más objetivas —y por tanto más neutrales— que las decisiones tomadas por personas. La razón principal de la atribución de una supuesta mayor neutralidad a los sistemas de IA no era otra que la visión encomiástica, apologética y acrítica de la tecnología: dado que los resultados que arrojan las máquinas arrojan derivan del uso de mecanismos técnicos, son más objetivas que las personas. Sin embargo, cuando intervienen en procesos decisorios que influyen en las personas, las máquinas tienen los mismos sesgos que estas (pues han sido creadas por

personas). Es precisamente la tecnofilia la que sigue amparando la supuesta neutralidad de la IA y dificulta que esta pueda ser evaluada por la población. Birhane *et al.* (2022: 178) publicaron un estudio orientado a analizar una serie de artículos influyentes en la disciplina del aprendizaje automático; los resultados mostraron que tan solo un 1 % de los encuestados mencionó o discutió sus posibles efectos negativos. Los autores concluyen que la tecnología no es una disciplina neutral en relación con los valores que promueve, sino que tiene una carga social y política, dado que fomenta la concentración de recursos, herramientas, conocimiento y poder en manos de actores que ya son poderosos, y prioriza y operacionaliza valores como el desempeño, la generalización, la eficiencia y la novedad (Birhane *et al.,* 2022: 182). Además, el tecnicismo también establece distancias entre las personas y permite que las reacciones habituales de indignación ante la discriminación se diluyan. En este sentido, Bigman *et al.* (2023: 4-6) afirman que la discriminación algorítmica causa menos indignación moral que la discriminación humana y sugieren que esta actitud puede estar motivada por el hecho de que las personas tienen menos predisposición a atribuir motivaciones negativas a un algoritmo —ya que las personas son percibidas como seres capaces de intención y antipatía y, por tanto, es más probable que tengan una motivación perjudicial, mientras que los algoritmos no son percibidos como capaces de intención y antipatía y, por tanto, es más improbable que se atribuya a su resultado una motivación perjudicial—.

Green (2021: 250-253 y 259) plantea que las personas que trabajan en el análisis masivo de datos deben reconocerse como actores políticos implicados en las construcciones normativas de la sociedad. En autor compara el impacto de los analistas de datos actuales con el de los ingenieros que organizaron la circulación cuando apareció el coche; a pesar de que estos últimos consideraban que su intervención era meramente tecnológica (establecer las señales de tráfico y los sistemas de temporización de los semáforos), en realidad diseñaron todo el entramado circulatorio para privilegiar la ubicuidad del automóvil, que desde ese momento se adueñó de las calles. En este sentido, el hecho de orientar el análisis masivo de datos hacia nuevas aplicaciones es fundamentalmente antidemocrático, dado que permite a las personas que trabajan en este campo (mayoritariamente del mismo grupo social) dar forma a la sociedad sin deliberación ni responsabilidad y pone de manifiesto la ilusoria creencia de que el desarrollo científico requiere pocas responsabilidades morales o políticas y que se lleva a cabo con neutralidad, un apriorismo que omite el hecho de que resulta imposible llevar a cabo ninguna tarea sin estar influenciado por determinados antecedentes, valores e intereses. Los esfuerzos encaminados a favorecer la neutralidad no conforman una opción política neutra, sino más bien una opción fundamentalmente conservadora que perpetúa el *statu quo* y refleja una actitud aquiescente con los valores sociales y políticos dominantes.

Green (2021: 254 y 259) señala críticamente que, si bien las personas que se dedican al análisis masivo de datos reconocen que este no pude proporcionar soluciones perfectas a problemas sociales, generalmente dan por descontado que contribuyen al «bien social» y que esa aplicación de la tecnología es una estrategia adecuada para el progreso social. Sin embargo, considerados desde una perspectiva de la igualdad sustantiva y antiopresión,

muchos esfuerzos de la ciencia de datos para hacer el bien no lo están haciendo de manera consistente. Sería, por tanto, necesario desarrollar una metodología rigurosa que relacionara las intervenciones algorítmicas con los impactos sociales (*ibid.*: 255), sobre todo respecto a sus impactos sobre colectivos sociales determinados, dado que muchos actores políticos y empresas de tecnología se benefician del cumplimiento de estándares en relación con la equidad algorítmica formal sin hacer planteamientos políticos o económicos significativos (Green, 2022: 23). Por ejemplo, no basta desarrollar algoritmos reconociendo que los datos sobre los delitos están sesgados; es necesario, además, reconocer que las definiciones de delito, el conjunto de instituciones encargadas de sancionarlo y las intervenciones que esas instituciones brindan son el resultado de procesos políticos históricos cargados de discriminación (Green, 2021: 258).

Por otra parte, Barocas *et al.* (2019: 24) apuntan que focalizar la atención en el aseguramiento de la equidad algorítmica formal supone eclipsar los debates sobre la legitimidad de las decisiones tomadas mediante sistemas algorítmicos. Los autores afirman que muchas empresas han adoptado el discurso de la equidad y que les resulta relativamente fácil garantizar la paridad en las decisiones entre grupos demográficos sin abordar las preocupaciones de legitimidad. Precisamente, la incorporación de la perspectiva antisubordinatoria a este análisis podría evidenciar que la toma de decisiones no es legítima (en cuanto justa) porque el sistema algorítmico participa en la perpetuación de la situación de privilegio de ciertos grupos sociales frente a otros. En este sentido, resultan especialmente relevantes los planteamientos que apuntan que las intervenciones orientadas a abordar los perjuicios causados por sistemas algorítmicos deberían centrarse en las instituciones subyacentes, dado que, en realidad, la adopción de decisiones mediante sistemas automatizados permite y aun legitima la inacción de las instituciones e impide el impulso de reformas estructurales necesarias (Barocas *et al.*, 2019: 218). Además, el planteamiento anclado en la pertinencia de la equidad algorítmica formal implica la difusión de una narrativa que prioriza la resolución problemas sociales a través de soluciones técnicas (Powles y Nissenbaum, 2018) y asume que la eliminación de las dinámicas de opresión/subordinación entre grupos sociales es un objetivo asumible cuando la lucha contra la discriminación es constante, pero que nunca llegará a un punto culminante de resolución, dado que, intrínsecamente, es una lucha de fuerzas.

La lógica estándar de la ciencia de datos se basa en la precisión y la eficiencia, y tiende a trabajar en el marco de los parámetros de los sistemas existentes, a los que acepta sin cuestionamiento alguno (Green, 2021: 256). Es decir, se considera que las «buenas» decisiones son aquellas que dan respuestas precisas al objetivo dado. Ahora bien, si la atención se limita a la precisión, es casi inevitable que surjan problemáticas relativas a la legitimidad de las decisiones. ¿Es legítimo perfeccionar la precisión de un sistema de vigilancia masiva para que reconozca mejor a las mujeres y a las personas no blancas? El objetivo de que un sistema funcione con más precisión no es necesariamente correcto desde el punto de vista normativo (West, 2020: 3). En contraste, podrían considerarse «buenas» decisiones aquellas que se centraran en las cualidades de los sujetos o en la obtenión de resultados más inclusivos, que considerasen solo el conjunto completo de factores relevantes, que

incorporasen principios normativos (por ejemplo, los de necesidad o proporcionalidad), que permitiesen que las personas comprendieran el sistema y pudieran impugnarlo (Barocas, 2019: 33) o que prevean buenas protecciones procesales. Sin embargo, aunque resulta posible introducir protecciones procesales en los sistemas automatizados para justificar su legitimidad, ello implicaría socavar el ahorro de costes que la automatización pretende lograr (Barocas *et al.*, 2019: 43).

3.3. DESCONTEXTUALIZACIÓN

En este apartado pretendemos explicar, en primer lugar, que el hecho de desvincular la discriminación de su dimensión colectiva implica obviar también la historia de minusvaloración que han sufrido las personas perjudicadas, que el tratamiento de datos a través de la IA es un gran método para manejar información ignorando el contexto en el cual esta ha sido extraída (3.3.1). Además, estudiaremos el modo a través del cual la descontextualización del concepto de «discriminación» ha influido en el desarrollo de la «discriminación inversa» (3.3.2).

3.3.1. Tratamiento de la discriminación obviando la dimensión colectiva

Otro aspecto primordial para abordar la discriminación jurídicamente desde la perspectiva antisubordinatoria es el reconocimiento de que el colectivo oprimido debe situarse en el centro de la protección jurídica. El trato desigual o injusto es experimentado por personas individualmente consideradas, pero la razón de este trato es que se les atribuyen ciertas características, rasgos o prejuicios propios de una colectividad (Añón Roig, 2013: 134) independientemente de la heterogeneidad interna del grupo, y que la opresión hacia el grupo se traduce en situaciones de discriminación particulares (Barrère Unzueta, 2008: 60-62). La discriminación tiene una dimensión colectiva y grupal definitoria y no eliminable. No solo eso, sino que, tal y como señala Álvarez del Cuvillo (2022: 191), la prohibición de la discriminación basada en la dimensión colectiva parece ser la única perspectiva que se sostiene desde una interpretación histórica, sociológica, teleológica, sistemática e incluso literal del derecho positivo vigente. En suma, de acuerdo con la perspectiva antisubordinatoria, la discriminación es la manifestación individualizada de sistemas de opresión y dominación. La opresión es una condición de grupos y esta caracterización resulta fundamental (Barrère Unzueta y Morondo Taramundi, 2011: 20).

Sin embargo, el principio de igualdad ante la ley (la cláusula de igual protección) ignora determinados fenómenos estructurales que subyacen a las situaciones injusticia social. Dado que aquel principio está basado en la igualdad de trato individual, no solo deja sin resolver la parte estructural del conflicto, sino que incluso contribuye a perpetuar la injusticia social (Barrère Unzueta, 2008: 58). La disociación entre la normativa antidiscriminatoria y la lucha contra las desigualdades sistémicas que realmente operan en la sociedad constituye una distorsión manifiesta de su finalidad que puede acabar desactivando su operatividad, dado que una visión puramente individualista concibe las causas de discriminación enun-

ciadas en las normas antidiscriminatorias simplemente como clasificaciones abstractas (Álvarez del Cuvillo, 2022: 192).

Este proceso de abstracción invisibiliza las desigualdades realmente existentes en la sociedad, proyectándolas sobre un plano lógico-formal desconectado de la realidad material concreta. Semejante forma de proceder permite concebir la discriminación como una situación anómala, excepcional o patológica, ocultando así situaciones cotidianas que reproducen las relaciones de dominación y subordinación que acaban legitimando el orden de poder desigual (Álvarez del Cuvillo, 2022: 192 y 193). En esta línea, Barrère Unzueta y Morondo Taramundi (2011: 39) señalan que, desde aquel enfoque, la discriminación se percibe como un fenómeno excepcional o accidental que reclama una respuesta puntual del Derecho, cuando en realidad se trata de un fenómeno sistémico. Por su parte, Sebastiani (2021: 756) observa que, en relación con el racismo, si bien el debate hegemónico bascula entre la consideración de la discriminación racial como una consecuencia de prejuicios difundidos y su caracterización como un resultado de predisposiciones patológicas y extremistas, en realidad concibe los episodios racistas como actos «aislados» imputables a individuos concretos y desconectados de dinámicas sociales más profundas. Posiblemente, uso de la IA ha servido para evidenciar el carácter sistémico y difuso de la discriminación y para debilitar la idea de que la responsabilidad radica en un individuo en concreto, dificultando así el planteamiento de respuestas por parte del sistema jurídico, todavía deudor de una perspectiva centrada la responsabilidad individual.

Desde nuestra perspectiva, en el momento en que se pierde de vista el grupo, también se pierde de vista el contexto. La corriente anticlasificatoria desvincula completamente la decisión del contexto, es decir, ignora todas las dinámicas sociales —no solo históricas, sino también actuales— que determinan la previsión jurídica de un atributo protegido. En cuanto a la discriminación algorítmica, los elementos que están implicados en el trato equitativo (por ejemplo, la consideración de la existencia de la discriminación estructural) rara vez se tienen en cuenta en la toma de decisiones concretas, ya que aquellos solo tienen en cuenta causas relativamente próximas a los puntos de decisión (Barocas *et al.*, 2019: 185). Por ejemplo, cuando se toman decisiones con base en el expediente académico, no se tiene en cuenta el hecho de que las diferencias en los resultados a menudo surgen de la inestabilidad en el hogar y otras circunstancias sociales, económicas y de salud (Barocas *et al.,* 2019: 158). El tratamiento de datos se basa, por tanto, en la descontextualización, es decir, traduce conceptos complejos en simples datos descontextualizados. De la misma manera, las clasificaciones que se realizan mediante el análisis masivo de datos pretenden imponer orden en el desorden de la vida humana, pero al hacerlo ocultan la distribución desigual de sus impactos en las comunidades (West, 2020: 4).

La lucha contra los sistemas de opresión y las dinámicas de subordinación que operan en la sociedad exige ir más allá de las simplificaciones y analizar las sutilezas involucradas en las relaciones y las decisiones humanas. La IA hace lo contrario, puesto que el coste de tomar el mayor núnero de decisiones en el mínimo tiempo posible y con el mínimo coste (incremento del valor cuantitativo) es la simplificación de los factores considerados a la hora de tomar la decisión (disminución valor cualitativo). Podríamos decir, siguiendo a

Barocas *et al.* (2019: 25), que estos fatores ponen en relación el aprendizaje automático con la burocracia, que la IA sobresale en la simplificación de los procesos para automatizar las decisiones, a despecho de la pérdida cualitativa que ello conlleva. Indefectiblemente, el traslado de la perspectiva burocrática a todos los ámbitos de adopción de decisión implica indefectiblemente una pérdida de riqueza en estos procesos que afecta más intensamente a colectivos oprimidos, dado que su participación social se está más restringida (Ranchordás y Scarcella, 2022: 418). Además, lo que hace la IA es adoptar los sistemas de dominación/subordinación dominantes (patrones, reglas y normas) y extrapolarlos a todos los ámbitos de toma de decisión, ignorando que contextualizar es una tarea que las personas sí pueden llevar a cabo. Por ejemplo, las personas que viven cerca de una intersección pueden tener más probabilidades de sufrir accidentes porque aquella está mal diseñada y, por tanto, es peligrosa. Esta contingencia raramente podrá ser tenida en cuenta por un sistema algorítmico. El ahorro de costes que podría lograrse a través de la automatización de ciertas decisiones (sobre todo, reemplazando a los trabajadores por *software*) se hace a expensas de privar a las personas de la oportunidad de resaltar aspectos relevantes de la información que no tienen cabida en el proceso automatizado (Barocas *et al.*, 2019: 37). Adicionalmente, cuando el sistema en sí es injusto, las personas encargadas de implementarlo pueden ser una fuente importante de resistencia ante el incumplimiento o la denuncia de irregularidades (Barocas *et al.*, 2019: 217).

3.3.2. Especial mención a la discriminación inversa

El enfoque de la discriminación predominante en el CEDH (Convenio Europeo de Derechos Humanos) y la jurisprudencia del TEDH (Tribunal Europeo de Derechos Humanos) se sitúa claramente en la órbita del principio de igualdad formal. El tratamiento de la discriminación en la normativa antidiscriminatoria de la UE —así como en la jurisprudencia del TJUE (Tribunal de Justicia de la Unión Europea)— es especialmente ambiguo y complejo en lo que respecta a la integración de las perspectivas individual y colectiva: aunque reconoce expresamente la legitimidad de medidas como las acciones positivas, que solo pueden entenderse desde una perspectiva antisubordinatoria, tiende a identificar las causas de discriminación con las clasificaciones contrarias al principio general de igualdad de trato entendido en términos abstractos —una comprensión que, por tanto, permite la tutela simétrica—. De hecho, la jurisprudencia del TJUE admite abiertamente la posibilidad de discriminación inversa, dado que ha tomado en consideración a menudo ha atendido las reclamaciones basadas en la supuesta discriminación de miembros de grupos dominantes (*vid.,* por ejemplo, las sentencias *Eckhard Kalanke c. Freie Hansestadt Bremen*, de 17 de octubre de 1995 (C-450/93), y *WA c. Instituto Nacional de la Seguridad Social* (INSS), de 12 de diciembre de 2019 (C-450/18)) (Álvarez del Cuvillo, 2022: 195).

La admisión de la discriminación inversa se puede entendida como una consecuencia de la omisión del contexto en el que se produce la discriminación. Generalmente, la discriminación es caracterizada como una figura híbrida que abarca tanto una dimensión individual como una dimensión colectiva. A pesar de que existen algunos puntos de con-

vergencia entre ambas dimensiones, en realidad hacen referencia a paradigmas contradictorios y mutuamente excluyentes, lo que no impide que se apliquen en la práctica de forma combinada al precio de incurrir numerosas incoherencias y contradicciones lógicas. Una de estas contradicciones es el concepto de «discriminación inversa» que se aplica a aquellos tratos que generar un perjuicio a personas que pertenecen a grupos sociales dominantes o mayoritarios precisamente por su adscripción a estos colectivos. Se trata de una apropiación de la categoría jurídica de la discriminación llevada a cabo por los grupos dominantes que acaba diluyendo su potencial transformador y legitimando así el orden de poder desigual. En estos supuestos, la lucha contra la discriminación no solo deja de cumplir su función emancipadora, sino que, además, se convierte en un instrumento de dominación (Álvarez del Cuvillo, 2022: 190, 191 y 196). Barocas y Selbst (2016: 723-728) explican que en varios litigios de Estados Unidos de América los órganos de adjudicación han redefinido la distinción entre categorizaciones buenas y perjudiciales de raza reconduciéndola al principio formalista de anticlasificación; de este modo, los tribunales suprimen la función originaria del derecho antidiscriminación —proteger a miembros de grupos históricamente desaventajados— y facilitan que los hombres blancos puedan denunciar que han sufrido discriminación. Al decir de los autores, el fallo recaído en el caso *Ricci vs. DeStefano* 557 U.S. 557 (2009) constituye un claro ejemplo de esta renovada línea jurisprudencial.

La perspectiva antisubordinatoria está claramente confrontada con el individualismo imperante en la sociedad y sostiene que para cambiar los patrones históricos de subordinación es preciso cambiar dinámicas de privilegio que se perciben como «justas». Por ejemplo, que la acción positiva genera animadversión en grupos históricamente privilegiados porque consideran que no es «justa» o que están pagando ellos por opresiones históricas que ahora ya no existen. Si no hay un cambio en la percepción de lo que se considera justo (por ejemplo, la confrontación entre la simple meritocracia y la toma de decisiones teniendo en cuenta el contexto de la persona para conseguir esos méritos —que conlleva un análisis mucho más complejo—), las personas individualmente perjudicadas por las políticas antisubordinatorias grupales (por ejemplo, aquellas que no pueden acceder a un trabajo que consideran que deberían obtener porque es lo más justo desde la perspectiva de los grupos sociales privilegiados que históricamente han establecido los criterios a partir de los cuales se toman las decisiones), van a oponer una fuerte resistencia a la adopción e implementación de estas políticas, resistencia cuya razón de ser radica en su capacidad de abstraerse y desvincularse del contexto social opresivo cuando las decisiones les afectan individualmente.

La conducta discriminatoria está conformada por la confluencia de dos elementos constitutivos: por un lado, la pertenencia a uno o varios grupos o categorías sociales, y, por otro, el perjuicio individual y social provocado por el tratamiento derivado de esta adscripción, que lesiona la dignidad humana porque tiende a situar a determinados grupos en una posición socio-jurídica de inferioridad. Así pues, solo serían discriminatorios los tratamientos desiguales que estuvieran funcional o sistemáticamente orientados a la marginación de determinados grupos sociales y las personas que los integran. Para que una conducta sea considerada discriminatoria, se exige que pueda ser contemplada como la manifestación de

una pauta sistemática de degradación de los integrantes de un grupo que, de generalizarse o reproducirse, contribuye a situarlos en una posición de inferioridad social (Álvarez del Cuvillo, 2022: 190, 191 y 196). Este modelo teórico parte de la base de que todo acto de discriminación debe causar un perjuicio a determinados grupos humanos, aunque el perjuicio sea solo potencial (Álvarez del Cuvillo, 2022: 208). Desde el momento en que se admite que una norma puede tener un trasfondo discriminatorio, también se admite que no se trata de proteger solo a las personas discriminadas, sino también a los grupos o colectivos que sufren un perjuicio por sus condiciones personales o sociales (Rodríguez- Piñero y Bravo-Ferrer, 2022: 15).

3.4. MONOCULTURA ALGORÍTMICA

La búsqueda de la equidad algorítmica formal, entendida como la garantía de un trato igual y no arbitrario por parte de los sistemas algorítmicos, es un compromiso con las personas individualmente consideradas y no con los grupos sociales a los cuales pertenecen, ya que para adoptar un compromiso de equidad en relación con ciertos colectivos sociales subordinados debe existir un trato diferenciado en favor de estos. Green (2022: 3) plantea que es imposible que un algoritmo satisfaga todas las definiciones matemáticas en la adopción de una decisión equitativa.

En lugar de plantear la necesidad de dispensar un trato favorable a los grupos históricamente (y actualmente) subordinados, el camino que se está siguiendo es el de uniformizar las decisiones, dado que los sistemas de aprendizaje automático que se están generalizando son los de la cultura mayoritaria (basada en sistemas de opresión y subordinación), sistemas que generan altas tasas de error que perjudican a los grupos minoritarios (Barocas *et al.*, 2019: 12). Kleinberg y Raghavan (2021: 1) llaman a esta situación «monocultura algorítmica», noción que hace referencia al hecho de que las elecciones y preferencias devienen homogéneas cuando se ven implicados sistemas algorítmicos, ya que, si muchos sistemas de aprendizaje automático utilizan los mismos datos de entrenamiento y la misma variable objetivo, tenderán a generar las mismas clasificaciones, incluso si los algoritmos de aprendizaje son muy distintos. Por ejemplo, un pequeño número de empresas de medios sociales determinan en conjunto qué tipos de discurso pueden formar parte del discurso principal en línea y qué comunidades pueden movilizarse en línea (Barocas *et al.*, 2019: 215). Por otro lado, cuando se utiliza el mismo algoritmo o diferentes algoritmos basados en los mismos datos en múltiples contextos, una persona puede ser excluida arbitrariamente de una amplia gama de oportunidades (Creel y Hellman, 2021: 35 y 36).

Un ejemplo de interés en relación con el impacto de los sistemas algorítmicos en culturas minoritarias son los sistemas de reconocimiento de voz, que son menos precisos cuando se les habla en dialectos locales, circunstancia que motiva que cada vez sea más común ajustar la pronunciación y neutralizar el acento y la entonación, es decir, adoptar un comportamiento autocorrectivo para asegurar que el *software* de reconocimiento de voz recoja las palabras correspondientes integrando los sesgos del aprendizaje automático a la propia sociedad (Pasquinelli, 2022: 27).

Hay que tener presente que el hecho de haber ostentado históricamente (y actualmente) el poder de crear normas que rijan la sociedad implica que se hayan creado unos estándares con base en los cuales se medirá la igualdad; el proceso para garantizar la igualdad de derechos y oportunidades puede ser, por ello, confundido con un proceso para asimilar los usos y particularidades de los grupos sociales subordinados a los usos y particularidades de los grupos privilegiados. Mackinnon (1991:82) planteó que la vigencia del estándar de acuerdo con el cual todas las personas son iguales provoca que las mujeres se midan según la correspondencia con los hombres: por consiguiente, la neutralidad es simplemente el estándar masculino. Para apreciar la existencia de discriminación, es necesario que existan diferentes personas o grupos de personas que son tratados de forma diferente, y para probar su existencia se ha requerido un término comparativo, es decir, un referente comparativo que se erige como modelo de un trato privilegiado frente a otro. Esta concepción ha suscitado la crítica que apunta al «asimilacionismo», entendido como la absorción y convalidación implícita del modelo con el que se compara (Barrère, 2001: 151). Holtmatt (2010: 203) plantea que el hecho de no integrar el trato desigual a las mujeres cuando se las compara con los hombres —y asimilar, por tanto, las necesidades y preocupaciones de las mujeres a la de los hombres— deja intactos los patrones masculinos ya existentes. En la misma dirección apunta Green (2022:10) cuanto sostiene que, si ciertos colectivos sociales se enfrentan a oportunidades de desarrollo muy diferentes (que generalmente dependen de atributos que se distribuyen de manera desigual entre los grupos debido a la opresión, agravando las desventajas existentes), es imposible crear sistemas competitivos justos, de manera que el objetivo no debería limitarse a ayudar a algunas personas que forman parte de grupos sociales subordinados a que reciban decisiones favorables a través de un trato especial, aceptando como dada la estructura de oportunidades, sino cambiar la propia estructura de oportunidades. En este sentido, y para considerar métodos que favorezcan la aplicación de medidas contra la opresión en los sistemas algorítmicos hay que tener en cuenta que la pretensión de los grupos sociales subordinados es obtener el mismo respeto social, independientemente de sus capacidades y atributos, y no ser situados en los peldaños inferiores de una escala jerárquica. No obstante, esta pretensión no debe ser confundida con la voluntad de absorber y convalidar implícitamente las dinámicas, los roles y los estándares atribuidos al grupo privilegiado, que es el que promueve —en el caso de la IA, de manera exacerbada— unos estándares y una cultura concretos. Es decir, el objetivo es conseguir un respeto equivalente entre grupos sociales, no la asimilación de los roles, las actitudes y la cultura de los grupos sociales privilegiados que tradicionalmente han operado como estándares.

4. MEDIDAS ANTISUBORDINATORIAS A APLICAR

Siguiendo la perspectiva según la cual la lucha contra la discriminación es constante que nunca llegará a un punto culminante de resolución porque es una lucha de fuerzas intrínseca, proporcionar soluciones implica convencer a aquellas personas que forman parte

de los grupos sociales privilegiados de que las prerrogativas que entienden como «justas» y «no arbitrarias» únicamente lo son a la luz de sus estándares, así como de la necesidad de llegar a acuerdos sobre el modo de desarrollar formas de razonamiento para juzgar cuando y qué cantidad de discriminación es tolerable (Barocas y Selbst, 2016: 728). Hay que tener en cuenta, especialmente en un campo como el de la IA, en el que la inmediatez es tan relevante, que las medidas antisubordinatorias están concebidas para que tengan una incidencia a largo plazo, circunstancia que puede propiciar que se las considere poco atrayentes para abordar los problemas de discriminación que se plantean en sistemas algorítmicos concretos con base en los cuales se pretende obtener una rentabilidad coste-beneficio.

En relación con los datos, para paliar los problemas de equidad que presentan los resultados que generan los sistemas algorítmicos se ha planteado la posibilidad de recopilar más datos y realizar inversiones para mejorar la tecnología de clasificación a fin de de mitigar potencialmente la disparidad en la tasa de error (Barocas *et al.*, 2019: 98), o divulgar masivamente datos sociales que sustenten el desarrollo de la IA teniendo en cuenta que ya es tarde para hacer una reevaluación radical del modo en que las grandes empresas tecnológicas controlan las enormes cantidades de datos que se utilizan para desarrollar la IA (Powles y Nissenbaum, 2018). Estas medidas se plantean después de aceptar que los algoritmos de aprendizaje automático utilizan y han utilizado bases de datos masivos, que normalmente han obtenido los datos que revelan información sensible e íntima sobre las personas de manera ilícita o con base en un consentimiento no significativo (Green, 2021: 249). Por un lado, podría interpretarse que las normativas de protección de datos han intentado regular el uso exacerbado de datos para crear información útil con finalidades comerciales o de eficiencia en los procesos de decisión. Por otro, también se puede interpretar que la laxitud de las normativas de protección de datos ha permitido el desarrollo sin control de la IA. Al respecto, es interesante la propuesta de Jo y Gebru (2020), consistente en extraer lecciones de la archivística y la bibliotecología para documentar conjuntos de datos de aprendizaje automático prestando atención a cuestiones de consentimiento, inclusión, poder, transparencia, ética y privacidad.

En relación con la legitimidad, antes de utilizar la toma de decisiones mediante sistemas algorítmicos sería interesante plantear cuál será el impacto que tendrá en la persistencia y la magnitud de los sistemas sociales de opresión y subordinación. En este sentido, se ha presentado la idea de reconocer a determinados colectivos sociales —por ejemplo, comunidades geográficas— el derecho a consentir o rechazar colectivamente la adopción de herramientas tecnológicas (Barocas *et al.,* 2019: 219). Se plantea aquí un escenario en el que tanto las personas que deciden de qué manera se van a tomar las decisiones como las personas afectadas por estas decisiones disponen de margen para actuar. En muchos casos, si no en la mayoría, no va a ser así, de manera que no es descabellado plantear la prohibición normativa del uso de sistemas de IA en contextos sociales si perpetúan y agraven los sistemas de opresión y desigualdad, o que se someta a los sistemas de IA que se utilicen en el sector público —al menos— a un escrutinio exhaustivo y de prueba en relación con el impacto que pueda tener su uso en los colectivos desaventajados de la sociedad.

En relación con los estereotipos, Pou (2016: 175) plantea que para examinar una acción o norma presuntamente discriminatoria se incluya el razonamiento sobre si esta crea, perpetúa o agrava alguno de los tratos que mantienen cierto tipo distintivo de desventaja. En la misma dirección, Ghidoni y Morondo Taramundi (2022: 41 y 42) ofrecen ejemplos de medidas orientadas a reducir el impacto de los estereotipos en los resultados de los procesos decisorios (extrapolables al campo de la IA), a diversificar la composición social, de género, racial, de edad o clase de las personas que se ocupan del análisis de datos, a exponer a las personas que tienen que tomar decisiones a contra-narrativas e imágenes positivas de los grupos estereotipados. Asimismo, dan cuenta de diversas estrategias de sensibilización, entre ellas las técnicas que promueven la «autocorrección» en la toma de decisiones, que podría incluirse en los sistemas algorítmicos a través de procesos de revisión por parte de personas para discernir entre los casos en los que el sistema algorítmico ha encontrado patrones que representan conocimientos o los que representan estereotipos (nos remitimos a la ejemplificación que hemos presentado en el apartado 3.1). Por otra parte, las autoras sostienen que, si bien el Derecho está compuesto por normas abstractas (y generales) y funciona a través de categorías, esto no significa que no sea posible discutir sobre las formas de generalización del Derecho, pues este es un ámbito idóneo para introducir la discusión sobre los estereotipos (*ibid.*: 43). Aunque en el ámbito del Derecho —un campo más controlado en el que existen procesos de decisión pautados— la discusión puede ser fructífera, parece que esta posibilidad es más complicada en el ámbito de la IA, terreno en el que la generalización se produce de manera menos calculada y no hay tantos controles en los que puedan introducirse criterios orientados a luchar contra los estereotipos. Por tanto, es más útil sensibilizar directamente las personas que van a trabajar en el campo de la IA (por ejemplo, cuando cursan sus estudios universitarios).

Es necesario reevaluar la posición de los mecanismos de análisis masivo de datos en la sociedad evitando de la presunción de acuerdo con la cual el aprendizaje automático mejora los procesos decisorios en cualquier caso (Green, 2021: 257). También genera suspicacias el hecho de que se plantee como una alternativa para mejorar los procesos decisorios cuando, en el ámbito de las relaciones y las estructuras sociales, parece que lo que hace es perpetuar el *statu quo* opresivo a gran escala.

5. CONCLUSIONES

Actualmente predomina la perspectiva que concibe el principio de igualdad de trato como indiferenciación o trato razonable no arbitrario y esta preponderancia tiene consecuencias para el abordaje de la discriminación algorítmica. Hemos analizado cuatro elementos clave de la perspectiva antisubordinatoria que tienen efectos en las vías a través de las que se pretende mitigar la discriminación que producen los sistemas algorítmicos. En primer lugar, hemos mostrado que el hecho de no tener en cuenta los estereotipos que operan en la sociedad —que impregnan también el ámbito de la IA— no solo tiene como consecuencia que los métodos creados para hacer frente a la discriminación algorítmica

puedan ser inefectivos, sino que también que estos puedan contribuir a perpetuar la discriminación. En segundo lugar, hemos puesto en evidencia que la neutralidad que se atribuye tanto al Derecho como a la IA solo es aparente y tiende a promover la conformidad con los valores sociales y políticos dominantes. En tercer lugar, hemos analizado las implicaciones asociadas a la comprensión del Derecho antidiscriminatorio desde la perspectiva anticlasificatoria, entre ellas la abstracción de la realidad y la omisión del contexto histórico-social en el que se producen las discriminaciones. En el análisis masivo de datos también se produce esta descontextualización, pues aquellos se utilizan como información sin tener en cuenta de dónde proviene. El hecho de que en estas circunstancias no se contextualice adecuadamente puede contribuir a robustecer la idea de que la dimensión colectiva de la discriminación es poco relevante, y facilitar que se aprecien casos de discriminación en los cuales la persona perjudicada no pertenece a un colectivo socialmente subordinado. En cuarto lugar, hemos explicado que la toma de decisiones mediante sistemas algorítmicos (que, al difundirse, interiorizan estándares de la cultura mayoritaria) supone un incremento en la uniformidad de las elecciones y preferencias que tiene una incidencia directa negativa en la consideración y prevalencia de los usos y culturas de grupos subordinados.

Los elementos estudiados generan preocupaciones desde la perspectiva antisubordinatoria tanto en el ámbito del Derecho como en el de la IA, dado que son dos esferas en las que se concentra el poder y que han estado y están dominadas por colectivos sociales privilegiados. La perspectiva social y cultural de los hombres ha construido la totalidad de un ordenamiento jurídico que se pretende neutral, del mismo modo que la perspectiva social y cultural de los ha construido la totalidad de una «inteligencia» (énfasis en el entrecomillado) artificial que pretende aparentar neutralidad. Es importante tener en cuenta que los sistemas algorítmicos magnifican la discriminación que opera en la sociedad, ya que tienen un impacto potencial sobre muchas personas, y que las soluciones que se están proponiendo para abordar la discriminación que generan se limitan a los mismos mecanismos tradicionales que se han utilizado hasta ahora para hacer frente a la discriminación, mayoritariamente inspirados en el principio de igualdas formal. En la sección 4 hemos presentado algunas propuestas orientadas a evitar la perpetuación de las dinámicas de opresión/subordinación causadas por los sistemas algorítmicos, entre ellas la formación contra los estereotipos de las personas que se dedican al análisis de datos o la prohibición normativa del uso de sistemas de IA en contextos sociales en aquellos casos en los que perpetúen y agraven (o puedan perpetuar o agravar) la opresión y la desigualdad.

En cualquier caso, es necesario preguntarse si, en términos normativos, el objetivo es conseguir que los sistemas de IA se limiten a no incrementar la discriminación existente en la sociedad o si lo que se pretende es que estos sistemas también participen en su corrección. La respuesta a esta cuestión conlleva la adopción de diferentes medidas y diferentes perspectivas de análisis.

Desde nuestra perspectiva, la utilización de los sistemas de IA contribuye a poner de manifiesto que la discriminación es un problema complejo, acaso más complejo de lo que muestra su tratamiento jurídico por ello, están cobrando relevancia planteamientos hasta ahora más minoritarios. Si bien es cierto que ante una sofisticación tecnológica debe

adoptarse una sofisticación normativa y de análisis de los conflictos (por ejemplo, la anticlasificación está basada en un análisis mucho más simplista de las relaciones de poder y las dinámicas de opresión/subordinación que operan en la sociedad que la antisubordinación). Esta sofisticación no se ha conseguido aún en relación con el Derecho ¿Es posible que se consiga en el campo de la IA? Las medidas antisubordinatorias son aquellas que más lejos han llegado a la hora de conseguir una para los grupos sociales subordinados una renuncia de privilegios por parte de los grupos sociales favorecidos. Las medidas antisubordinatorias difícilmente pueden traducirse en conceptos técnicos y raramente pueden traducirse en resultados concretos a corto plazo que justifiquen la introducción (o no) en el mercado de un bien o servicio de la IA (contrariamente a las medidas anticlasificatorias: por ejemplo, conseguir que el sistema algorítmico ignore ciertos atributos protegidos). A pesar de que pueden revelarse como una buena estrategia para luchar contra la opresión, es necesario que buena parte de la población esté convencida de que su implementación es necesaria, efectiva y deseable.

BIBLIOGRAFÍA

ÁLVAREZ DEL CUVILLO, Antonio (2022): «El problema de la discriminación inversa: ¿Es posible discriminar a quienes pertene cen a los grupos sociales dominantes?», *Trabajo, Persona, Derecho, Mercado,* 5, 187-209.

AÑÓN ROIG, María José (2013): «Principio antidiscriminatorio y determinación de la desventaja», *Isonomía: Revista de Teoría y Filosofía del Derecho,* 39, 127-157.

— (2021): «Transformaciones en el derecho antidiscriminatorio: avances frente a la subordinación», *Revista Electrónica del Instituto de Investigaciones Jurídicas y Sociales Ambrosio Lucas Gioja*, 26, 29-54.

— (2022): «Desigualdades algorítmicas: conductas de alto riesgo para los derechos humanos», *Derechos y libertades,* 47, 17-49.

BARRÈRE UNZUETA, María Ángeles (2001): «Problemas del derecho antidiscriminatorio: subordinación versus discriminación y acción positiva versus igualdad de oportunidades», *Revista Vasca de Administración Pública. Herri-Arduralaritzako Euskal Aldizkaria*, 60, 145-166.

— (2008): «Iusfeminismo y derecho antidiscriminatorio: hacia la igualdad por la discriminación», en R. M. Mestre i Mestre (coord.), *Mujeres, derechos y ciudadanías,* Valencia: Tirant Lo Blanch, 45-71.

— (2018): «Filosofías del Derecho Positivo ¿Qué Derecho y qué discriminación? Una visión contrahegemónica del Derecho Antidiscriminatorio», *Anuario de Filosofía del Derecho,* XXXIV, 11-42.

BARRÈRE UNZUETA, María Ángeles y Dolores MORONDO TARAMUNDI (2011): «Subordinación y discriminación interseccional: elementos para una teoría del derecho antidiscriminatorio», *Anales de la Cátedra Francisco Suárez,* 45, 15-42.

BAROCAS, Solon *et al.* (2019): «Fairness and Machine Learning: Limitations and Opportunities» [en línea] <https://fairmlbook.org/>. [Consulta: 12/02/2024.]

BAROCAS, Solon y Anderw D. SELBST (2016): «Big Data's Disparate Impact», *California Law Review*, 104, 671-732.

BIGMAN, Yochanan E. *et al.,* «Algorithmic Discrimination Causes Less Moral Outrage Than Human Discrimination», *Journal of Experimental Psychology: General*, 152(1), 4-27.

BIRHANE, Abeba *et al. (2022):* «The values encoded in machine learning research», *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 173-184.

COOK, Rebecca J. y Simone CUSACK (2010): *Estereotipos de género. Perspectivas Legales Transnacionales*, Bogota: Profamilia.

CRAWFORD, Kate (2021): *Atlas of AI*, New Haven and London: Yale University Press.

CREEL, Kathleen y Deborah HELLMAN (2022): «The algorithmic leviathan: Arbitrariness, fairness, and opportunity in algorithmic decision-making systems», *Canadian Journal of Philosophy*, 52 (1), 26-43.

FARKAS, Lilla (2014): «Collective actions under European anti-discrimination law», *European Anti-Discrimination Law Review*, 19, 25-40.

GONEN, Hila y Yoav GOLDBERG (2019): «Lipstick on a Pig: Debiasing Methods Cover up Systematic Gender Biases in Word Embeddings But do not Remove Them» [en línea] <https://arxiv.org/pdf/1903.03862.pdf>. [Consulta: 20/02/2024.]

GREEN, Ben (2021): «Data Science as Political Action: Grounding Data Science in a Politics of Justice», *Journal of Social Computing*, 2(3), 249-265.

— (2022): «Escaping the Impossibility of Fairness: From Formal to Substantive Algorithmic Fairness», *Philosophy & Technology,* 35, 1-32.

GHIDONI, Elena y Dolores MORONDO TARAMUNDI (2022): «El papel de los estereotipos en las formas de la desigualdad compleja: algunos apuntes desde la teoría feminista del derecho antidiscriminatorio», *Discusiones,* 28, 37-70.

HOLTMAAT, Rikki (2010): «Equal Treatment to Equal Right», en E. Bodelón y D. Heim (coords.), *Law, Gender and Equality*, Barcelona: Universitat Autònoma de Barcelona, 209-228.

JO, Eun Seo y Timnit GEBRU (2020): «Lessons from archives: Strategies for collecting sociocultural data in machine learning», *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency,* 306-316.

KLEINBERG, Jon y Manish RAGHAVAN (2021): «Algorithmic monoculture and social welfare», *Proceedings of the National Academy of Sciences*, 118(22), 1-7.

LOUSADA AROCHENA, José Fernando (2022): «Evolución de la igualdad desde la constitución de 1978: del patriarcado fuerte hacia la igualdad de género», *IQUAL. Revista de género e igualdad,* 5, 1-27.

MACKINNON, Catherine (1991): «Difference and Dominance: on sex discrimination», en K. Bartlett y R. Kennedy (eds.), *Feminist legal theory: readings in law and gender*, Boulder: Westview Press, 81-94.

MORONDO TARAMUNDI, Dolores (2023): «Los estereotipos como mecanismos de desigualdad y alienación: un análisis desde el derecho antidiscriminatorio», *Oñati Socio-Legal Series,* 13(3), 710-729.

PASQUINELLI, Matteo (2022): «Cómo una màquina aprende y falla. Una gramàtica del error para la Inteligencia Artificial», *Revista Hipertextos,* 10(17), 13-29.

POU, Francisca (2016): «Estereotipos, daño dignitario y patrones sistémicos: la discriminación por edad y género en el mercado laboral», *Discusiones*, 16(1), 147-188.

POWLES, Julia y Helen NISSENBAUM (2018): «The seductive diversion of 'solving' bias in artificial intelligence» [en línea] <https://onezero.medium.com/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53>. [Consulta: 12/02/2024.]

RANCHORDÁS, Sofia y Luisa SCARCELLA (2022): «Automated Government for Vulnerable Citizens: Intermediating Rights», *William & Mary Bill of Rights Journal*, 30(2), 373-418.

RIVAS VALLEJO, Pilar (2020): *La aplicación de la Inteligencia Artificial del trabajo y su impacto discriminatorio*, Navarra: Aranzadi.

ROBERTS, Jessica L. (2015): «Protecting Privacy to prevent discrimination», *William and Mary Law Review,* 56(6), 2123-2124.

RODRÍGUEZ-PIÑERO y BRAVO-FERRER, Miguel (2022): «Los contornos de la discriminación», T*emas Laborales,* 162, 11-18.

SCHREURS, Wim *et al.* (2008): «Cogitas, Ergo Sum. The role of data protection Law and non-discrimination law in group profiling in the private sector», en M. Hidelbrandt y S. Gutwirth (eds.)*, Profiling the European citizen: Cross-Disciplinary Perspectives*, Berlin: Springer Dordrecht, 241-270.

SEBASTIANI, Luca (2021): «Investigando los límites de la lucha legal contra el racismo: El marco español de antidiscriminación por origen racial o étnico», *Oñati Socio-legal series*, 11(3), 753-786.

SORIANO ARNANZ, Alba (2021): «Decisiones automatizadas y discriminación: aproximación y propuestas generales», *Revista General de Derecho Administrativo*, 56, 1-45.
WACHTER, Sandra (2021): «How Fair AI Can Make Us Richer», *European Data Protection Law Review*, 7(3), 367-372.
WEST, Sarah Myers (2020): «Redistribution and Rekognition: A Feminist Critique of Algorithmic Fairness», *Catalyst: Feminism, Theory, Technoscience*, 6(2), 1-24.

Fecha de recepción: 3 de julio de 2024.
Fecha de aceptación: 20 de octubre de 2024.

RETOS DE LA FUNCIÓN JURISDICCIONAL PARA UN MUNDO INTERDEPENDIENTE Y ECODEPENDIENTE*

CHALLENGES OF THE JURISDICTIONAL FUNCTION FOR AN INTERDEPENDENT AND ECODEPENDENT WORLD

Elena Martínez García

Catedrática de Derecho procesal
Universitat de València

RESUMEN

En un mundo en el que las relaciones jurídicas son interdependientes y en el que dependemos de los recursos naturales para mantener nuestro nivel de producción y consumo, un nutrido grupo de juristas y científicos avalan la necesidad de poner ciertos límites a la libertad del mercado para proteger adecuadamente los derechos humanos y, por tanto, la democracia. No hacerlo conlleva privar de futuro a las próximas generaciones. Se abre así ante la justicia un nuevo elenco de intereses dignos de tutela y de legitimación. Previsiblemente, la utilización de la tecnología —especialmente, de la inteligencia artificial— propiciará la aceleración y la complejización de la conflictividad, y los tribunales se verán enfrentados a la necesidad de resolver las nuevas controversias en escenarios globales, multinivel y, a menudo, anómicos. Los derechos de solidaridad y la diligencia debida de los Estados aparecen como conceptos clave en esta nueva era.

PALABRAS CLAVE

Retos para la Justicia, ecodependencia, anomia, solidaridad, diligencia debida.

ABSTRACT

In a world where legal relations are interdependent and were we depend on natural resources to maintain our level of production and consumption, a large group of jurists and scientists support the need to set certain limits to market freedom to adequately protect human rights and, therefore, democracy. Failure to do so means depriving future generations of a future. This opens a new range of interest's worthy of protection and legitimacy before Justice. The use of technology —especially artificial intelligence— will predictably lead to the acceleration and complexity of conflicts, and the courts will be faced with the need to resolve new disputes in global, multilevel and often anomic scenarios. Solidarity rights and due diligence of States appear as key concepts in this new era.

KEYWORDS

Challenges for Justice, ecodependece, anomie, solidarity, due diligence.

DOI: https://doi.org/10.36151/TD.2024.113

* Este artículo ha sido realizado en el Marco del Proyecto Prometeo (2023) 64 "JUSTSOS".

RETOS DE LA FUNCIÓN JURISDICCIONAL PARA UN MUNDO INTERDEPENDIENTE Y ECODEPENDIENTE

Elena Martínez García
Catedrática de Derecho procesal
Universitat de València

1. ¿QUÉ ES TUTELAR CON JUSTICIA EN EL SIGLO XXI?

El mundo está cambiando a una velocidad de vértigo y, como diría Tony Judt (2011), «algo va mal». Nada es lo que parece y ninguna cuestión relevante merece un análisis unidireccional, basado en compartimentos estancos. Bauman (2011) nos hizo ver que la liquidez constituye el mayor riesgo para la seguridad jurídica, Han (2017) ha puesto en evidencia que en la sociedad de la desinformación reina un desencanto inducido que incita a la inacción (Han, 2017), Innerarity (2023) recuerda que la libertad solo puede ser libertad si es democrática, y Ferrajoli (2022) invita a la cultura jurídica a trabajar holísticamente para garantizar los derechos sociales, parte esencial, sostiene el autor, de los derechos humanos en un contexto global capaz de soñar con la materialización de una Constitución de la Tierra. Ahora más que nunca, el Derecho tiene ante sí el reto de responder a las transformaciones sistémicas desde una perspectiva ecosocial para garantizar nuestro futuro

como especie (Rockström y Lade, 2023). Este desafío implica a los poderes ejecutivo, legislativo y judicial, pero también a todas y cada una de las personas que formamos parte la sociedad como integrantes de un ecosistema que se pretende sostenible, y alcanza también al tejido empresarial y asociativo y a propia la ciudadanía. Ante estos desafíos, el trabajo colaborativo orientado a robustecer unos valores compartidos constituye el acicate para que impulsemos con emoción el reemplazo del «cansancio» y la «inacción» a los que alude Han. Aquellos valores ya están positivados: son los derechos humanos de los tratados y convenios internacionales de los que España forma parte (art.10.2 CE). Por tanto, no se trata solo de cuestiones éticas o ideológicas, sino de normas de Derecho positivo vinculantes *erga omnes* y muy a menudo olvidadas por nuestra jurisprudencia y los operadores jurídicos que la crean.

La Agencia de los Derechos Fundamentales de la Unión Europea (FRA, por sus siglas en inglés) alerta de que el mercado global necesitará nuevas reglas y limitaciones si queremos seguir hablando en serio del respeto a los derechos fundamentales (Sánchez Ramos, 2023). De hecho, la FRA propone la revisión de ciertos elementos clave del nuevo entramado de relaciones e interacciones entre el mercado-tecnología y la justicia social. Como procesalista, me corresponde analizar las tensiones y los obstáculos a los que puede verse enfrentado el ejercicio de la función jurisdiccional en un contexto global caracterizado por la ausencia de fronteras para los negocios jurídicos, la profusión de lagunas normativas atractivas para la criminalidad organizada (o sin organizar), la ampliación de los espacios sustraídos a la acción de la Justicia, la opacidad de las redes digitales y la generación de sesgos asociada a la aplicación de las nuevas tecnologías, particularmente la inteligencia artificial. Semejante panorama constituye un caldo de cultivo para que el resultado del análisis coste beneficio se incline en favor de la generación de conflictos en marcos permeados por posibles anomias y antinomias —usando la terminología del profesor Ferrajoli— que, si bien son resueltas en los marcos nacionales, tienen claras repercusiones internacionales. Es probable, por ende, que en los próximos años cambien la naturaleza de los conflictos, el perfil de las acciones y las legitimaciones, así como las tipologías procesales, y que aquella inflexión se oriente a asegurar que la respuesta jurisdiccional esté a la altura de las nuevas circunstancias interdependientes que desbordan el ejercicio de la potestad jurisdiccional en un solo país. Pero los cambios en curso también afectarán (y mucho) a los poderes legislativo y ejecutivo, de modo que también es esperable que se produzcan modificaciones de calado en la producción y la aplicación del Derecho para abordar la nueva realidad. Veamos una aproximación a estas ideas desde la perspectiva de nuestra disciplina.

Traigo a colación las palabras del profesor Moreno Catena sobre la obligación de la judicatura de tutelar con justicia. «Tutelar» con justicia no necesariamente significa que la interpretación y aplicación del Derecho deben siempre ser justas —o, al menos, parecérnoslo—, sino más bien que las decisiones judiciales que derivan de ambos procesos estén revestidas de mínimo de justicia para garantizar el contenido esencial del derecho a la tutela judicial efectiva». Como afirma el autor citado, la función jurisdiccional no solo debe solventar el conflicto en el caso concreto, sino también «[…] enviar señales que prevengan el nacimiento de otros nuevos conflictos en el futuro» (Morena Catena, [11]2021). «Los

conflictos —sostiene el autor— nacen porque a alguien le ha compensado provocarlos. En otras palabras, y con matizaciones, alguien pensó que el beneficio esperado de ignorar o apropiarse de un derecho, incumplir un contrato, no adoptar precauciones y causar daños o, en general, de infringir una norma era superior al coste de hacerlo. El concepto de valor esperado del Derecho (lo que dicen las normas, multiplicado por la probabilidad de su exigibilidad efectiva) es esencial para entender su funcionamiento y su efecto sobre las conductas de sus destinatarios» (*ibid.*). En una sociedad *glocal*, la dualidad de intereses entre los derechos violados en el caso concreto y el mensaje de seguridad que debe expandirse en la sociedad cobran especial importancia a nivel global; en un mundo globalizado, las injusticias de aquí tienen amplio eco en el otro lado del planeta. En este sentido, las soluciones deben adoptar un perfil global, pues el coste de renunciar a ello no es otro que la posible multiplicación exponencial del caos.

2. NUEVOS DERECHOS, NUEVOS SUJETOS, NUEVAS DUDAS Y UNA PROHIBICIÓN DE *NON LIQUET* EN UN MUNDO INTERDEPENDIENTE Y GLOBALIZADO

El paradigma político, social y cultural está cambiando. Recientemente, ha empezado a hablarse de la existencia de bienes demaniales que pertenecen a toda la humanidad, hemos visto que el legislador ha dotado de personalidad jurídica al Mar Menor (una concreción del reconocimiento de derechos a la naturaleza y los ecosistemas), que los animales no humanos han son considerados sujetos titulares de derechos en algunos ordenamientos, que la legislación ha intensificado la protección que brinda a las personas mayores debido a que conforman un colectiva en situación de especial vulnerabilidad, que la exigibilidad los derechos sociales se ha abierto paso en el campo jurisprudencial a través de condenas que imponen a los Estados el cumplimiento de obligaciones en la materia, que las futuras generaciones tienen intereses concretos y que las decisiones que tomemos hoy modularán los derechos de los que aquellas disfruten mañana. Estas y otras muchas tendencias de cambio desbordan los parámetros tradicionales del Derecho procesal y la función jurisdiccional, marcados fuertemente por la globalización y por esa la *lex mercatoria* invisible que en no pocas ocasiones constriñe los márgenes decisionales de los Estados y sus gobiernos. También la filosofía del Derecho ha teorizado en términos propositivos una nueva etapa en la evolución de los derechos humanos concretada en la posible transición desde los derechos sociales (entendidos como derechos individuales) hacia los denominados derechos de solidaridad (un concepto abarcativo y global que requiere renuncias y autocontención para asegurar los propios derechos sociales) (De Lucas, 1993). En este marco de transformaciones, la presente investigación adopta como marco de análisis una incipiente perspectiva ecosocial[1] de la vida y de la conflictividad que parte de la premisa de que el ecosistema planetario y los recursos naturales de nuestro planeta son bienes limitados. En esta línea, somos ya conscientes de que el modelo productivo extractivista que mantiene nuestro modo de vida resulta simplemente insostenible (Puleo, 2099: 69-72). La toma de

conciencia de nuestra condición ecodependiente e interdependiente debería detonar la transición hacia un modelo económico que incorpore como valor[2] no solo la persecución del beneficio y la competitividad, sino también la calidad de vida de las personas. Se trata, por tanto, de ampliar el catálogo de valores y de integrarlos en un modelo *biocéntrico*[3], dado que nuestra vida cotidiana influye decisivamente en el tratamiento que damos al planeta. Este parámetro de cambio se sostiene sobre la premisa del reconocimiento de que hay bienes de naturaleza común y no privativa, es decir, *bienes demaniales* (Ferrajoli, 2022) que pertenen a la humanidad, y que, al igual que los derechos fundamentales, constituyen límites y vínculos impuestos a los poderes públicos y a la propia ciudadanía. Así lo ha sostenido Naciones Unidas (De Luis, 2020). Hablamos de un cambio cultural, económico, político capaz de dotar de sentido real a la democracia y a los sus derechos y libertades (Ferrajoli, 2014; Innerarity, 2023) o, si se quiere, de reescribir el «contrato social» para incorporar en su articulado implícito una vertiente transformadora de todas las relaciones sociales, jurídicas y políticas que incluya la naturaleza biocéntrica o ecocéntrica. A nivel jurídico, este nuevo «contrato social» se está fraguando entre un nutrido grupo conformado de juristas a través de lo que Ferrajoli ha denominado «Una Constitución de la tierra», texto llamado a recepcionar en su articulado los valores, derechos y principios protectores de la vida de las personas y del planeta desde el prisma supranacional, y a reflejar la transición al paradigma ecológico[4].

Un ejemplo de la aplicación de la filosofía de este movimiento ha sido la reciente aprobación de la Ley 19/2022, de 30 de septiembre, para el reconocimiento de la personalidad jurídica a la laguna del Mar Menor y su cuenca, aprobada por iniciativa legislativa popular, que ha convertido en sujeto titular de derechos subjetivos al planeta, una iniciativa indudablemente novedosa en nuestra legislación (Martínez García, 2023). Lo mismo ocurre con las últimas legislaciones en materia de protección de animales (Cuerda, 2021)[5]. O las regulaciones en favor de colectivos vulnerables, entre ellos ancianos, los menores, personas discapacitadas o discriminadas por su orientación sexual o de género, cuyos derechos y marcos normativos se han ganado siempre en la calle, es decir, a través de las reivindicaciones populares frente al *statu quo* del poder.

La judicatura no ha sido ajena a esta nueva conflictividad, entre otras razones porque debe resolver controversias encuadradas en ella. La carrera de los juzgadores se torna cada vez más exigente porque tienen que conocer la complejidad que subyace a conflictos aparentemente lineales y sencillos en contextos supranacionales de averiguación de los hechos, y para cuya resolución los instrumentes existentes son vagos, técnicamente deficitarios en términos jurídicos, a menudo pobres… La libertad como pretexto, parafraseando al profesor Tomas Vives, la libertad como pretexto no puede ser un límite a la hora de generar condiciones materiales de igualdad (Vives Antón, 1995). Todo lo contrario, la libertad tiene límites y estos vienen impuestos por el resto de los derechos fundamentales y los derechos humanos. No es extraño, entonces, que la igualdad del art. 9.2 CE, que impone a los poderes públicos la remoción de los obstáculos (derivados de la libertad) que impiden que la libertad e igualdad del individuo y los grupos sean reales y efectivas, precepto que

constituye el fundamento de las nuevas acciones positivas u obligaciones de hacer exigibles a los poderes públicos.

La importancia de asumir el poder transformador de las sentencias en un mundo lleno de retos «glocales» (globales y locales al tiempo) (Guterres, 2020) resulta evidente. Es sabido que las personas que tienen encomendado el ejercicio de la función jurisdiccional en régimen de exclusividad (casi) siempre van por delante en el abordaje de la realidad. Para los titulares del Poder Judicial es difícil solucionar conflictos que afectan a las personas cuando no disponen de normas concretas aplicables al caso sometido a su enjuiciamiento. Precisamente por ello, son titulares de un poder del Estado y está sujetos a tan alta responsabilidad. Previsiblemente, Dado que estos operadores jurídicos van a enfrentarse en el futuro inmediato a una realidad jurídica y social muy cambiante, presumiblemente serán testigos privilegiados de la modificación de las bases de la teoría del Derecho tal y como han sido concebidas y estudiadas hasta ahora. Este reto solo acaba de empezar (Guterres, 2020). Pensemos que «[…] tras la Justicia y la función jurisdiccional ya no hay nada más para dar esa tutela de derechos a la ciudadanía y dicha potestad es el medio para fiscalizar la actividad administrativa, desde la imparcialidad objetiva y subjetiva de los jueces y magistrados, individualizada en el caso concreto para tener sentido y resolver conflictos […] Se crea derecho en un sentido auténtico de la expresión, (porque) la sentencia es la norma jurídica individualizada» (Eibar, 2023: 54-55), de acuerdo con la Constitución, las leyes, los principios generales del derecho (art. 1.4 CC) y el contexto histórico y social en el que vivimos (art. 3 Código Civil), y en el marco jurídico internacional al que nos hemos adherido (arts. 93 y 10.2 CE), conformado, entre otras muchas normas, por la Carta de los Derechos Fundamentales de la Unión Europea, la Declaración Universal de los Derechos Humanos o el Pacto Internacional de Derechos Económicos, Sociales y Culturales, que son el anclaje al que debemos sujetarnos para resolver las nuevas controversias de manera no simplista y con una visión transversal y multinivel de los problemas. Por todo lo dicho hasta aquí, no debe extrañar que, en un mundo globalizado, la prohibición de *non liquet* (art. 1.7 CC) adquiera un renovado protagonismo.

3. NUEVOS DERECHOS EN TIEMPOS DE ANOMIA

Los ejemplos de las tendencias en materia de justicia social que avalan la aparición de nuevos intereses dignos de tutela son numerosos. Tal vez el más claro es el que la propia Asamblea General de la ONU ha denominado *jurisprudencia sobre la Tierra o justicia climática* (*earth jurisprudence)* una resolución de 22 de abril de 2009. La ONU aprobó el Programa Armonía con la Naturaleza[6] e invitó a los Estados a realizar la denominada «teoría del cambio», que exige a los gobiernos y las sociedades la reorganización de sus prioridades con el fin de que internalicen el valor de la naturaleza y eviten los costes futuros asociados a la no adopción de medidas orientadas a la consecución de aquel objetivo[7]. En este mismo sentido, la importantísima Resolución de Naciones Unidas de 28 de julio de 2022 ha reconocido por primera vez como *derecho humano* el derecho a disfrutar un medio

ambiente sano y a asegurar que la naturaleza garantice la vida saludable y sostenible para las generaciones futuras[8]. Este cambio de paradigma ya se inició con el reconocimiento de los denominados derechos de los animales no humanos, que generó una ruptura de esquemas, dado que hasta ahora solo el ser humano era titular de derechos y digno de defensa y tutela por los tribunales (Martínez Dalmau, 2019: 21-22). Fue una de las creaciones normativas más disruptivas para la concepción antropocéntrica de poder.

En esta lenta construcción de los derechos humanos, percibimos que muchos de ellos van integrando nuevas garantías y que, en otras ocasiones, se crean *ex novo*, tal y como sucede con el derecho al medio ambiente sano. Sin embargo, este camino es complejo porque corresponde a la ley definir quién y cómo debe regular esta manera de reconstruir la sociedad y el orden jurídico al que tendemos, por la simple razón de que es la ley el límite y la garantía de la democracia; puentear la ley es tanto como acabar con el principio de separación de poderes y la garantía de imparcialidad objetiva y subjetiva (Ferrajoli, 2011). Sin embargo, sabemos que en esta lenta construcción muy a menudo la jurisprudencia va por delante del legislador, es decir, la ley que surge de los parlamentos es fruto de una andadura previa que normalmente nace en reivindicaciones en la ciudadanía, llega a la práctica de los tribunales y que *a posteriori* esta acaba siendo desarrollada el legislador. Con frecuencia el proceso sigue este itinerario y corresponde a la práctica judicial hallar soluciones novedosas a través de la aplicación de normas pertenecientes al Derecho internacional de los derechos humanos, que es Derecho obligatorio para los Estados firmantes y que, por tanto, puede ofrecer una respuesta *legal* a muchas de las anomias y antinomias de los ordenamientos jurídicos nacionales. Esta posibilidad es especialmente relevante en las materias que, como las analizadas en estas páginas, evolucionan a una gran velocidad.

Aquí encontramos un viejo y conocido tema, a saber, la delimitación entre los poderes ejecutivo, legislativo y judicial que impida que uno de ellos invada las atribuciones de los otros. Ante una anomia ¿dónde está el límite de la creación judicial del Derecho que garantiza que no sea invadida la potestad discrecional de la Administración pública, por ejemplo? En España, los autores que con mayor profundidad han analizado el alcance de labor interpretativa de los juzgadores son los profesores Atienza (2017) y García Amado (2009). Sin entrar en el examen detallado de esta cuestión, como procesalista considero que el punto más importante sobre el que hay que reflexionar es el de la exigibilidad de los derechos humanos, entendidos como derecho positivo aplicable cuando estamos en presencia una anomia que surge al resolver el caso concreto ante la judicatura. En otros términos, recurrir a los valores *normativizados* de los tratados internacionales de derechos humanos es obligatorio y esas normas producen eficacia *erga omnes* en los Estados que han ratificado aquellos instrumentos. Añón Roig ha analizado ampliamente el carácter vinculante del Derecho internacional de los derechos humanos para los Estados, a menudo olvidado en nuestras sentencias, cuya incidencia en tiempos de crisis es indudable (Añón, 2021: 1-26; Cardona Llorens, 2015: 135-172). El «sistema de protección de los derechos humanos» incluye en el este acervo normativo no solo los textos internacional internacionales y sus protocolos, sino también los documentos emitidos por los Comités de control de los tratados instituidos en el marco de las Naciones Unidas (típicamente,

las Observaciones Generales), que establecen pautas sobre la naturaleza, los contornos y el alcance interpretativo de cada derecho con el fin de facilitar a los operadores judiciales la aplicación del instrumento internacional en cuestión a un caso concreto. Por tanto, para resolver las anomias la judicatura cuenta con un corpus normativo multinivel que suele ser mucho más extenso de lo que a menudo reflejan las sentencias dictadas en el ámbito nacional.

En un entorno caracterizado por la evolución constante de los derechos humanos, la revolución tecnológica y la emergencia climática, es frecuente encontrar normas vagas que no generan la seguridad jurídica suficiente para resolver conflictos y asegurar la paz social. En este sentido, nos atrevemos a aventurar la paulatina y progresiva aplicación de la Carta Social Europea (revisada), de 3 de mayo de 1996, y del Pacto internacional de derechos económicos, sociales y culturales de 3 de enero de 1976, dado que ofrecen muchas respuestas a los retos que plantean los derechos sociales de las personas y su justiciabilidad, especialmente en contextos disruptivos. La concreción de los contenidos y las obligaciones de estos derechos es un proceso evolutivo. De hecho, recientemente se han incorporado al corpus los derechos ambientales para encarar la amenaza del cambio climático en el nuevo horizonte que tienen ante sí las sociedades en democracia. A título ilustrativo, cabe señalar que las vulneraciones del derecho humano al medio ambiente saludable y limpio[9] no solo afectan a la salud y a la vida privada de los habitantes del presente, sino también al planeta y a aspectos no menores de la vida de las generaciones futuras, entre ellos el trabajo, la vivienda, los servicios básicos como agua o la salud y el propio acceso a justicia[10].

4. MODELO ECONÓMICO, DERECHOS SOCIALES Y DERECHOS DE SOLIDARIDAD: ¿CUÁL ES EL LÍMITE DE LA CREACIÓN JUDICIAL EN ESTE CONTEXTO?

La Agencia Europea de Derechos Fundamentales (FRA) ha creado un marco de protección de los derechos fundamentales aplicado al mercado y al modelo económico actual. Conocedora de la omnipotencia del mercado sobre nuestras vidas y consciente de que este diseña el mundo a través del comercio y los poderes financieros, la FRA reivindica que los derechos humanos deben de ser límite a estos poderes privados porque la naturaleza de la democracia y la esencia de los derechos humanos así lo exigen[11].

En este renovado marco, vamos a seguir usando las herramientas clásicas para garantizar el haz de derechos que conforman la tutela judicial efectiva. Ahora más que nunca, el Estado social y democrático y de Derecho debe imponer su sentido, sus métodos método para resolver los nuevos conflictos y lograr la paz social en un mundo global y conectado con el fin de que, como señala Moreno Catena ([11]2021), la generación de conflictos no sea ventajosa.

4.1. LAS GARANTÍAS PRIMARIAS Y SECUNDARIAS DE LOS DERECHOS COMO OBJETO DE TUTELA POR LA JUDICATURA

En el nuevo escenario habrá materias que ya no podrán ser decididas por mayorías porque afecta a los derechos y garantías fundamentales reconocidos en la Constitución española (Todorov, 2012: 9). Como afirma Ferrajoli, hay una «esfera de lo indecidible» (Ferrajoli, 2011) que se erige como límite para los poderes públicos (y para las personas), y vincula especialmente a los primeros no solo en lo que respecta al procedimiento de producción de las normas, sino también a su contenido sustantivo o, si se prefiere, a su significado. «De este modo, política y mercado […] forman la *esfera de lo decidible*, rígidamente delimitadas por *la esfera de lo indecidible* integrada por el conjunto de derechos fundamentales» (Ferrajoli, 2011). Ambas esferas son inseparables desde la perspectiva constitucional (*ibid.*).

Las garantías de los derechos fundamentales son complejas y variadas. Ferrajoli distingue dos tipos (primarias y segundarias) que integran siempre el derecho fundamental. Las *garantías primarias* de los derechos fundamentales consisten en prohibiciones de hacer (en el caso de *los derechos de libertads*) y en obligaciones de hacer (en el caso de los *derechos sociales*). Las primeras operan a modo de *garantías negativas* impuestas al legislativo y el ejecutivo y tendentes a que desarrollen normas que aseguren que los derechos de libertad no sean lesionados por otros[12], y las segundas como *garantías positivas*, que obligan al legislador y a la administración a prever en las normas infraconstitucionales determinadas prestaciones activas para satisfacer adecuadamente los derechos sociales[13]. Por su parte, las *garantías secundarias* o *las garantías de proteccionabilidad o justiciabilidad* operan en caso de incumplimiento de las garantías primarias y se concretan en la posibilidad de recurrir a la jurisdicción para reparar aquella inobservancia, o, dicho de manera más sencilla, se materializan en el ejercicio del derecho a la tutela judicial efectiva del art. 24 CE (constituyen, pues, los derechos derivados de su *status activus procesualis)*[14]:

i) Cuando un juzgador opera para evitar o reparar una violación de un derecho y sus garantías, aborda lo que Ferrajoli denomina *antinomia.* Estas son resueltas directamente por los jueces y tribunales. En nuestra opinión, no merecen mayor precaución en el futuro porque es lo hacen los órganos jurisdiccionales cotidianamente.

ii) Sin embargo, cuando este mismo juzgador se enfrenta a la inexistencia de una regulación de estos derechos o sus garantías estamos ante lo que Ferrajoli denomina una *anomia*, es decir, ante vacíos legales o *lagunas*. Dado que en el ejercicio de la función jurisdiccional los jueces y tribunales no pueden infringir la prohibición del *non liquet,* la persona juzgadora debe tomar la decisión que proteja el derecho de la manera que considere más adecuada (Ferrajoli, 2011). Las lagunas deben ser identificadas por los órganos jurisdiccionales, que habrán de indicar a la Administración y al poder legislativo las acciones que deben llevar a cabo para cumplir con sus obligaciones de hacer a fin de suplir el vacío normativo. Por esta razón Ferrajoli denomina derechos y garantías débiles a los preceptos constitucionales que reconocen determinados derechos (educación, protección de la salud, seguridad social y

otros), aunque no por ello dejan de tener la consideración de derecho fundamental y de ser parte esencial del sistema democrático (Ferrajoli, 2014).

4.2. LA EVOLUCIÓN DE LOS DERECHOS HUMANOS HACIA LOS DERECHOS DE SOLIDARIDAD

Desde la Filosofía del Derecho se nos ha enseñado que la *primera generación de derechos humanos* nació a partir de una «matriz ideológica individualista» que con el tiempo sufriría un proceso de progresiva erosión a raíz de su impugnación por parte los movimientos obreros del siglo XIX, cuestionamiento que darían lugar al surgimiento de los derechos humanos económicos, sociales y culturales (o *derechos de segunda generación*). Los primeros protegen la *libertad* y la *igualdad formal* del individuo; los segundos, *la igualdad material* del individuo y los grupos sociales (Pérez Luño, 1991: 24). Este proceso supuso el tránsito del denominado Estado liberal de Derecho al Estado social de Derecho[15]. Como era esperable, esta dinámica de agregación de derechos no se ha detenido y en la actualidad se ha abierto paso una tercera generación de derechos caracterizados por su peculiar fisonomía. La nueva generación de derechos humanos complementa a las anteriores y da «respuesta a la "contaminación de las libertades", (dada) la erosión y degradación que aqueja a los derechos fundamentales ante determinados usos de nuevas tecnologías» (*ibid.*: 25). Estos derechos tienen como principal objeto de tutela la *solidaridad,* que es esgrimida frente a los excesos a los que han conducido la libertad, el mercado y la tecnología. Siguiendo a Pérez Luño (*ibid.*: 29-40) los derechos de tercera generación presentan unos rasgos caracteriales que los diferencian de los otros dos tipos de derechos (*ibid.*: 30).

La propuesta explicativa de Pérez Luño apunta al cambio de paradigma ante el que nos encontramos, uno de cuyos núcleos temáticos es el binomio ciudadano-tecnología. A raíz de la teorización del surgimiento de la tercera generación de derechos, la solidaridad y el altruismo se presentan como elementos nuevos que deben ser integrados en el concepto de Estado social y democrático de Derecho. Decíamos al inicio de este estudio que la naturaleza que da soporte a nuestro modelo de vida se está agotando. De hecho, este se mantiene sobre la base de esgrimir recursos que pertenecen al imaginario colectivo (naturaleza y vida sostenibles en el futuro). Por esta razón, debemos autolimitarnos en la producción y consumo y transformar nuestros hábitos para generar un sistema próspero que garantice el futuro a las próximas generaciones. A esto nos referimos cuando hablamos de los *derechos de solidaridad*. En nuestra opinión, todas las personas tienen derecho al progreso y a salir de la pobreza. Parece, no obstante, que el sistema actual no puede satisfacer de manera equitativa las necesidades básicas de todos; es más: son los países pobres los que facilitan que los países ricos mantengamos un ritmo frenético de extracción, producción y consumo (Ferrajoli, 2022: 93). Esta dinámica es absolutamente incompatible los derechos sociales establecidos en el Pacto Internacional de Derechos Económicos, Sociales y Culturales, cuyo contenido ha sido extensamente desarrollado a través de la jurisprudencia nacional e internacional, y las Observaciones Generales del Comité DESC.

Para superar estas trabas, Ferrajoli apuesta por dar un paso más y valorar la posibilidad de desarrollar un «Constitucionalismo de los mercados» o, lo que es lo mismo, integrar la perspectiva del Estado de Derecho en el ámbito de los intercambios económicos cuestiones mediante el diseño de un constitucionalismo de derecho privado (Ferrajoli, 2022: 93) que ponga límites a las relaciones privadas y a la creación de sus reglas al margen del Estado y de la *res pública* (Rodot*à*, 2014: 69). Semejante propuesta requiere una renovada alianza entre el poder legislativo y el judicial orientada a la construcción de un «constitucionalismo global» (Rodot*à*, 2014: 70). Este es el nuevo contrato social al que nos venimos refiriendo.

Entretanto, es previsible que el legislador vaya por detrás en el proceso de la regulación de estos nuevos contextos, especialmente el de la tecnología y la IA (Ferrajoli, 2022: 99). Por su parte, la prohibición de *non liquet* impone a los operadores jurisdicionales la obligación de decidir *sobre lo nuevo*. Urge, entonces, que el legislador elabore y apruebe principios básicos sobre las nuevas relaciones jurídicas[16] para conformar un marco normativo que permita al juez moverse con seguridad en la maraña legislativa casi inmanejable que caracteriza a los ordenamientos contemporáneos y proceda a hacerlos realidad a través de su aplicación al caso concreto. En este sentido, la función jurisdiccional deberá proteger la Constitución y reordenar la política de los Estados en materia de derechos (a través de la protección de su ciudadanía) para imponer límites a los poderes económicos y financieros. En puridad, aplicar la Constitución española es conseguir aquel objetivo porque, de otra forma, el sistema se vaciaría de Derecho público. Esta es la razón de ser de la propuesta de articular un constitucionalismo de Derecho privado[17].

A tal fin, los tratados internacionales sobre derechos humanos son Derecho internacional vinculante para los Estados firmantes. No son meras declaraciones morales *de lege ferenda.* Son Derecho positivo vinculante *erga omnes*. Lo único que sucede es que, con frecuencia, las obligaciones positivas y negativas asociadas a los derechos no están bien delimitadas porque, como hemos señalado, se encuentran inmersos en un proceso de constante cambio y evolución. Por tanto, la responsabilidad del Estado, también en su faceta de Estado-juez, es clara cuando las controversias jurídicas se resuelven administra o judicialmente sin tomar en consideración los derechos humanos.

4.3. LA NUEVA ERA DE LA DILIGENCIA DEBIDA DE LOS ESTADOS

Presumimos que a través de este *iter* evolutivo hemos ingresado en la era de la diligencia debida de los Estados (Courtis, 2022; Huerta Viesca y Rodríguez, 2023; Doménech Pascual, 22; Ponce i Solé, 2019), es decir, en una fase en la que el legislador internacional exige a los Estados parte de un convenio o tratado no solo que tutelen el derecho en el caso concreto, sino también algo mucho más transcendental y transformador, a saber, que generen los cambios sistémicos necesarios para transformar este modelo de Estado con el fin de que simultanee la protección de los derechos de la ciudadanía y los derechos humanos colectivos sobre bienes que son de demanio público. No podemos seguir creciendo sin prosperidad, término que no solo tiene connotaciones económicas, sino también biocén-

tricas y ecosociales. No solo implica al presente, sino también al mañana y a las futuras generaciones.

Son cada vez más los tratados internacionales y las directivas europeas que apuntan al deber de cuidado del Estado respecto a su ciudadanía —al conjunto de sujetos individuales titulares de derechos, pero también a la sociedad, entendida como colectivo interdependiente que comparte misma ética y democracia—, y que reclaman a las instituciones públicas que emprendan un camino que permita garantizar la vida de las próximas generaciones. La pandemia de COVID-19 ha sido un un ejemplo traumático reciente con la pandemia[18]. Este episodio nos ha abierto los ojos y nos ha hecho ver cuánta «estructura de Estado» hace falta para sobrevivir a una crisis sanitaria de gran magnitud y cuánta intervención hace falta para que exista la solidaridad. En otros términos, en el ejercicio de los derechos el Estado no solo tiene el deber de protegerme y reparar mi daño en un conflicto *inter partes* a través de la intervención de las fuerzas y cuerpos de seguridad del estado, la judicatura y Fiscalía, el derecho a la justicia gratuita, etc., sino que también tiene el deber de legislar, de limitar la libertad económica en pro de la igualdad para remover los obstáculos que la impiden —restricciones que en ocasiones afectan a la cultura, otras a los hábitos de naturaleza pública y privada—, delimitando todas las garantías primarias que dan forma y fondo al Derecho en cuestión[19]. Es bien cierto que, si estas garantías fallan, operarán las garantías secundarias o de justiciabilidad, pero como cabe inferir de lo dicho hasta aquí, corresponde al Estado *abstenerse* de hacer y también *obligarse a hacer* para que las personas gocen de la tutela de sus derechos. De no hacerlo, podrán incurrir en responsabilidad (Ponce i Solé, 2019; Doménech Pascual, 2022) por no legislar, por legislar mal, por no hacer, por no cambiar, etc., muy especialmente en materia de derechos sociales, unos derechos que, si bien han sido tradicionalmente calificados como derechos débiles, al fin y al cabo son los que nos dignifican como personas en la vida[20]. Hemos ingresado, pues, en la era de la justicia climática, los cuidados, la vivienda digna, servicios públicos accesibles para las personas, la defensa del territorio, la soberanía alimenticia, la igualdad de la mujer como forma de transición hacia la prosperidad y la reducción del consumo, etc. Nuevas controversias, nuevos derechos y nuevas soluciones judiciales.

Ahora bien, la función jurisdiccional no lo puede todo; al contrario, su sometimiento a la ley es la garantía de su independencia. Así, ante situaciones de anomia podrá recurrir a los acuerdos internacionales, pero no podrá suplir la libertad de la potestad discrecional de la administración para decidir cómo regular los derechos. Lo que sí podrá hacer, sin lugar a duda, es conminar a la Administración o al Parlamento a que regule de la forma que estime apropiada, siempre y cuando tutele los derechos conforme a la Constitución y los tratados internacionales de los que España forma parte. Es más, los conflictos venideros motivados por el más que presumible desarrollo acelerado de la IA son casi inimaginables. En este sentido, la diligencia debida del Estado cobra mucho sentido, dado que lo importante no será, o no será tanto, lograr una indemnización a cargo del Estado —que se podrá obtener si media daño y responsabilidad objetiva— cuanto obligar al ejecutivo o al legislativo a que actúen y hagan sus *deberes.* La prohibición de *non liquet* exige siempre a la judicatura que dé una solución, aunque no pueda crear Derecho.

La creación del Derecho queda sujeta a sus justos términos, tal y como hemos señalado *supra*. En esta nueva era, no podemos permitirnos el lujo de sucumbir a la tentación de introducir parámetros ideológicos en la función jurisdiccional. En ningún caso. Nuestra aspiración es normativista, pues pretendemos que el juez aplique la Ley de forma razonada expresando su íntima convicción judicial (art.120 CE) y que, ante la presencia de lagunas o anomias, no pueda dar saltos al vacío porque «[…] en definitiva, todas las Declaraciones y Convenios internacionales reconocen derechos fundamentales con las consiguientes garantías para los seres humanos por el hecho de serlo, con independencia de su condición, que han de servir no solo de guía y pauta al tiempo de interpretar y aplicar cualquier ordenamiento jurídico, sino de paradigma legal para la decisión judicial. […] No son solo meros criterios éticos, políticos y sociales desprovistos de sanción jurídica, sino auténticos cuerpos normativos de proyección universal y cuya fuerza de obligar son los jueces quienes la tienen que fijar al momento de dirimir la cuestión litigiosa» (Peces Morate, 1992: 96). Por esta razón, los instrumentos normativos internacionales, cada vez más frecuentemente, incorporan cada vez con mayor frecuencia la responsabilidad del Estado por la falta de diligencia exigible en la prevención y protección de los intereses y valores inspiradores de estos marcos normativos21. Se trata del deber de precaución de las administraciones públicas (Doménech Pascual, 2022; Ponce i Solé, 2019). El ejemplo más claro es el Convenio de Estambul[22], pero lo podemos encontrar esta previsión en la Convención para la protección de personas con discapacidad[23], en la Convención para la protección de menores[24] o en materia climática[25] y en relación con los derechos de las personas mayores[26]. En este camino, Naciones Unidas nos alerta de que, si no viremos el rumbo del barco, no tendremos futuro[27].

5. CONCLUSIÓN: LOS CAMBIOS ESTRUCTURALES E INDIVIDUALES VAN DE LA MANO EN LA NUEVA ERA

Necesitamos tener una perspectiva planetaria del momento en que vivimos, donde las personas y el planeta son las claves para invertir la cúpula de relaciones de poder que parecen no dejar futuro a la vida. En ese sentido apunta la determinante propuesta de Ferrajoli destinada a crear una Constitución para la Tierra.

Es necesario un cambio de paradigma que genere un nuevo constitucionalismo, una economía diferente y diverso entramado jurídico diverso. Un nuevo contrato social que admita como prioridad la preservación de la democracia y el aseguramiento del futuro de las generaciones venideras: «Semejante vacío de Derecho público en una sociedad global cada vez más frágil, transparente e interdependiente, no es sostenible a la larga sin avanzar hacia un futuro de guerras y violencias capaces de arrollar a nuestras democracias» (Ferrajoli, 2022: 103). Son muchas las cosas que podemos hacer desde la ciudadanía, la política, las empresas y las instituciones públicas para evitar este futuro indeseable.

Urge que nuestra judicatura tome en consideración el nuevo contexto en el ejercicio de la función que le es propia y que haga todo lo posible en este sentido, ya sea a través de

la protección directa y activa de la ciudadanía cuando reclama la tutela de sus derechos e intereses de conformidad con las leyes, ya mediante el señalamiento en sus sentencias de las instancias a las que corresponde esa protección, a saber, los poderes ejecutivo y legislativo, en el bien entendido de que su potestad jurisdiccional tiene límites: la potestad discrecional de la administración pública y las facultades asociadas al el ejercicio de la soberanía por parte del parlamento. Es decir, si bien no será lícito que la judicatura le indique a la Administración o al Parlamento *lo que debe hacer*, sí lo será, por el contrario, que *señale los espacios de anomia en los que deben intervenir para que la protección de los derechos e intereses de las personas sea garantizada con justicia*. Desde la academia debemos coadyuvar en esta profundización compleja y multinivel.

Inteligencia artificial, transhumanismo, identidad, igualdad de la mujer, lucha contra la pobreza, valor económico y ético de los cuidados de la vida, titularidad de derechos de la naturaleza para ejercitar acciones judiciales para su defensa, derecho de las futuras generaciones, transición climática… Se trata de retos importantísimos cuyo abordaje demanda la articulación de un renovado diálogo entre los derechos sociales, la libertad y la justicia social (solidaridad). La evolución desde la Modernidad hacia una consciencia sobre la vida interdependiente o ecodependiente nos exige tomar conciencia sobre los efectos perniciosos de nuestros hábitos de consumo, pues ahora disponemos de la información necesaria para saber que lo que compramos tiene, género, raza, clase, un pasaporte con más o menos derechos, un trasfondo más o menos democrático y evidentes repercusiones medioambientales—. El Estado debe impulsar cambios legislativos y sistémicos o estructurales que equilibren la situación actual y permitan compatibilizar la libertad de mercado y el ejercicio de los derechos individuales con la transición ecológica y los derechos colectivos. La Justicia, por tanto, deberá señalar a los demás poderes su deber de diligencia debida y precaución tanto en sus acciones directas como en el cumplimiento del deber de vigilancia y fiscalización de la actividad de las empresas. Los instrumentos internacionales de los que España forma parte son, ahora más que nunca, los anclajes jurídicos que resolverán las lagunas en un marco de conflictividad compleja, novedosa y global, y deben seguir marcando las líneas maestras vinculantes que habrán de servir como guía del ejercicio de la función jurisdiccional en tiempos de anomia.

NOTAS

1. Según la RAE, el término ecología significa «Ciencia que estudia los seres vivos como habitantes de un medio, y las relaciones que mantienen entre sí y con el propio medio».

2. Sobre la transición ecológica y la intervención activa del Estado *vid*. Plan de Recuperación, Transformación y Resiliencia del Gobierno de España, aprobado por Real Decreto-ley 36/2020, cuyo principal objetivo es impulsar la transformación de las cadenas de valor estratégicas del sector de los cuidados y la economía social, de forma multisectorial y transversal. La línea de transición ecológica se lleva a cabo a través de los derechos relativos a la salud, los cuidados, el territorio, la energía, la igualdad de oportunidades y de género, el desarrollo del tejido empresarial y asociativo y también la ciencia.

3. *Vid*. Ley 10/2023, de 5 de abril, de bienestar para las generaciones presentes y futuras de les Illes Balears (BOE 12 de junio 2023). Allí explica claramente estos conceptos transformadores de la naturaleza de las cosas y conflictos venideros. Para un análisis de estos conceptos desde el punto de vista científico, *vid.* Rockströn y Lade (2023).

4. *La aspiración más elevada: llamamiento a la acción en favor de los derechos humanos*, presentado por el secretario general de la ONU al Consejo de derechos Humanos el 24 de febrero 2020.

5. Con respecto a la atribución a los animales no humanos de la condición de sujetos de derechos, *vid.* la Ley 7/2023, de 28 de marzo, de protección de los derechos y el bienestar de los animales (BOE 7936/2023), la Ley 17/2021, de 15 de diciembre, de modificación del Código Civil, la Ley Hipotecaria y la Ley de Enjuiciamiento Civil, sobre el régimen jurídico de los animales (BOE 300 de 16 de diciembre). *Vid.,* asimismo*,* el Código de Protección y Bienestar animal publicado por el BOE, que recoge y regula todas las formas de relación entre el hombre y los animales no humanos (alimentación, usos científicos, sanidad, seguridad, explotaciones. Igualmente es importante el giro que en esta materia ofrece el Proyecto de Ley de protección, derechos y bienestar de los animales 121/000117, de 12 de septiembre. Sobre los retos que plantea este tema*, vid.* Brage Cendán (2018); también la STS 229/2022, de 11 de marzo 2022, Rec. 2142/2020, con un voto particular interesante sobre la proporcionalidad entre el concepto de «sufrimiento animal» y «dolo» del comportamiento humano. Cfr., recientemente, Cuerda Arnau (2021).

6. Disponible en: http://www.harmonywithnatureun.org/. Este pronunciamiento ha ido actualizándose año a año hasta llegar a 75th sesión de la Asamblea general de la ONU de 20 diciembre de 2020 donde se adoptó la Resolucion On Harmony with Nature (A/RES/75/220). Cfr., igualmente, el Convenio sobre Diversidad Biológica posterior a 2020 (Borrador) CBD/WG2020/2/3 de 6 de enero 2020 de Naciones Unidas, que lamentablemente todavía ha sido aprobado, pandemia mediante. Actualmente se encuentra en las últimas fases de actualización tras su última reunión WG2020/4, 21 - 26 de junio de 2022, (Nairobi, Kenia).

7. A tal fin, afirman que «[…] para lograr ese cambio alcanzar la igualdad de género, empoderamiento de las mujeres, jóvenes, enfoques sensibles a cuestiones de género, participación plena y efectiva de pueblos indígenas y comunidades locales en la implementación de este marco […] A través de la colaboración de muchas organizaciones a nivel mundial, nacional y local a fin de generar impulso. Se implementará con un enfoque basado en derechos y reconociendo el principio de equidad intergeneracional». *Vid.* la Resolución del Parlamento Europeo, de 9 de junio de 2021, sobre la Estrategia de la UE sobre la biodiversidad de aquí a 2030: Reintegrar la naturaleza en nuestras vidas 2020/2273(INI) y el Convenio sobre Diversidad Biológica posterior a 2020 (Borrador)*,* página 7. Sbre esta teoría del cambio, *vid.,* asimismo el informe*, Estrategias mundiales del sector de la salud contra el VIH, las hepatitis víricas y las infecciones de transmisión sexual para el periodo 2022-2030*, publicado por la Organización Mundial de la Salud en 2022. Sobre los derechos de la naturaleza, *vid.* «¿Puede la Naturaleza lograrlo? Estudio sobre los derechos de la naturaleza en el contexto europeo», elaborado por el Parlamento Europeo: <https://www.europarl.europa.eu/thinktank/es/document/IPOL_STU(2021)689328>. [Consulta: 24/02/2023.]

8. Más recientemente, la Recomendación del Comité de Ministros del Consejo de Europa para considerar la protección del medioambiente como derecho humano aprobado el 22 de septiembre de 2022 (CM/Rec (2022)20, donde se recomienda revisar la legislación nacional y la práctica a fin de respetar el derecho internacional medioambiental, los derechos de la CEDH y los preceptos de la Carta social Europea en relativos al medioambiente, asegurando el derecho de acceso a la justicia en la materia, implementando las políticas necesarias para hacerlo realidad en la sociedad civil, las instituciones destinadas a proteger los derechos humanos, las empresas, los *stakeholders*, los grupos de personas, las comunidades, las ciudades etc.

9. Ejemplo de ello es, entre otras muchas, la sentencia del Tribunal Europeo de Derechos Humanos dictada en el caso *Paulov c. Rusia*, de 11 de octubre de 2022, que a partir del art.8 del CEDH sobre el derecho a la vida privada, deduce en un voto particular la creación de un «[…] subderecho», entendido «[…] como un derecho humano implícito o emergente de carácter medioambiental». De no considerarlo así, continua la sentencia «[…] un aspecto del derecho a la vida privada se perdería, completamente desprotegido, y en peligro por los riesgos medioambientales. Por esta razón, este subderecho o derecho indirectamente derivado de la efectividad de la norma es extremadamente importante para la protección del medioambiente». En su § 13, afirma: «Debe aclararse que, al ampliarse para proteger el derecho en cuestión (a una vida privada) de riesgos presentes y futuros, la norma de efectividad y el derecho en cuestión siguen siendo los mismos. La ampliación de la norma de efectividad para proteger el derecho a no sufrir contaminación, ruido y otros problemas medioambientales también debe examinarse a la luz del Derecho internacional y puede verse influida por el avance de la conciencia medioambiental en Europa y en todo el mundo, que es un valor de civilización estrechamente vinculado al respeto de la dignidad humana. Y la dignidad sustenta todos los derechos humanos, incluido, por supuesto, el artículo 8». *Vid.*, también, Sentencia del Tribunal Europeo de Derechos Humanos en los casos *López Ostra c. España*, de 9 de diciembre de 1994, y *Cordella y otros c. Italia*, de 24 de enero 2019; y en la Sentencia de la Corte interamericana de Derechos Humanos acaba de resolver el Caso Comunidad de *La Oroya vs. Perú*, de 27 de noviembre de 2023. Más recientemente, *vid.* la sentencia del TEDH dictada en el caso *Kilmaseniorinnen c. Suiza* de 4 de abril de 2024.

10. Afortunadamente, adoptan este enfoque holístico la Resolución *2021* Council of Europe Parliamentary Assembly Resolution N. 2396 (Anchoring the right to a healthy environment: need for enhanced action by the Council of Europe), de 29 de septiembre de 2021, y el posterior desarrollo de la Recomendación del Comité de Ministros del Consejo de Europa para considerar la protección del medioambiente como derecho humano aprobado el 22 de septiembre de 2022 (CM/Rec (2022)20, que recomienda revisar la legislación nacional y la práctica destinada a respetar el Derecho internacional medioambiental, los derechos de la CEDH, los preceptos de la Carta social Europea relativos al medioambiente asegurando el derecho de acceso a la justicia en la materia, implementando las políticas necesarias para hacerlo realidad en la sociedad civil, las instituciones destinadas a proteger los derechos humanos, las empresas, los *stakeholders*, los grupos de personas, las comunidades, las ciudades etc.

11. Disponible en: https://fra.europa.eu/es/themes/business-and-human-rights. El texto contiene dos documentos muy interesantes sobre este tema. El primero es Bussiness and Huma rights: Access to remedy (FRA 2020) y Fredom to conduct a bussiness: Exploring the dimensions of a fundamental right (FRA 2015). Un año antes, Naciones Unidas emitió el documento Principios rectores relativos a las evaluaciones de los efectos de las reformas económicas en los derechos humanos. Informe del Experto Independiente sobre las consecuencias de la deuda externa y las obligaciones financieras internacionales conexas de los Estados para el pleno goce de todos los derechos humanos, sobre todo los derechos económicos, sociales y culturales, que establecía límites a la libertad de mercado en los Estados y obligaciones positivas y negativas para estos a fin de salvaguardar los derechos humanos. También en este sentido es muy relevante el documento de las Naciones Unidas (A/HRC/40/57).

12. Su ausencia da lugar a antinomias. Junto a estas, existen garantías primarias que contienen obligaciones de hacer (en *los* derechos sociales o *status positivus socialis*)

13. Su ausencia da lugar a lagunas.

14. Así lo reconoce Naciones Unidas cuando exige a los Estados que, a través del ejercicio de la función jurisdiccional, aseguren los derechos humanos de la ciudadanía: «Los Estados deben velar por que el acceso a la justicia y el derecho a un recurso efectivo estén garantizados, mediante mecanismos judiciales, cuasijudiciales, administrativos y políticos, con respecto a las acciones y omisiones en la preparación y/o la aplicación de políticas de reforma económica que puedan menoscabar los derechos humanos. Los

Estados deberían velar por que la población esté plenamente informada de los procedimientos, los mecanismos y los recursos de que dispone y porque esos mecanismos sean física y económicamente accesibles para todos. (21.1) El derecho a un recurso efectivo incluye reparaciones y garantías de no repetición. Un poder judicial independiente, bien financiado y proactivo es esencial tanto para evitar que las reformas económicas perjudiquen a los derechos humanos como para proporcionar recursos efectivos en caso de que se produzca un daño. Las evaluaciones de los efectos en los derechos humanos pueden servir para asegurar que haya procedimientos de rendición de cuentas y mecanismos exigiendo opciones de políticas claramente articuladas y justificadas que se hayan preparado mediante la participación inclusiva de la población potencialmente afectada».

15. No es extraño que esta idea fuera acuñada por una mujer Beatrice Webb, junto a su marido Sidney Webb. Se les atribuye la creación de la idea de red pública de protección desde el nacimiento hasta su muerte, *vid. The Minority Report*, elaborado en el año 1908.

16. Así, se observa que se ha llevado a cabo a partir de la aprobación del Reglamento relativo al Mecanismo de Recuperación y Resiliencia del Parlamento Europeo y Consejo de diciembre de 2021/241, de 12 de febrero. A partir de este texto, se aprobó la Comunicación de la Comisión. Guía técnica sobre la aplicación del principio «no causar perjuicio significativo» en virtud del citado Reglamento (DOUE núm.58, de 18 de febrero de 2021), que de una manera muy didáctica y práctica explica las diferentes situaciones con las que se pueden encontrar los jueces o Administración a la hora de aplicar el Reglamento, aportando los parámetros valorativos de solución de la infinidad de problemas aplicativos que puede tener una norma marco como esta. Consideramos que, a partir de ahora, la complejidad de los conflictos, su transversalidad, su carácter tecnológico y sus implicaciones éticas y supranacionales van a requerir la elaboración de determinadas guías interpretativas que den soporte a las decisiones inimaginables que deberán adoptar los jueces.

17. Un documento básico que he analizado para entender los retos de la función jurisdiccional para el siglo XXI es el s «Principios rectores relativos a las evaluaciones de los efectos de las reformas económicas en los derechos humanos Informe del Experto Independiente sobre las consecuencias de la deuda externa y las obligaciones financieras internacionales conexas de los Estados para el pleno goce de todos los derechos humanos, sobre todo los derechos económicos, sociales y culturales», Consejo de Derechos Humanos 40.º período de sesiones 25 de febrero a 22 de marzo de 2019 (A/HRC/40/57)

18. Durante la pandemia de COVID-19 pudimos comprobar que el ejercicio de la libertad individual pasaba por la salud colectiva, quedó en evidencia que para ejercer derechos individuales era necesario un mínimo de justicia para todas las personas y que la libertad individual y su ejercicio necesitan que la salud colectiva esté protegida. Pudimos comprobar en directo cuánta solidaridad e insolidaridad es posible ver en una sociedad, aunque, afortunadamente, la empatía y solidaridad reinaron frente al egoísmo. Parece que lo vivido en aquella coyuntura fue una antesala de la nueva época, es decir, de los derechos humanos de solidaridad: Sin ciertas dosis mínimas y comunes de igualdad se complica el ejercicio de la libertad individual. Los *otros* van a ser parte de mi y del mío.

19. *Vid*. Courtis (2023).

20. *Vid*. Folleto informativo N.º 16 (Rev. 1) - Comité de Derechos Económicos, Sociales y Culturales, que desarrolla el DESCA, entre otros muchos instrumentos.

21. Esta panoplia de normas internacionales crea los contenidos *objetivos* y *subjetivos* de los derechos que, seguramente, no tenemos en el ámbito nacional, pero que, sin embargo, se han conseguido gracias a la presión de la comunidad internacional y que elevan los estándares de exigibilidad en España. Pongamos como ejemplo la regulación relativa a la mujer (Convenio de Estambul arts.4 y 5), el medio ambiente (Tratado de París arts. 2, 6 y 7) o la inclusión de las personas con discapacidad (Convención europea art. 4). En los tres casos hay textos internacionales que sugieren la pertinencia de regular un deber de diligen-

cia debida de los Estados e introducen en nuestro Estado una doble dimensión de responsabilidad en el resarcimiento a la víctima del daño, a saber, una dimensión sistémica y una dimensión individual. Muy interesante es el modelo propuesto de Directiva europea de Diligencia debida de las empresas en materia de sostenibilidad, por la que se modifica la Directiva 2019/1937.

22. En el artículo 4 del Convenio de Estambul, que dispone: «Las partes adoptarán las medidas legislativas necesarias para promover y proteger el derecho de todos, en particular de las mujeres, a vivir a salvo de la violencia tanto en el ámbito público como en el privado». Se trata del derecho de las mujeres a vivir una vida sin violencia de género, en condiciones de igualdad y sin discriminación.

En su art. 5 el *Convenio de Estambul,* bajo el epígrafe «Obligaciones del Estado y Diligencia debida», establece la responsabilidad del Estado en su dimensión sistémica destinada a modificar la legislación (civil, penal, administrativa y laboral) adecuada para transformar esta sociedad y proteger a la mujer frente a la discriminación y la violencia, y prevé sanciones adecuadas para las personas que no cumplan, así como la responsabilidad del Estado en su dimensión individual (obligaciones destinadas a proporcionar protección a las víctimas que sufren la violencia de su pareja o expareja).

23. Firmado en Nueva York en 2006 y ratificado por España en abril de 2008.

24. Firmado el 20 de noviembre de 1989 y ratificado por España en diciembre de 1990.

25. Convención Interamericana para la Protección de los derechos de las personas mayores, aprobada por la Organización de Estados Americanos en diciembre de 2015; todavía no existe un texto de estas características en el Consejo de Europa o la UE, a pesar de los derechos reconocidos en el Pacto de Derechos Económicos Sociales y Culturales y el Pilar Social de la UE.

26. Tras la pandemia de COVID-19, Naciones unidas ha creado la Mesa Estatal por los Derechos de las Personas Mayores en el Gobierno español, tal y como se afirma en España 2050. Fundamentos y propuestas para una Estrategia Nacional a largo plazo (desafío 5.º: «Preparar nuestro Estado de Bienestar para una sociedad más longeva». Actualmente, no existe una Convención internacional de Derechos de las personas mayores. Esta sería importante por dos razones. El primer lugar, representaría un avance para el reconocimiento de sus derechos y su justiciabilidad. En segundo término, constituiría un punto de partida básico para el diseño de políticas públicas dirigidas a las personas mayores, particularmente en el ámbito de cuidados, el primer Informe sobre el impacto del cambio demográfico, que trata sus causas y su repercusión en el contexto europeo. Como primer resultado de este informe, el 21 de enero de 2021 se aprobó el *Libro Verde sobre el envejecimiento. Fomentar la solidaridad y la responsabilidad entre generaciones* (Comisión Europea, 2020). Se trata de una trayectoria natural que a buen seguro finalizará con la adopción de una Convención internacional, al igual que ha ocurrido con las personas con discapacidad o las personas menores.

27. Entre otros muchos, *vid.* los documentos Resolución adoptada por la Asamblea General de Naciones Unidas el 21 de diciembre de 2020, denominada *Harmony with de Nature*, y el Convenio sobre Diversidad Biológica (CBD/WG2020/2/3 de 6 de enero de 2020), pendiente de aprobación; el Acuerdo de París para el cambio climático, aprobado por las Naciones Unidas en diciembre de 2015 y la *Ley Europea del clima* aprobada por el Reglamento (UE) 2021/1119 del Parlamento Europeo y del Consejo de 30 de junio de 2021 por el que se establece el marco para lograr la neutralidad climática y se modifican los Reglamentos (CE) n.º 401/2009 y (UE) 2018/1999 (Legislación europea sobre el clima). Especialmente interesante es la aproximación interseccional a la igualdad de género-clima que está desarrollando la Organización para la Cooperación y Desarrollo Económico (OECD). *Vid*. Gender and the Enviroment: Builing Evidence and Policies to achieve SDGs.

BIBLIOGRAFÍA

AÑÓN ROIG, Maria José (2021): «Human rights obligations, especially, in times of crisis», *The Age of Human Rights Journal*, 17, 1-26.

ATIENZA, Manuel (2017): *Filosofía del Derecho y transformación social*, Madrid: Trotta.

BAUMAN, Zygmunt (2011): *La Modernidad líquida*, Madrid: Taramara.

BRAGE CENDÁN, Santiago Bernardo (2018): «¿Es necesaria una nueva reforma penal en el ámbito de los delitos de maltrato y abandono de animales?» *Diario La Ley*, 9187.

HAN, Byung-Chul (2017): *La sociedad del cansancio*, Barcelona: Herder.

CARDONA LLORENS, Jorge (2015): «Hacia la configuración de un «sistema» de protección de los derechos humanos de las Naciones Unidas», *Cursos de derecho internacional y relaciones internacionales de Vitoria-Gasteiz*, 1, 135-172.

CASTAÑÓN DEL VALLE, Manuel (2023): «La protección jurídico medioambiental de las generaciones futuras», *Diario la Ley*, 10232.

CORTINA ORTS, Adela, (2016): «La manida palabra ética», *Revista Contraste* [en línea] <https://www.uv.es/~fores/contrastes/quince/cortina.html>. [Consulta: 02/02/2024.]

COURTIS, Christian, (2022): «La "debida diligencia" de la empresa como medio para el respeto de los derechos humanos. ¿herramienta útil o vía para postergar obligaciones legales?», *Revista de Derecho Laboral*, 65, 411-415.

— (2023): «*Peritaje ante la Corte Interamericana de Derechos Humanos». Caso Comunidad de La Oroya vs. Perú.*

CUERDA, Maria Luisa (2021): *De animales y normas. Protección animal y derecho sancionador*, Valencia: Tirant lo Blanch.

DE LUIS, Elena (2020): *El derecho al medio ambiente*, Valencia: Tirant lo Blanch.

DOMÉNECH PASCUAL, Gabriel (2022): «Repensar la responsabilidad patrimonial del Estado por normas contrarias a Derecho», *Indret*, 4, 168-228.

ESPARZA EIBAR, Iñaki (2021): *Introducción al Derecho Procesal*, Valencia: Tirant lo Blanch.

FERRAJOLI, Luigi (2011): *Poderes salvajes. La crisis de la democracia constitucional*, Madrid: Trotta.

— (2014): *Derechos y Garantías: La democracia a través de los derechos*, Madrid: Trotta.

— (2022): *Para una Constitución de la Tierra. La humanidad en la encrucijada*, Madrid: Trotta.

GARCÍA AMADO, Juan Antonio (2009): «Interpretar, argumentar, decidir», en F. Gutierrez-Alviz y J. Martínez Lázaro (coords.), *El juez y la cultura contemporánea*, vol. 3, Madrid: Editorial Universitaria Ramón Areces, 97-144.

GUTERRES, Antonio (2020): Secretario General de las Naciones Unidas. La aspiración más elevada. Llamamiento a la acción en favor de los derechos humanos, con motivo del 75 Aniversario de las Naciones Unidas.

HISKES, Richard (2014): *Human Right to a Green Future – Environmental Rights and Intergenerational Justice*, Cambridge: Cambridge University Press.

HUERTA VIESCA, Isabel y Daniel RODRÍGUEZ (2023): «Nuevas obligaciones y nuevas responsabilidades de la Administración en materia de sostenibilidad y medioambiente», *La Ley*, 10352.

INNERARITY, Daniel, (2023): *Libertad democrática*, Barcelona: Galaxia Gutenberg.

LANDO, Henrik (2024): «Should courts decide climate policies», *Review Law and Economics*, 23(6).

MARTÍNEZ GARCÍA, Elena (2023): «La personalidad jurídica del Mar Menor», en R. Martínez Dalmau y A. Pedro Bueno, *Debates y perspectivas sobre los derechos de la naturaleza*, Valencia: Editorial Pireo.

— (2023): «Nuevos usuarios, nuevos derechos, nuevas legitimaciones: La titularidad del ecosistema del Mar Menor en España», en S. Barona Vilar (ed. lit), *Derecho del consumo y protección del consumidor sustentable en la sociedad digital del siglo XXI*, Ediciones Universidad Autónoma de Chile.

MORENO CATENA, Víctor (112021): «La jurisdicción y el Poder Judicial», *Introducción al Derecho Procesal*, Valencia: Tirant lo Blanch.

PECES MORATE, Jesús Ernesto (1992): «Las Normas y la Jurisdicción», *Revista Jueces para la Democracia*, 15, 47-56.

PÉREZ LUÑO, Antonio-Enrique (2009): «La concepción generacional de los derechos fundamentales», en F. Gutiérrez-Alviz y Conradi y J. Martínez Lázaro (coords.), *El Juez y la Cultura jurídica conteporánea,* t. I*, La tercera Generación de derechos fundamentales,* Madrid: Consejo General del Poder Judicial.
PONCE I SOLÉ, Juli (2019): *La Lucha por el buen gobierno y el derecho a una buena administración mediante el estándar Jurídico de la diligencia debida*, Alcalá: Editorial Universidad de Alcalá.
PULEO, Alicia (2009): *Claves del ecologismo social*, Madrid: Libros en Acción.
RODOTÀ Stefano (2014): *El derecho a tener derechos*, Madrid: Trotta.
SÁNCHEZ RAMOS, Begoña (2023): «Empresa y Derechos Humanos, Consejos de Administración y legislación den la Unión Europea», *Diario La Ley, 10296.*
TODOROV, Tzvetan (2012): *Los enemigos íntimos de la democracia*, Barcelona: Galaxia Gutenberg.
JUDT, Tony (2011): *Algo va mal*, Barcelona: Taurus.
VIVES ANTÓN, Tomás (1995): *La libertad como pretexto*, Valencia: Tirant lo Blanch.
VICIANO PASTOR, Roberto (2019) *Los derechos de la naturaleza: teoría, Política y Práctica*, en R. Martínez Dalmau, F. de Carlvalho, L. Estupiñan Achury y C. Storini (coords.), *La naturaleza como sujeto de derechos en el constitucionalismo democrático*, Valencia: Ediciones Pireo.
ROCKSTRÖM, Johan y Steve LADE *et al.* (2023): «Safe and just Earth system boundaries», *Nature*, 619, 102-111.
VV AA, (2024): Monográfico sobre la obra *Por una Constitución para la Tierra*, *Revista Teoría & Derecho*, 36.

Fecha de recepción: 12 de julio de 2024.
Fecha de aceptación: 20 de octubre de 2024.

BLOCKCHAIN: LA NUEVA ERA DE LA IDENTIFICACIÓN DIGITAL EN INTERNET

BLOCKCHAIN: THE NEW ERA OF DIGITAL IDENTIFICATION ON THE INTERNET

Javier Martínez Boada

Abogado
Investigador en formación
Universidad Camilo José Cela

Francisco José Santamaría Ramos

Profesor Ayudante Doctor
Universidad Complutense de Madrid

RESUMEN

Internet permite a las personas realizar en el ámbito digital prácticamente las mismas cosas que hacen en los entornos reales. En el mundo físico, las personas pueden ser identificadas a través de un documento único que contiene los datos esenciales que las individualizan frente a las restantes. Por el contrario, hoy por hoy en la esfera digital, los usuarios son capaces de tener bajo su control diferentes identidades que pueden identificarlos de manera fehaciente o no. Bajo esta situación, en los entornos digitales los usuarios pueden actuar desde el anonimato o pseudoanonimato con el fin de atentar contra la confianza en las transacciones y, en consecuencia, generar inseguridad jurídica. A la vista de tal situación, entidades como la UE han elaborado propuestas para crear carteras de identidad digital en las que se registren los datos de identificación de los usuarios. Sin embargo, no se presta la debida atención al hecho de que Internet es un espacio digital que no conoce fronteras y en el que difícilmente puede obligarse a los operadores del resto del mundo a hacer uso de una identidad digital única. En este contexto y dada la importancia de implantar un mecanismo de identidad digital única, aparece *blockchain*, un sistema trasnacional, inmutable, distribuido, seguro y descentralizado que puede ser opción idónea para instaurar una identidad digital única que permita identificar fehacientemente a los usuarios que operen en Internet. El objetivo de este artículo es analizar la situación actual de la identidad digital y argumentar que *blockchain* es un sistema idóneo para garantizar la identidad digital y facilitar la identificación en Internet.

PALABRAS CLAVE

Identidad digital, identificación, anonimato, Internet, *blockchain*.

ABSTRACT

The Internet allows people to do virtually the same things in the digital realm as they do in real enviroments. Within the physical world, every person can be identified by a unique document containing the essential data that individualizes them from everyone else. On the contrary, today in the digital sphere, users can have a multitude of different identities under their control that can reliably identify then or not. Thus, in digital environments, users can act anonymously or pseudo-anonymity to undermine trust in transactions and, consequently, producing legal uncertainty. In view of such a situation, entities such as the EU have drawn up proposals for the creation of digital identity wallets in wich users' identification data are recorded. However, due attention is not paid to the fact that the Internet is a digital space that knows no borders, and in wich it is difficult to force operators from the rest of the world to use a unique digital identity. In this context and given the importance of implementing a unique digital identity mechanism, blockchain appears as a transnational, immutable, distributed, secure and decentralized system that may be an ideal option to establish a unique digital identity that allows users operating on the Internet to be reliably identified. The objective of this article is to analyze the current situation of digital identity. We also argue that blockchain is a suitable system to ensure digital identity and facilitate the identification on the Internet.

KEYWORDS

Digital identity, ID, anonymity, Internet, Blockchain.

DOI: https://doi.org/10.36151/TD.2024.114

BLOCKCHAIN: LA NUEVA ERA DE LA IDENTIFICACIÓN DIGITAL EN INTERNET

Javier Martínez Boada
Abogado
Investigador en formación
Universidad Camilo José Cela

Francisco José Santamaría Ramos
Profesor Ayudante Doctor
Universidad Complutense de Madrid

1. INTRODUCCIÓN

Hasta hace poco, la sociedad gestionaba, procesaba y archivaba la información de una forma elemental, utilizando como soporte el papel. Sin embargo, el siglo XXI dio lugar a la digitalización y a la llegada de la denominada sociedad de la información (Santamaría Ramos, 2024: 38).

La aparición de la sociedad de la información la humanidad ha sido un punto de inflexión cuyas manifestaciones más relevantes son la digitalización y la irrupción de las llamadas tecnologías disruptivas[1].

En estas circunstancias, se han desarrollado espacios virtuales que hasta hace apenas unas décadas eran inimaginables. En ellos, las personas pueden realizar actuaciones similares a las que desarrollan en los entornos tradicionales.

En el mundo físico las personas cuentan con una única identidad[2] y una serie de datos que las diferencian del resto de individuos. Sin embargo, a la hora de acceder a los servicios de la sociedad de la información en los nuevos entornos virtuales los usuarios tienen la

posibilidad operar en multitud de plataformas utilizando diferentes identidades digitales (creadas por ellos mismos) que imposibilitan su trazabilidad y detección[3].

En contraste con lo que ocurre en los entornos reales —donde la identificación de las personas se lleva a cabo a través de documentos únicos como el DNI—, en el entorno digital los procedimientos de identificación electrónica están basados en datos alfanuméricos (por ejemplo, un usuario y una contraseña diseñados por el propio usuario). Esta información puede haber sido inventada por el usuario, de tal forma que ni el nombre de usuario ni los datos introducidos para crearlo correspondan con su identidad, es decir, puede actuar bajo diferentes seudónimos que hacen imposible su identificación. También es posible que la información utilizada por el usuario corresponda a otra persona y, de este modo, suplante su identidad (Batuecas Caletrío, 2022: 947).

A medida que la esfera digital se expande, cobra mayor importancia el tratamiento de los problemas que genera. En el ámbito de la identidad personal, el abordaje de algunos de estos inconvenientes puede incluso comportar la necesidad de replantear su configuración en el nuevo entorno. Cabe incluso concebir la posibilidad de que bajo el control de cada usuario exista una identidad digital y que ésta sea única, como ocurre en el mundo físico.

El modo de administrar la identidad digital en la red ha ido evolucionando de modo paralelo a las transformaciones que ha experimentado Internet, entre las que cabe destacar el incremento de la oferta de servicios que de un modo u otro afectan a la identidad personal. En la Web 1.0 los usuarios únicamente estaban capacitados para leer páginas y realizar comentarios en foros y blogs. Es decir, no podían interactuar de ninguna forma, por lo que su identidad se encontraba limitada. La Web 2.0 permite a los internautas introducir información para crear perfiles digitales y cada usuario es libre de conformar su identidad con los datos que decida incluir a la hora de crear su usuario. Además, la información y las cuentas de usuario están bajo control de los proveedores de servicios, que en cualquier momento pueden privar de identidad a los usuarios según sus propios criterios (Allen, 2016). Finalmente, la Web 3.0 se erige como un espacio de lectura, escritura y confianza que todavía se encuentra en desarrollo y se considera como el futuro de Internet. Se trata de un Internet descentralizado donde se distribuirá la información y los datos cuando los usuarios lo deseen y sin que ninguna entidad sea propietaria o gestione la información; esta posibilidad viene facilitada por determinadas tecnologías disruptivas, por ejemplo, *blockchain* (Gallardo Rodríguez, 2023: 1016).

Así como Internet revolucionó la forma en la que se compartía y se accedía a la información, la tecnología *blockchain* promete transformar de forma determinante la manera en la que se realizan las transacciones entre las personas físicas y jurídicas o, incluso, entre máquinas (Alvarado Bayo y Supo Calderón, 2021: 346).

La evolución tecnológica de los últimos años ha modificado nuestras vidas en multitud de aspectos. A medida que se ha explorado el potencial de la transformación digital se ha descubierto que la naturaleza segura, inmutable y descentralizada de *blockchain* tiene el potencial de superar los desafíos actuales y rediseñar la forma en la que se almacenan,

comparten, custodian y gestionan los datos de los usuarios (Albiol-Perarnau y Alarcón Belmonte, 2024: 2).

La tecnología *blockchain* comenzó a estructurarse como sistema de intercambio de activos digitales, especialmente del bitcoin, pero con el tiempo la evolución de este mecanismo ha propiciado la ampliación de sus funcionalidades, que pueden revolucionar la forma de actuar en la red (Corredor Higuera y Díaz Guzmán, 2018: 407). Una de ellas es su utilización como protocolo de identidad digital, es decir, como sistema de almacenaje de los datos de los usuarios que les permita identificarse dentro de Internet.

A la vista de las ventajas de *blockchain* a la hora de conformar una identidad única digital que permita identificar fehacientemente a los usuarios dentro de Internet, el Reglamento (UE) 2024/1183 del Parlamento Europeo y del Consejo, de 11 de abril de 2024, por el que se modifica el Reglamento (UE) n.º 910/2014 en lo que respecta al establecimiento del marco europeo de identidad digital (en adelante, eIDAS 2) regula una cartera de identidad digital europea que permitirá a los internautas almacenar, gestionar y controlar sus credenciales en los entornos virtuales (Flamini *et al.*, 2024: 1).

Es importante subrayar que Internet es un espacio que carece de fronteras y que los contenidos circulan entre países en cuestión de milésimas de segundos (Ruiloba Castilla, 2006: 56). Por esta razón, difícilmente puede regularse y crearse una cartera de identidad digital para un solo Estado o un conjunto de Estados como la UE, pues los internautas que no pertenezcan a estos continuarán navegando y actuarán en Internet con identidades no trazables y, en consecuencia, sin que se les pueda identificar.

Como se verá a lo largo de este artículo, para identificarse digitalmente los usuarios requieren de un protocolo internacional sencillo de usar y seguro que les permita administrar exclusivamente sus datos digitales en Internet e identificarse fehacientemente en los entornos virtuales para garantizar la confianza y la seguridad en las transacciones que se realizan en la red.

Bajo todos estos presupuestos, este estudio trabajo pretende, primeramente, estudiar las cuestiones de interés que plantea la identidad digital en la actualidad y analizar cómo están conformados los métodos de identificación. Una vez examinada la situación, se propondrá el uso de la tecnología *blockchain* como sistema de gestión de identidad digital y protocolo de identificación digital internacional. A tal efecto, se analizarán fuentes bibliográficas e informaciones de diversa procedencia que muestran en qué situación se encuentra la identidad digital de las personas, cómo se gestionan y regulan actualmente los datos de identidad digital de los usuarios y cómo *blockchain* puede servir para identificar digital y globalmente a los internautas.

2. LA IDENTIDAD DIGITAL

Desde el momento en el que una persona nace, adquiere personalidad jurídica y se encuentra amparada y protegida por el ordenamiento jurídico. Tal y como establece el artícu-

lo 50 de la Ley 20/2011, de 21 de julio, del Registro Civil, toda persona tiene derecho a un nombre —y, por lo tanto, a ser identificado por su nombre y sus apellidos—. En nombre constituye, además, una de las formas a través de las que los poderes públicos controlan la identidad de los individuos (Benavente Moreda, 2013).

Por su parte, la identidad de una persona[4] es el conjunto de rasgos propios que la definen como individuo y la diferencian del resto de sujetos que conforman la sociedad (Gete-Alonso y Calera, 2018: 35).

Este concepto es de vital importancia para el ser humano, puesto que es el que otorga la capacidad jurídica o, dicho de otro modo, la personalidad jurídica. En resumen, la identidad confiere la capacidad para ser sujeto de derecho (Santamaría Ramos, 2015: 39)[5].

Justamente es esta capacidad para ser sujeto de derecho la que se reconoce y protege como derecho fundamental dentro del artículo 6 de la Declaración Universal de los Derechos Humanos de 10 de diciembre de 1948. Este precepto dispone que todo ser humano tiene derecho a que se le reconozca en cualquier lugar su personalidad jurídica. De esta forma, la identidad, entendida como el conjunto de rasgos que caracterizan y distinguen a unas personas de otras y que le atribuyen su personalidad jurídica, es un derecho fundamental que debe ser reconocido con indepencia del lugar en el que actúe, es decir, tanto en el ámbito físico como en la esfera digital (Santamaría Ramos, 2015: 39)[6].

En el mundo real, la identidad se configura a partir de una serie de datos vinculados a la persona, entre ellos el nombre, los apellidos, la fecha de nacimiento, domicilio, número de identificación fiscal, etc., información que está contenida en su DNI, es decir, en su identificación a nivel nacional, o en su pasaporte, documento de identificación a nivel internacional (Merchán Murillo, 202: 185).

Como vemos, en un entorno físico, la identificación de una persona y la comprobación de su identidad como persona se lleva a cabo a través de ciertos instrumentos (por ejemplo, el DNI o el pasaporte) en los que aparezcan todos los datos que individualizan al sujeto frente a los restantes miembros de la sociedad. En cambio, en el entorno virtual, los usuarios se diferencian del resto en virtud de su identidad digital[7]. Esta puede ser definida como la expresión electrónica del conjunto de caracteres con los que una persona, ya sea física o jurídica, se individualiza frente a las demás (Fernández Burgueño, 2012: 127).

A la hora de determinar qué se entiende por identidad digital, algunos autores sostienen que este concepto alude exclusivamente a un conjunto limitado de datos, entre ellos el nombre y apellidos, los datos biométricos y el DNI. Otros, sin embargo, consideran que la identidad digital estará conformada por todos los datos que existan de una persona en de Internet, es decir, tanto los reconocidos de forma oficial (como el lugar de nacimiento, la fecha de nacimiento, la vecindad civil, la nacionalidad o el DNI) como los creados por el propio usuario (usuarios, avatares, ideologías, creencias, chats, etc.) (Batuecas Caletrío, 2022: 957)[8].

Como se ve, no existe consenso en la definición de identidad digital, pues, ya se ha dicho, algunos autores incluyen dentro de ella toda la información que existe en Internet sobre un sujeto, es decir, tanto los datos que hacen referencia a su identidad como la in-

formación que se refiere a su personalidad digital y, por ende, a su reputación digital. De acuerdo con lo expuesto, la identidad digital debe estar compuesta por los datos que sean únicamente consustanciales a la propia persona, es decir, los que hacen al usuario único y diferente dentro de la red (Batuecas Caletrío, 2022: 957).

Actualmente, las personas no poseen una identidad digital única que las diferencie fehacientemente de los demás usuarios dentro del amplio entorno de Internet, sino que cada proveedor de servicios, red social o web requiere diferentes datos personales, por ejemplo, usuarios, contraseñas, direcciones de correo electrónico, claves o datos de recuperación que conforman diversas identidades digitales de la persona para cada sitio web, red social o entorno virtual en cuestión (Ibáñez Jiménez, 2018: 322-323). De esta forma, puede asumirse como identidad digital desde el perfil que un usuario ha creado por sí mismo en la red social Facebook hasta el expediente electrónico en el que una entidad bancaria conserva actualizados los datos personales de cada uno de sus clientes (Fernández Burgueño, 2012: 127).

Junto a la amplitud de información existente para determinar la identidad digital de los internautas[9], dentro del ámbito virtual los usuarios también pueden tener intenciones maliciosas. Tal y como se encuentra conformado actualmente el espacio digital los usuarios pueden cometer el denominado «robo de la identidad digital», es decir, tienen la capacidad de hacerse pasar por otra persona bien accediendo al propio perfil del usuario que quieren suplantar, bien creando una cuenta utilizando los datos de identificación de dicha persona (Batuecas Caletrío, 2022: 952).

Asimismo, la identidad de las personas que actúan en Internet puede ser desconocida o de difícil determinación, ya que los usuarios pueden utilizar datos falsos o que dificulten la identificación (por ejemplo, nombres de usuario, direcciones de correo electrónico o *nicknames*) para registrarse en el espacio digital. Además, ni siquiera mediante la dirección IP puede identificarse al usuario, dado que este puede haberse accedido a la red a través de dispositivos electrónicos que no sean de su propiedad (Martínez Calvo, 2020: 158).

2.1. NORMATIVA APLICABLE A LA IDENTIDAD DIGITAL

Conviene enumerar brevemente las normas que el legislador ha ido aprobando para solventar estos problemas para que el lector pueda analizar rápidamente la legislación dedicada a la materia y observar en qué situación nos encontramos actualmente.

Dado que el ámbito digital no tiene fronteras y que la transmisión electrónica de datos no se circunscribe únicamente a un territorio determinado, la UE se encargó de elaborar diferentes textos normativos para ofrecer soluciones y contribuir a mejorar la identificación de los internautas (Merchán Murillo, 2022: 1389).

El primer cuerpo legal que hace breves referencia a la identificación de los usuarios en la red fue la actualmente derogada Directiva 1999/93/CE del Parlamento Europeo y del Consejo, de 13 de diciembre de 1999, por la que se establece un marco comunitario para la firma electrónica. La Directiva pretendía instituir un marco común en el que los Estados

miembros ajustaran su Derecho interno a una serie de condiciones sobre firma electrónica. Sin embargo, al transponerla a los ordenamientos jurídicos internos, los países miembros establecieron marcos jurídicos dispares que hacían imposible realizar transacciones electrónicas transfronterizas con confianza (Merchán Murillo, 2022: 1390).

En vista de este problema, se elaboró un segundo instrumento legal que derogó la Directiva 1999/93/CE. Hablamos del Reglamento (UE) n.º 910/2014 del Parlamento Europeo y del Consejo de 23 de julio de 2014 relativo a la identificación electrónica y los servicios de confianza para las transacciones electrónicas en el mercado interior y por la que se deroga la Directiva 1999/93/CE, en adelante eIDAS. Este Reglamento creó un marco normativo encargado de regular la identificación digital y los servicios de confianza dentro del territorio de la UE. En otras palabras, la norma consagró un marco jurídico común a todos los Estados miembros que establece los requisitos y las condiciones para identificar digitalmente a las personas físicas y jurídicas a través de medios de identificación electrónica (Alamillo, 2019: 9).

A título de ejemplo práctico, en España la identidad digital de una persona —y, por consiguiente, su identificación en el ámbito virtual— se obtiene a través de figuras como el DNI electrónico o los certificados electrónicos. Este tipo de instrumentos surgen con el objetivo de dar respuesta a la necesidad de conferir seguridad a las operaciones que se realizan a través de Internet y hacer posible que un usuario actúe con su identidad real. El DNI electrónico es único y no se puede falsificar. Está certificado por un tercero y, al igual que los certificados electrónicos, se utiliza principalmente para acceder a los procedimientos en la Administración pública en los que requieren este tipo de identificación. Por otro lado, el certificado digital, ya sea de persona física o jurídica, es un conjunto de datos asociados a una identidad avalada por un prestador de servicios de certificación y se instala dentro de los navegadores de los usuarios para identificarse al actuar frente a organismos públicos (Lucas Durán y García Martínez, 2019: 179).

Como puede verse, estas herramientas de identificación electrónica se utilizan en las relaciones de los individuos con el sector público. El eIDAS el que dota de libertad a los Estados miembros para que decidan si las implantan en del sector privado o no[10].

Finalmente, con el transcurso del tiempo, y a la vista de que el eIDAS dejaba cierta libertad a los Estados miembros para implantar estos mecanismos de identificación digital dentro del sector privado, surgió la idea de elaborar un espacio de confianza transfronterizo. A tal efecto, el 3 de junio de 2021 la Comisión Europea desarrolla la Propuesta[11] del eIDAS 2. Este nuevo Reglamento pretende colmar las lagunas que existen en el actual eIDAS y desarrollar una forma de identificación digital, a saber, una cartera de identidad digital europea que involucre también al ámbito privado (Merchán Murillo, 2022: 1400).

Por el momento, tan solo es una proposición que han realizado los organismos europeos. De acuerdo con la legislación vigente, puede afirmarse que garantizar la identidad digital en los espacios virtuales es, actualmente, un proceso complejo, a pesar de los mecanismos de los que disponemos actualmente para generar propiedades digitales identitarias,

asociarlas a un usuario concreto y autentificarlas mediante recursos biométricos, tarjetas digitales, certificados electrónicos etc. (Benach y Pueyo, 2013: 75).

Como se ha adelantado, Internet es una herramienta que no conoce de fronteras, es decir, que tiene alcance global (Salvador García, 2024: 82). De esta forma, la implantación de una identidad única destinada a los internautas pertenecientes a los Estados miembros de la UE se enfrenta a un problema añadido: ¿Qué sucede con aquellos usuarios pertenecientes a otros territorios que no son parte de la UE?

Hace años que nos encontramos inmersos en un contexto mundial que reclama la creación de un Derecho global llamado a ordenar las relaciones de acuerdo con criterios de justicia (López-Medel Báscones, 2024: 28). De esta forma, se hace necesaria la elaboración de una regulación internacional de la identidad digital individualizada para cada sujeto en de Internet, pues, como se ha dicho, cualquier persona desde cualquier parte del mundo puede actuar en la red.

Como se ha visto, tal y como se encuentran constituidos los entornos digitales, los internautas pueden crearse perfiles introduciendo datos creados por ellos mismos que no son confiables, pues con ellos no se puede averiguar la identidad digital que «*anda detrás*» de ese usuario o cuenta. Para solucionarlo, resulta más necesario que nunca la creación de una identidad única y universal que sirva de base para imputar derechos y obligaciones y conocer qué sujeto en concreto está actuando en el entorno virtual (Lucas Durán y García Martínez, 2019: 178).

Dado que actualmente no disponemos de una identidad digital única, los usuarios pueden llevar a cabo actuaciones en la red contrarias a la legislación, cometer delitos o realizar actuaciones para las que no poseen capacidad legal. Un ejemplo es el libre acceso de los menores a contenidos a los que, según su edad, no tendrían que acceder. Hablamos de los contenidos pornográficos[12]. A través de una identidad digital única se garantizaría que la persona que navega por Internet puede ser identificada; es decir, es quien dice ser, puede probarlo y puede demostrar su capacidad de obrar, edad o libertad de actuación (Merchán Murillo, 2022: 1388)[13].

Por el momento, las personas que navegan a través de Internet e interactúan en los entornos virtuales pueden poseer múltiples identidades, utilizar identidades de terceros a través de la suplantación de identidad o participar en actuaciones electrónicas para las que no tienen capacidad legal (por ejemplo, acceder a contenidos pornográficos, comprar[14] productos a través de Internet o registrarse[15] en redes sociales en las que no cumplen la edad mínima de inscripción).

En resumen, parece evidente la necesidad de que exista una identidad digital única por cada usuario que opere en la red. En este sentido, seguidamente se analizan algunos aspectos cruciales en la materia: el sistema actual de garantía de la identidad digital, la verificación de la identidad real de los internautas y las razones que hacen necesaria la implementación de una identidad digital única, individual e internacional para cada usuario.

3. SISTEMA ACTUAL DE GARANTÍA DE IDENTIDAD DIGITAL

Con el propósito de situar al lector en el escenario actual, es preciso analizar la evolución de los sistemas de identificación digital con el fin de conocer la modalidad ante la que nos encontramos y el esquema hacia el que nos conduce el desarrollo tecnológico y la realidad actual.

Primeramente, como sistema inicial de identidad digital, hay que destacar el modelo de identidad centralizada-jerarquizada. En él, las identidades digitales dependen de cada sitio web o de las autoridades de certificación que las entregan a los usuarios y pueden retirarlas en cualquier instante. Como puede observarse, en este sistema el control de la identidad está centralizado, es decir, es una autoridad central la que supervisa y maneja las identidades digitales de los operadores y no los propios usuarios —quienes, en contra de su voluntad, pueden ser despojados de su identidad o sufrir los efectos de una identidad falsa— (Allen, 2016).

En segundo lugar, hay que destacar el modelo de identidad federada. Su objetivo era la creación de una identidad digital que los usuarios pudieran emplear en diferentes entornos virtuales (González Granado, 2023: 16). En este modelo, el control de las identidades correspondía a las múltiples autoridades federadas. Sin embargo, solo se consiguió elaborar una oligarquía, pues la única diferencia con el sistema centralizado es que en este modelo el poder de decisión y control correspondía a unas pocas autoridades (Allen, 2016).

En tercer lugar, se creó el modelo de identidad centrada en cada usuario, es decir, aquel en cuya elaboración se partía de la premisa de que cada individuo tenía el derecho de controlar su propia identidad digital. Desde un punto de vista teórico, esta idea es fundamental para lo que hoy se conoce como identidad auto-soberana. No obstante, desde la óptica práctica, distintas compañías tecnológicas impidieron alcanzar el objetivo de que los usuarios controlasen sus propias identidades digitales y la propiedad final de las identidades continuó siendo de las autoridades encargadas de registrarlas (Allen, 2016).

Finalmente, emerge el denominado modelo de identidad digital autosoberana, que propone que cada usuario disponga de una única identidad digital validable en cualquier entorno digital y que cada persona sea la que ejerza el control exclusivo y la gobernanza de la misma (González Granado, 2023: 17).

Para determinar qué modelo está actualmente vigente en España, habrá que recurrir a dos normas: a nivel europeo el Reglamento eIDAS y el Reglamento eIDAS 2 y a nivel estatal la Ley 6/2020, de 11 de noviembre, reguladora de determinados aspectos de los servicios electrónicos de confianza, pues estos cuerpos normativos son los que regulan los sistemas con los que se identifica a las personas en el ámbito digital, los servicios de confianza que permiten su identificación y las instancias que garantizan los datos contenidos en esos servicios.

Por el momento, la legislación europea —los Reglamentos eIDAS— destina la identificación electrónica mediante servicios de confianza a las actuaciones que los usuarios quieran realizar en relación con los servicios públicos a nivel nacional o comunitario[16].

De acuerdo con el artículo 3 del eIDAS, un servicio de confianza es aquel servicio electrónico prestado normalmente a cambio de una remuneración que consiste en: *i)* la creación, verificación y validación de firmas electrónicas, sellos de tiempo electrónicos, sellos electrónicos, servicios de entrega electrónica certificada y certificados relativos a estos servicios; *ii)* la creación, verificación y validación de certificados para la autenticación de sitios web; o *iii)* la preservación de sellos, firmas o certificados electrónicos relativos a estos servicios (Ávila Gonzales 2023: 118).

Por otro lado, encontramos los servicios de confianza cualificados, es decir, los servicios de confianza que, además, cumplen con los requisitos que establece el artículo 21 del eIDAS (Canut Zazurca, 2015: 45).

Así las cosas, tal y como está configurado el sistema, es necesario que un tercero garantice que el instrumento de identificación que se está utilizando contiene efectivamente los datos verdaderos del usuario y que no han sido falsificados. Es decir, la legislación requiere que exista un prestador de servicios que actúe como intermediario para la verificación y validación de la identidad digital.

Los prestadores de servicios de confianza pueden ser cualificados o no[17]. Los no cualificados son aquellas personas físicas o jurídicas que proporcionan todos o alguno de los servicios de confianza que se han mencionado. En cambio, los prestadores cualificados son aquellos que han sido calificados como tales por parte del organismo de supervisión, previa verificación de cumplimiento de las condiciones establecidas en el artículo 24 del eIDAS y de los servicios de confianza cualificados (Canut Zazurca, 2015: 45-47)[18].

Como puede observarse, la identidad digital es gestionada por el sector de los prestadores de servicios de confianza. Este sector está formado por las personas físicas o jurídicas que tienen bajo su control los datos de la identidad de los usuarios (Rodríguez Illaraz y Palomo Zurdo, 2023: 378) y que, en consecuencia, pueden revocar o suspender los servicios de confianza[19].

Por lo tanto, hoy en día nos encontramos ante un sistema centralizado donde la gestión y control de la identidad digital queda bajo la supervisión y verificación de una autoridad.

Dado que, como se ha dicho, los servicios de confianza regulados por eIDAS están destinados a hacer posible una interacción fehaciente con los servicios públicos locales y comunitarios, en 2021 surgió la propuesta del eIDAS 2[20]. El objetivo del nuevo Reglamento es configurar un cuerpo normativo que regule la identificación digital transfronteriza dentro de la UE (Gallardo Rodríguez, 2023: 1014). Además, con esta nueva normativa se desea extender los efectos de sus previsiones al sector privado[21].

La legislación europea se orienta a cubrir las necesidades de identidad digital en las transacciones que realizan los usuarios con los organismos públicos. Sin embargo, parece que la intención del legislador es que estos efectos se extiendan al sector privado y los usuarios pueda operar bajo una identidad digital única por cada usuario.

Por lo que respecta a la identidad digital, el eIDAS 2 persigue la creación de una cartera de identidad digital europea que se implantará a través de las tecnologías descentralizadas

y de 5G (Rodríguez Illaraz y Palomo Zurdo, 2023: 364). Para poder implantar este monedero digital que recoge la identidad de un usuario concreto, es necesario elaborar un identificador descentralizado (denominado DID) que, en sí mismo, no es una identidad, sino una serie alfanumérica única y aleatoria que estará bajo el control del internauta (Llaneza González, 2021: 84).

Uno de los rasgos distintivos de un DID es que para configurarlo se utiliza la tecnología de registros distribuidos (DLT) —por ejemplo, *blockchain*—; es decir, para implantar esta cartera de identidad digital no sería necesario el modelo centralizado que existe actualmente. De esta forma, se permite la aplicación de una suerte de esquema de clave pública descentralizada (DPKI), en contraposición a los sistemas clásicos de infraestructura de clave pública (PKI), que normalmente se apoyan en la centralización de la función de emisión bajo la dirección de un prestador (Alamillo, 2019: 3).

Aun así, la propuesta eIDAS 2 afirma que «[…] el Reglamento propuesto generará costes financieros y administrativos que deberán ser asumidos por los Estados miembros como emisores de las carteras de identidad digital europea, así como por los prestadores de servicios de confianza y en línea». Es decir, el eIDAS2 continúa relacionando el poder de creación, expedición, gestión y verificación con la figura de los Estados y los prestadores de servicios de confianza, lo que, en consecuencia, evidencia los deseos de la UE de mantener un sistema centralizado, aunque la propuesta se aproxima al modelo de identidad digital autosoberana[22].

En resumen, el proyecto pretende implantar una identidad digital única a nivel europeo que se encontrará disponible para los ciudadanos y residentes de la UE que requieran identificarse o confirmar una información personal determinada y que podrá aprovecharse para acceder a servicios tanto públicos como privados a través de una cartera de identidad digital personal e individual en la que, mediante los datos de identificación básicos de una persona (como son nombre, apellidos, DNI y fecha de nacimiento), a los que podrán añadirse otros (titulaciones, salud, bancos etc.); en cada caso particular, los usuasior gestionarán su cartera de identidad y decidirán qué datos transferir a cada prestador o proveedor de servicios en línea (González Granado, 2023: 14-15).

Como se ha dicho, la UE traspasa la competencia de emisión de las carteras de identidad digital europea a los Estados miembros. De esta forma, surge la cuestión de cómo va a responder España.

Hasta ahora, en virtud de la disposición adicional sexta del artículo 3 del Real Decreto-ley 14/2019, de 31 de octubre, por el que se adoptan medidas urgentes por razones de seguridad pública en materia de administración digital, contratación del sector público y telecomunicaciones, no se permitían en las relaciones con los servicios públicos los sistemas de identificación o de firma basados en las tecnologías de registros distribuidos. Es más, como indica el segundo apartado de la disposición, solo podrán ser utilizados si lo prevé la legislación y siempre y cuando la Administración General del Estado actúe como intermediaria para garantizar la seguridad pública.

El precepto mencionado mantiene que no serán admisibles «[…] en tanto que no sean objeto de regulación específica por el Estado en el marco del Derecho de la Unión Europea».

Con la llegada del eIDAS 2, se cumple dicha previsión, y España, mediante la Resolución de 6 de julio de 2023, de la Secretaría de Estado de Digitalización e Inteligencia Artificial, por la que se publica el Acuerdo del Consejo de Ministros de 27 de junio de 2023, por el que se determinan los supuestos de validez de sistemas de identificación y firma electrónica en la Administración del Estado cuya verificación se lleve a cabo por medio de un sistema de tecnología de registro distribuido, pasa a

> «[…] considerar válido un sistema de identificación y firma de los interesados por medio de una credencial incorporada a la Cartera digital cuya verificación se lleve a cabo por medio de un sistema de tecnología de registro distribuido basado en la Infraestructura Europea de Servicios de Blockchain».

Como se explicará más abajo, *blockchain* es un sistema descentralizado[23] capaz de garantizar la confianza entre los participantes sin necesidad de que una autoridad central actúe como intermediario[24]. Sin embargo, el segundo precepto de la Resolución mencionada afirma que «[…] la Administración General del Estado actuará como autoridad intermedia que ejercerá las funciones que corresponda para garantizar la seguridad pública con relación a los sistemas que se refiere este Acuerdo». De esta forma, se distorsiona la esencia de *blockchain* y se pierde uno de sus caracteres fundamentales que hacen única a esta tecnología: la desintermediación.

Realizado el análisis de la situación jurídica en la que se encuentra actualmente la identidad digital y de las normas encaminadas a paliar los problemas de identidad, persiste la duda de qué ocurre con el resto de los internautas que se encuentran fuera del acervo legislativo de la UE y que tienen que identificarse para acceder a determinados servicios electrónicos.

Internet es un sistema que tiene un alcance global, es decir, no tiene fronteras (Andrade Salazar, Contreras Cedeño, Gómez Tuarez, Cuzco Piedra y Sinchiguano Moreira, 2023: 951). El Derecho surge de la persona. Por consiguiente, este concepto es el fundamento del Derecho global, y se ha impuesto a la idea, tan difundida actualmente, de un Derecho eminentemente legal que emana de disposiciones jurídicas controladas por un Estado o grupo de Estados a través de una comunidad, por ejemplo la UE (Domingo, Ortega, Rodríguez-Antolín y Zambrana, 2006:128).

Ya se ha dicho que la identidad digital hace referencia a los caracteres esenciales de la persona que hacen que pueda ser identificable en del mundo virtual. A este respecto, la Agenda de Naciones Unidas, mediante su Proyecto de resolución remitido a la cumbre de las Naciones Unidas para la aprobación de la Agenda para el desarrollo después de 2015 por la Asamblea General en su sexagésimo noveno período de sesiones, obliga a los países a proporcionar antes del año 2030 una identidad legal para todos los ciudadanos (*ex* objetivo 16.9). Esta cuestión es muy importante porque: *i)* la identidad es un concepto que no tiene un significado global ni es un término reconocido tradicionalmente en muchos países; y

ii) un organismo internacional como la ONU indica que, según el Derecho internacional, debe existir un derecho individual a la identidad y que, por consiguiente, los Estados tendrán que reconocer en sus ordenamientos jurídicos la identidad digital como derecho global (Gallardo Rodríguez, 2023: 1013-1014).

De esta forma y a la vista de lo expuesto, nos encontramos ante un derecho que debe ser reconocido internacionalmente y que se ejerce en una red mundial sin fronteras, por lo que de poco sirve una regulación estatal o europao.

Con ello, ¿qué posibilidades tenemos actualmente para garantizar el derecho a la identidad digital en una red que no conoce de fronteras? En efecto, anteriormente se ha hecho alusión a la tecnología que puede hacer posible esto: *blockchain*, otro sistema que no conoce de fronteras porque su alcance es transnacional[25].

4. IDENTIDAD DIGITAL ÚNICA A NIVEL INTERNACIONAL CON *BLOCKCHAIN*

Actualmente, pertenecemos a la denominada sociedad digital, en la que el papel de los Estados se ha reducido claramente debido al carácter global de Internet (De Salamanca Rodríguez, 2016: 37).

A la luz de esta consideración, puede colegirse que las dimensiones de Internet escapan de las regulaciones locales y comunitaria, y que es necesario buscar una alternativa que posibilite garantizar la identidad digital de los usuarios a nivel mundial[26].

La *blockchain* puede ayudarnos a registrar cada una de las identidades digitales de las personas y, como veremos a continuación, se puede utilizar como tecnología que permita al usuario identificarse dentro de los espacios digitales[27].

Cuando nos enfrentamos al concepto de *blockchain*, debemos tener presente que existen muchas formas de definir esta tecnología (Alvarado Bayo y Supo Calderón, 2021: 350). Esto se debe a que el alcance funcional de *blockchain* es muy amplio[28]. Por ello, las definiciones de esta tecnología normalmente centran su atención en las múltiples aplicaciones a las que se puede destinar (Hierro Viétez, 2021: 300).

A la hora de analizar *blockchain* como un sistema que nos permita identificarnos en Internet deberemos entender esta tecnología como un libro digital compartido formado por bloques conectados y archivados en una red distribuida, descentralizada y protegida a través de criptografía que nos proporciona un depósito de información incorruptible e inmutable (Pacheco Jiménez, 2019: 63).

Además, algunos autores proponen una definición que hace referencia a la capacidad de esta tecnología para identificar a las personas dentro de los espacios digitales. En este sentido, la *blockchain* puede ser entendida como un «[…] protocolo de verificación de identidad descentralizado emitido por un proveedor de servicios que no es una autoridad

estatal, capaz de verificar la identidad de un participante y luego publicar una firma de identidad» (Lucas durán y García Martínez, 2019: 180).

Los bloques que conforman la *blockchain* contienen información. Esta información refleja los datos de las transacciones de cualquier índole, es decir, los eslabones de la cadena pueden contener datos de cualquier tipo: valores, bienes, dinero, propiedades, votos (Beck y Müller-Bloch, 201: 5390), pero también, los datos personales que conforman la identidad digital de las personas.

Además, la tecnología *blockchain* posee una serie de características que la hacen idónea para ser utilizada como sistema de identificación dentro del ámbito digital. A continuación, se expondrán los rasgos más importantes que la conforman como herramienta apropiada para registrar la identidad digital y, por ende, para identificar a los usuarios dentro del entorno virtual.

La información que contiene *blockchain* es inmutable, es decir, todas las transacciones que se almacenan y registran en los bloques no pueden manipularse ni eliminarse. En caso de cambio, el *software* que se está introducido en los bloques es capaz de registrar de forma automática en un nuevo eslabón la nueva transacción sin variar el bloque anterior, facilitando que cualquier usuario pueda observar si la información ha variado o no (Charles, Emrouznejad y Gherman, 2023: 10).

Junto con la inmutabilidad, *blockchain* es una tecnología que aporta una enorme seguridad a los usuarios de la red. Los bloques que conforman la cadena se ordenan por orden cronológico y poseen un código alfanumérico denominado *hash* que corresponde al eslabón que les precede. Gracias a este código, todos los bloques se encuentran verificados con el bloque que los creó y, por lo tanto, solo se introducen y distribuyen los bloques que contengan ese código válido, eliminando las posibilidades de que se registren bloques cuya información no corresponda con el código de la red, minimizando la posibilidad de fraude, suprimiendo la pérdida de datos y conformando un sistema totalmente trazable en el que verificar que la información es correcta (Yahari Navarro, 2017: 7).

Dentro de *blockchain* no existe una autoridad central que coordine y controle la red. La descentralización es tal que se elimina cualquier intermediario y/o autoridad central. Cada nodo tiene la misma copia del registro para mantenerlo actualizado de acuerdo con las normas del protocolo de consenso que se hayan establecido (Alvarado Bayo y Supo Calderón, 2021: 352).

El sistema *blockchain* se caracteriza también por su transparencia, es decir, los participantes de la red están capacitados para acceder a la información y verificarla mediante la cadena de bloques que contiene la información de las transacciones (Corredor Higuera y Díaz Guzmán, 2018: 410).

La *blockchain* es un sistema que está distribuido entre todos los participantes de la red. Cada ordenador (también denominado nodo) posee una copia idéntica del libro digital, lo que garantiza que los datos contenidos en los bloques sean más fácilmente identificables y verificables por el resto de los usuarios (Charles, Emrouznejad y Gherman, 2023: 10).

Habida cuenta de las características de la *blockchain* y de sus capacidades como tecnología, en el siguiente apartado analizaremos cómo podrá utilizarse este sistema como mecanismo en el que se encuentren registradas las identidades digitales de los usuarios y, por ende, que demuestre que las actividades u operaciones que realiza un usuario corresponde a este y no a otra que le esté suplantando y que acredite que el usuario no utilicza una identidad falsa que no corresponda a ningún sujeto o que emplee una identidad con la que no se le pueda trazar.

La identidad digital es compleja, multifuncional y de gran relevancia. Actualmente, los usuarios tienden a administrar múltiples versiones de uno mismo que se hacen visibles en los diferentes entornos virtuales en los que operan las personas. Este hecho nos conduce a un nuevo desafío mundial: las posibles transgresiones de datos personales en línea y los sistemas de identificación vigentes (Merchán Murillo, 2021: 187).

Uno de los problemas que se presentan Internet es la certificación de la identidad de los usuarios. Hasta ahora, ninguna persona que operaba dentro de la red poseía una identidad digital única y consolidada. Sin embargo, esto puede cambiar con la tecnología *blockchain* (Colle, 2018: 4)[29].

A la vista de la capacidad de *blockchain* para servir como registro de identidad digital y, por consiguiente, como sistema de identificación dentro de los entornos digitales, las Naciones Unidas junto con la compañía Microsoft y la alianza Hyperledger tomaron la iniciativa de desarrollar un sistema capaz de proporcionar identidades digitales a los usuarios mediante tecnología *blockchain* (Sundararajan, 2021).

Como vemos, la idea de emplear *blockchain* como sistema de registro de identidades digitales y de identificación dentro de Internet ya navega por la mente de las personas. Sin embargo, resulta erróneo su encuadre, pues, como se ha reiterado, Internet es una herramienta global que puede ser utilizada desde cualquier parte del mundo y por cualquier persona.

Con el fin de que el lector entienda la funcionalidad de *blockchain* como sistema certificador de identidad digital, es preciso analizar el ya mencionado concepto de identidad digital autosoberana (González Grandado, 2023: 16).

A simple vista, con la identidad digital autosoberana, a la hora de identificarse dentro del espacio digital los usuarios podrán acceder seleccionando su propia y única identidad digital, que estará bajo su propio control y gobierno. Con ella, los servicios web otorgarían su acceso prácticamente de forma automática, evitando que los usuarios tengan que cumplimentar continuamente formularios para introducir sus datos de identidad, los cuales pueden estar viciados en cada servicio al que quieran acceder individualmente (Muñoz Moruno, 2020: 6)[30].

Ahora bien, para que esta identidad pueda estar controlada y gestionada por cada uno de los usuarios sin que una autoridad la gobierne o sea capaz de controlarla, es necesario la implementación de la tecnología *blockchain*[31].

La identidad digital autosoberana no estará almacenada dentro de cada uno de los proveedores de servicios que existen en el entorno virtual, sino que dependerá directamente de *blockchain*. Como se ha visto, esta tecnología es capaz de descentralizar el panorama digital y contribuir a que exista una única identidad que sea interoperable, es decir, creada por el usuario y utilizable por igual ante cualquier proveedor de servicios (Gallardo Rodríguez, 2023: 1016).

Tal es la revolución introducida por este modelo que llevó a que algunos países comenzaran a emitir normas técnicas de configuración de la tecnología Blockchain. Así, el 20 de diciembre de 2020, en España, la Asociación Española de Normalización, UNE, publicó la Norma UNE 71307-1 titulada «*Tecnologías Habilitadoras Digitales. Modelo de Gestión de Identidades Descentralizadas sobre* Blockchain y otras Tecnologías de Registros Distribuidos. Parte 1: Marco de referencia», donde determina:

> «Un marco de referencia genérico para la emisión, administración y uso descentralizados de aquellos atributos que faciliten la caracterización (identificación) de individuos u organizaciones, permitiendo a estos últimos crear y controlar su propia identidad digital de forma autogestionada, sin la necesidad de recurrir a autoridades centralizadas»[32].

El objetivo principal del sistema de identidad digital autosoberana dentro de *blockchain* es unificar las credenciales y los datos personales de los usuarios dentro de un sistema legal y totalmente válido. En este entorno, el usuario será capaz de registrar sus datos y credenciales para que, una vez introducida la información, los proveedores de servicios puedan consultarla y verificarla a fin de permitir el acceso a sus entornos virtuales o efectuar transacciones (Santos-Cabaleiro, 2022: 30).

Así las cosas, el desarrollo de la identidad digital mediante *blockchain* permite implementar el concepto de identidad autosoberana descentralizada sin la intermediación de entidades, autoridades o terceros, identidad que está compuesta por testimonios, afirmaciones y evidencias, técnicamente llamadas credenciales verificables, autenticadas por su emisor y comprobadas a través de mecanismos criptográficos. De esta forma, la red *blockchain* almacenará los datos de identidad de cada usuario, el emisor, la fecha, el receptor, el destino y uso de esta información sin que exista la necesidad de que una autoridad se encargue de acreditarla o transmitirla (Madrazo Aguirre, 2020: 25).

Actualmente nos encontramos ante sistemas centralizados en los que pueden utilizarse identidades fraudulentas. La identidad digital está fragmentada y dividida entre los diferentes espacios virtuales que frecuentan los usuarios. Adicionalmente, la conexión entre las identidades digitales y las físicas es escasa, lo que facilita la creación y utilización identidades falsas. Con el uso de *blockchain* y de la criptografía pueden crearse sistemas de gestión de identidad basados en identificadores descentralizados protegidos por una clave privada esté bajo control de cada usuario para solucionar este entramado de problemas (Fiaño Rodríguez, 2022: 17).

Asimismo, el uso de *blockchain* para elaborar sistemas de gestión de identidad descentralizados comportaría la eliminación de los sistemas centralizados, acrecentería la confianza en los entornos virtuales, garantizaría la integridad de los datos personales de los

usuarios, eliminaría los intermediarios y traspasaría el control y gestión de los datos e identidades digitales a los propios usuarios siempre y cuando mantengan el dominio de las claves criptográficas que les dan acceso a ellos (Fiaño Rodríguez, 2022: 17).

A la vista de la posibilidad de crear un sistema de alcance internacional en el que cada usuario tenga en su poder su propia identidad digital y sea capaz de gestionarla y utilizarla con libertad dentro de los entornos virtuales, se plantea la cuestión de a quién corresponde acreditar que la información contenida en la *blockchain* es veraz, corresponde a la identidad real del usuario y no se ha manipulado a efectos de actuar fraudulentamente en el ámbito digital.

Para que la identidad en *blockchain* sea efectiva es necesario que exista un punto de conexión con el mundo físico, es decir, resulta imprescindible que alguien dé fe de que la identidad que se encuentra registrada en *blockchain* coincide con la identidad real del usuario, tarea que pueden llevar a cabo los notarios (en caso de España) o los fedatarios públicos de los distintos países del mundo (Puig Pascual, 2018: 4)[33].

Con todo, aunque esta tecnología fue creada pensando en las criptomonedas, se han ido descubriendo diferentes utilidades para diversos sectores. Actualmente, se utiliza en gran variedad de aplicaciones: pagos globales, compartición y registro de música, trazabilidad de ventas y cadena logística, sistemas de voto, etc. Como puede verse, *blockchain* permite registrar cualquier información y, dado su potencial, se ha dicho que puede llegar a ser la base de una nueva Internet (Colle, 2018: 3).

En definitiva, para que la identidad digital se convierta en una realidad universal, es necesario un protocolo estándar aceptado por todos los países del mundo que tenga como fin registrar y procesar todos los datos de las personas que se conecten a Internet (Colle, 2018: 7). Como se ha visto, actualmente la mejor tecnología que puede garantizar la identidad digital a nivel global es *blockchain* debido a las características que la conforman.

No debe olvidarse que, para implantar esta tecnología como sistema de identificación digital, debe realizarse una revisión normativa de los diferentes países y adaptarla a un marco estándar[34], común e internacional. En particular, podemos señalar el caso de la protección de datos dentro del ordenamiento jurídico de la UE. Para implantar la identidad digital descentralizada será necesario revisar, entre otras normas, el Reglamento (UE) 2016/679 Del Parlamento Europeo y del Consejo de 27 de abril de 2016 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos, en adelante RGPD), pues su diseño se efectuó de acuerdo con los sistemas centralizados y no con los descentralizados, por lo que, por ejemplo, la figura del encargado del tratamiento no encaja dentro de los modelos descentralizados que emplea *blockchain* (Fiaño Rodríguez, 2022: 18).

Además, como se ha indicado, una de las características que posee *blockchain* y que permiten catalogarla como un sistema garante de identidad digital es la inmutabilidad de la información de sus bloques. Sin embargo, la imposibilidad de modificar la información que contiene el libro distribuido choca directamente con el derecho al olvido regulado en

el artículo 17 del RGPD. Este derecho es un aspecto fundamental de la soberanía del ciudadano sobre su identidad y la relación de esta con las entidades privadas o públicas. El citado precepto reconoce el derecho de los usuarios a suprimir sus errores, es decir, cualquier internauta puede ejercer su derecho a eliminar cualquier aspecto de su identidad para que sus datos no se mantengan indefinidamente en la red y puedan seguir siendo empleados por terceros (Fiaño Rodríguez, 2022: 18). Por tanto, es bnecesario la adaptación de ciertas disposiciones para que el sistema *blockchain* de identidad digital internacional surta todos sus efectos y se ajuste a las disposiciones locales[35].

5. CONCLUSIONES

Hasta hace unos años, las personas solo tenían que ser identificadas mediante una identidad física conformada por una serie de características y datos individuales que distinguían a unos individuos de otros[36].

Con la aparición y posterior evolución de Internet, las personas comienzan a navegar dentro de la red y actúan prácticamente igual que en el mundo analógico, es decir, interactúan con otros sujetos, realizan compras, poseen perfiles en redes sociales, consumen material audiovisual, juegan a videojuegos en línea, etc.

A medida que los usuarios interactúan en la red como en el mundo real, se torna necesaria la implantación de una identidad digital única que permita identificar e individualizar a los internautas. La creación de una identidad digital única ayuda a prevenir los riesgos del anonimato en Internet y facilita la realización de operaciones en línea, garantizando la responsabilidad en caso de que surjan consecuencias legales.

Hoy en día, los usuarios disponen de múltiples identidades digitales creadas por ellos mismos. ¿Quién no tiene varias cuentas de correo electrónico o distintos perfiles dentro de las diversas redes sociales que existen? Internet permite a los usuarios tener múltiples identidades digitales. Asimismo, los usuarios pueden hacerse pasar por otras personas, ya que son ellos quienes crean los perfiles dentro de las redes.

A todo ello hay que sumar el acceso generalizado a Internet que se ha alcanzado a lo largo de los últimos años. Es extraño encontrar a alguien que no pueda acceder a la red, pues disponemos de ordenadores y otros dispositivos como teléfonos inteligentes que tienen acceso constante a Internet.

A la vista de tal situación, es necesario que los internautas dispongan de una única identidad digital que les individualice del resto y, en consecuencia, puedan ser identificados a la hora de investigar quién «anda detrás» de un determinado perfil o quién ha podido ser el responsable de determinadas actuaciones ilícitas en la red.

La UE propone como solución la implantación de una cartera única digital que recoja los datos esenciales que hacen posible la identificación fehaciente de una persona en Internet. Este mecanismo de garantía de la identidad digital se propone para que los Estados

miembros lo implanten con el fin de identificar dentro en la red a los individuos cuya residencia habitual se encuentre dentro del territorio de la Unión o a los ciudadanos que pertenezcan a ella.

Ahora bien, Internet no se circunscribe a territorios determinados, sino que es global y no tiene fronteras. De poco sirve que un determinado conjunto de personas sí posea una identidad digital única cuando en otras partes del mundo los internautas pueden continuar realizando transacciones y actuaciones en la red con identidades falsas o que no identifiquen verdaderamente a la persona que está detrás.

Además, el concepto de identidad digital sigue en constante evolución: todavía no existe una definición legal unificada y global, su regulación internacional es igualmente incierta y, aunque existe predisposición para regularla, aún queda un largo camino por recorrer.

A la vista de todo ello, resulta pertinente la implantación de algún sistema que, mientras se reconoce y regula internacionalmente el concepto de identidad digital, dé respuesta a la inseguridad jurídica que supone actuar en Internet sin una identidad clara y verdadera que permita identificar fiablemente a los usuarios. Este sistema puede ser *blockchain*.

Actualmente, la forma de administrar la identidad digital es centralizada, es decir, la gestión y supervisión de las distintas identidades está bajo el control de los proveedores de servicios de la red. Con *blockchain* este sistema puede evolucionar hacia una identidad digital única descentralizada controlada únicamente por los usuarios, que podrán utilizarla en los sitios que deseen sin que un tercero aporte confianza o controle sus datos.

Las características de la tecnología *blockchain* (inmutabilidad, seguridad, descentralización, etc.) ligadas al carácter trasnacional del libro distribuido que la conforma, hacen que este sistema sea una opción idónea para que en él se encuentren registradas los distintos datos de identidad de la persona y, además, sea el propio usuario quien los controle y decida qué información facilitar a los prestadores de servicios.

Si bien hoy por hoy la descentralización total del sistema es prácticamente imposible porque la información de identidad digital que contiene *blockchain* puede haberse introducido erróneamente o con el objetivo de crear una identidad falsa. De esta forma, será necesaria la figura de un tercero que dé fe de que los datos son veraces y corresponden a la persona que ha introducido dicha información. A través de, por ejemplo, organismos garantes de confianza o de un fedatario público, como en España los notarios, puede resolverse este inconveniente puede resolverse mediante la creación de un sistema de identidad digital internacional, que es lo que realmente interesa a quienes firmamos este artículo y, creemos, al mundo en general.

Como se ha visto, los problemas, incertidumbre y falta de seguridad que encontramos actualmente en la red debido a la falta de una identidad digital única que individualice a los usuarios y permita trazarlos hace necesaria la búsqueda de alternativas mediante las herramientas disponibles en tanto el marco jurídico internacional encuentre una regulación global para la identidad digital y construya un sistema que garantice la identidad digital que todo usuario debe tener en Internet.

En definitiva, con la implantación de una identidad digital única, global y descentralizada a través del uso de *blockchain*, los usuarios podrán controlar sus datos dentro de la red, facilitando la información que deseen a los distintos proveedores de servicios, se evitará que esos datos queden en manos de terceros que puedan comerciar con ellos, podrá prevenirse la comisión de actuaciones ilícitas —pues los internautas podrán ser identificados— y se aportará seguridad jurídica y confianza a las transacciones mundiales que se realicen en los entornos digitales.

NOTAS

1. *Vid*. Martín Diz (2024: 113).

2. Lavín, Olate y Zambrano (2013: 18) han sostenido que «[…] los seres humanos poseemos una identidad única e invariable». También Rodríguez Galván (2017: 28) señala que «[…] como individuos poseemos una identidad única la cual se debe de proteger».

3. Finocchiaro (2012: 737) afirma que «[…] Internet un sistema potente en el que se pueden asumir diversas identidades. Como expresa la autora es fácil poseer multitud de identidades dentro de las redes sociales, juegos y páginas donde se crean diversos avatares».

4. Batuecas Caletrío (2022: 927-928) ha observado que «[…] la identidad es inherente a la persona y, en este sentido, se reconoce identidad tanto a las personas físicas como a las jurídicas».

5. Ballesteros (2001) lo ha expresado en los siguientes términos: «[…] tener identidad, ser un sujeto, equivale a poder ser juzgado».

6. González Granado, (2023: 14) afirma que «[…] todos los derechos humanos que existen *offline* deben también estar protegidos online». Por su parte, Gallardo Rodríguez (2023) apunta que «[…] es de obligada referencia tratar el reconocimiento legal a nivel europeo por medio del Reglamento eIDAS, así como a nivel nacional en España, en la Carta de Derechos Digitales, al reconocer por primera vez el "Derecho a la identidad en el entorno digital"».

7. Según Gallardo Rodríguez (2023: 1010), «[…] desde la irrupción de Internet, la identidad física evoluciona a una identidad digital o identidad 2.0, a través de las acciones que realizamos en Internet».

8. Benach y Pueyo (2013: 75) han afirmado que «[…] el correo electrónico, los formularios, la imagen, la geolocalización, los avatares, direcciones y nicknames constituyen elementos muy importantes en la configuración de nuestra identidad».

9. Pues, como indica Batuecas Caletrío (2022), «[…] los datos existentes en Internet que deberán tenerse en cuenta para conformar la identidad digital serán […] datos "oficiales" que existan de la persona (entendiendo por tales los así reconocidos por el ordenamiento jurídico: nombre, apellidos, filiación, número del DNI, fecha y lugar de nacimiento, etc.), datos reconocidos auténticamente por la propia persona (ideología, creencias, etc.) (…) y, finalmente, datos que objetivamente pueda probarse que forman parte de la identidad de la persona (p. ej., pertenencia a asociaciones, intervenciones públicas que reflejen su pensamiento, etc.)».

10. En este sentido, el considerando 13 del Reglamento (UE) N.º 910/2014 del Parlamento Europeo y del Consejo de 23 de julio de 2014 relativo a la identificación electrónica y los servicios de confianza

para las transacciones electrónicas en el mercado interior y por la que se deroga la Directiva 1999/93/CE afirma que «[…] los Estados miembros deben seguir siendo libres de utilizar o introducir, a efectos de identificación electrónica, medios de acceder a los servicios en línea. También deben poder decidir si interviene o no el sector privado en la prestación de estos medios».

11. Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se modifica el Reglamento (UE) n.º 910/2014 en lo que respecta al establecimiento de un marco para una Identidad Digital Europea.

12. *Vid*. Sanmartín (2024).

13. Calderón Palomar (2024) indica que «[…] hecho significativo es que, los menores cuentan antes con su identidad digital que con su Documento Nacional de Identidad».

14. *Vid*. Pérez Gurrea (2012).

15. Por ejemplo, en Facebook la edad mínima para poder registrarse es de 13 años en general y 14 en España: <https://www.facebook.com/help/100532533374396/?helpref=faq_content>. [Consulta: 13/05/2024.]

16. Como han puesto de manifiesto Guerola-Navarro, Oltra Badenes, Gil Gómez y Stratu Strelet (2019: 69), «[…] el Reglamento eIDAS garantiza que las personas y las empresas puedan utilizar sus propias identificaciones electrónicas nacionales para acceder a servicios públicos en línea en otros países de la UE».

17. El Ministerio de Industria y Turismo pone a disposición de los usuarios una página donde poder consultar quiénes son prestadores cualificados y quiénes no. Disponible en: <https://sede.serviciosmin.gob.es/es

-es/firmaelectronica/paginas/Prestadores-de-servicios-electronicos-de-confianza.aspx>. [Consulta: 23/02/2024.]

18. España, a través del Ministerio para la transformación digital y de la función pública recoge un listado accesible para todos los usuarios donde recopila todos los prestadores cualificados de servicios de confianza. Disponible en: <https://sedediatid.mineco.gob.es/Prestadores/TSL/TSL.pdf>. [Consulta: 23/02/2024.]

19. *Vid.* Art. 5 de la Ley 6/2020, de 11 de noviembre, reguladora de determinados aspectos de los servicios electrónicos de confianza.

20. Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se modifica el Reglamento (UE) n.º 910/2014 en lo que respecta al establecimiento de un Marco para una Identidad Digital Europea.

21. Según el considerando 1 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se modifica el Reglamento (UE) n.º 910/2014 en lo que respecta al establecimiento de un Marco para una Identidad Digital Europea. Disponible en: <https://eur-lex.europa.eu/resource.html?uri=cellar:5d88943a-c458-11eb-a925-01aa75ed71a1.0018.02/DOC_1&format=PDF>. [Consulta: 14/04/2024.]

22. *Vid.* Ibáñez Jiménez (2018).

23. «Las ventajas señaladas son consecuencia de las dos principales características de esta «tecnología»: ser una base de datos online de carácter descentralizado» (Legerén Molina, 2019: 181).

24. «Esta novedosa tecnología permite eliminar intermediarios, pero garantizando –e incluso aumentando– la seguridad que aquellos proporcionaban» (*ibid.*: 177).

25. Padilla Sánchez (2020) ha explicado que «[…] se trata de un esquema de naturaleza transnacional en el que los partícipes pueden tener acceso a servicios globales desintermediados».

26. «Es importante que las credenciales verificables emitidas en tecnología Blockchain puedan interoperar entre redes, y por ello es necesario establecer unos estándares globales comunes a todas las redes DLT» (Latorre Salvador, 2021: 56).

27. Como indican Lucas Durán y García Martínez (2019) «[…]la identidad digital puede certificarse por una autoridad estatal tanto como por terceros con un sistema de confianza. Con todo ello podemos concluir que una de las claves para controlar las operaciones en el ámbito virtual, […], es la identidad digital y que ésta no está ya exclusivamente en manos de las autoridades, sino que a efectos de poder controlar la identidad […] pueden establecerse sistemas efectivos y eficientes impulsados desde iniciativas privadas».

28. «Todas las características de la tecnología *blockchain* confieren a esta una amplia variedad de usos» (Touriño Peña *et al.*, 2022: 15).

29. *Vid.* Sullivan y Burger (2017).

30. Puig Pascual (2018) sostiene que hay que tener presente que «[…] no solo identidades personales pueden ser certificadas en blockchain, también se pueden representar empresa, propiedades… Yo podría demostrar que soy el administrador único de mi empresa y permitir que esta u otros soportes de operaciones y contratos».

31. *Vid.* Merchán Murillo (2021).

32. UNE (2020).

33. De Salamanca Rodríguez (2016: 37) afirma que «[…] los propios internautas irán fijando las reglas de ese ciberespacio, y creo profundamente que los notarios podemos convertirnos en notarios 3.0 ofreciendo soluciones a los problemas que plantea y que planteará este entorno».

34. Es cierto que ya desde los órganos internacionales se efectúan diferentes estándares internacionales, prenormativos y de facto con el fin de adoptar globalmente la tecnología *blockchain,* sin embargo, queda un largo camino por recorrer. *Vid.* Latorre Salvador (2021).

35. Existen quienes elaboran proyectos *blockchain* en los cuales se puede, a petición de los usuarios, eliminar ciertos datos para garantizar el derecho al olvido de los usuarios. *Vid.* Dorri, Kanhere y Jurdak, (2019).

36. *Vid.* Fernández Sessarego (1997).

BIBLIOGRAFÍA

ALAMILLO, Ignacio (2019): «Uso de los Sistemas de Identidad Auto-Soberana en el Sector Público Español y de la Unión Europea», *Blockchain Intelligence*, marzo, 1-21

ALBIOL-PERARNAU, Marc e Iris ALARCÓN BELMONTE (2024): «Blockchain en salud: transformando la seguridad y la gestión de datos clínicos», *Atencion Primaria,* 56(5).

ALLEN, Christopher (2016): «The Path to Self-Sovereign Identity», *Lifewithalacrity* [en línea] <https://www.lifewithalacrity.com/article/the-path-to-self-sovereign-identity/>. [Consulta: 14/05/2024.]

ALVARADO BAYO, María del Carmen y Daniela SUPO CALDERÓN (2021): «Blockchain y propiedad intelectual: aplicando una tecnología innovadora en la gestión de derechos intangibles», *THEMIS: Revista de Derecho,* 79, 345-357.

ANDRADE SALAZAR, Milton Temistocles, Jessie Nicole CONTRERAS CEDEÑO, Jenniffer Narcisa GÓMEZ TUAREZ, Maylon Elian CUZCO PIEDRA y Jostyn Steveen SINCHIGUANO MOREIRA (2023): «Una exploración segura en la Deep Web», *Código Científico Revista de Investigación,* 4(2), 949-958.

ÁVILA GONZALES, Nadia Paola (2023): *El comercio electrónico en Bolivia: el nuevo horizonte y sus desafíos jurídicos* (Tesis Doctoral), Repositorio de la Universidad de Valencia.

BATUECAS CALETRÍO, Alfredo (2022): «El derecho a la identidad y la identidad digital», *Anuario de derecho civil*, 75(3), 923-986.

BECK, Roman y Christoph MÜLLER-BLOCH (2017): «Blockchain as Radical Innovation: A Framework for Engaging with Distributed Ledgers as Incumbent Organization», *Proceedings of the 50th Hawaii International Conference on System Sciences*.

BENACH, Ernest y Miquel PUEYO (2013): «La vida, el temps, la mort, la memòriai la identitat», *Item: revista de biblioteconomia i documentació*, 57, 71-80.

BENAVENTE MOREDA, Pilar (2013): «Identidad y contexto inmediato de la persona (identidad personal, el nombre de la persona, identidad sexual y su protección)», *Anuario de la Facultad de Derecho de la Universidad Autónoma de Madrid*, 17, 105-162.

CANUT ZAZURCA, Pedro J. (2015): «El prestador cualificado de servicios de confianza – seguridad jurídica en Internet», *PIDCC: Revista em propriedade intelectual direito contêmporaneo,* 9(2), 41-55.

CHARLES, Vincent, Ali EMROUZNEJAD y Tatiana GHERMAN (2023): «A critical analysis of the integration of blockchain and artificial intelligence for supply chain», *Annals of Operations Research*, 327, 1-41.

COLLE, Raymond (2018, 18 de mayo): «La identidad digital en la internet futura con blockchain», *Asociación Chilena de Investigadores en Comunicación*, 1-12.

CORREDOR HIGUERA, Jorge Armando y David DÍAZ GUZMÁN (2018): «Blockchain y mercados financieros: aspectos generales del impacto regulatorio de la aplicación de la tecnología blockchain en los mercados de crédito de América Latina», *Derecho PUCP,* 81, 405-439.

DE SALAMANCA RODRÍGUEZ, Francisco Rosales (2016): «Testamento digital», en R. Oliva León y S. Valero Barceló (coords.), *Testamento ¿Digital?*, España: Juristas con Futuro.

DOMINGO, Rafael, Javier ORTEGA, Benjamín RODRÍGUEZ-ANTOLÍN y Nicolás ZAMBRANA (2006): *Principios de Derecho Global.* Navarra: Thomson-Aranzadi.

FERNÁNDEZ BURGUEÑO, Pablo (2012): «Aspectos jurídicos de la identidad digital y la reputación online», *Revista de Estrategias, Tendencias e Innovación en Comunicación,* 3, 125-142.

FIAÑO RODRÍGUEZ, Jacobo (2022): *Sistema de Identidad Digital Soberana y Descentralizada basada en Blockchain* (Trabajo de Fin de Grado), Repositorio Univeridade Da Coruña.

FLAMINI, Andrea, Giada SCIARETTA, Mario SCURO, Amir SHARIF, Alessandro TOMASI y Silvio RANISE (2024): «On Cryptographic Mechanisms for the Selective Disclosure of Verifiable Credentials», *arXiv preprint*.

GALLARDO RODRÍGUEZ, Almudena (2023): «Identidad digital y responsabilidad civil de las plataformas digitales: de las redes sociales al metaverso», *Actualidad jurídica iberoamericana,* 18, 1008-1033.

GETE-ALONSO Y CALERA, María del Carmen (2018): «La inscripción de nacimiento en la ley 20/2011. Entre el derecho a la identidad de la persona y la reserva de la maternidad», *Revista de Derecho Civil,* 5(1), 1-54.

GONZÁLEZ GRANADO, Javier (2023): *De la identidad a la identidad digital soberana* [en línea] <https://tallerdederechos.com/de-la-identidad-a-la-identidad-digital-soberana/>. [Consulta: 12/05/2024.]

HIERRO VIÉTEZ, Gonzalo (2021): «Introducción al Blockchain, los contratos inteligentes y su relación con el arbitraje», *THÊMIS-Revista de Derecho*, 79, 299-309.

IBÁÑEZ JIMÉNEZ, Javier Wenceslao (2018): *Derecho de Blockchain,* Navarra: Aranzadi.

LLANEZA GONZÁLEZ, Paloma (2021): *Identidad digital, actualizado a la Orden ETD/465/2021, de 6 de mayo (sobre métodos de identificación remota) y a la propuesta de Reglamento eIDAS2*, Madrid: Wolters Kluwer.
LÓPEZ-MEDEL BÁSCONES, Manuel (2024): «La contemporaneidad del derecho digital y de los derechos fundamentales», en F. J. Santamaría Ramos (dir.), *Derechos Digitales*, Valencia: Tirant Lo Blanch.
LUCAS DURÁN, Manuel (dir.) (2019): *Residencia fiscal: problemática y cuestiones actuales*, Documentos de Trabajo del Instituto de Estudios Fiscales, 6/2019.
MADRAZO AGUIRRE, Teresa Juana (2020): *Identidad digital soberana y tecnología Blockchain. Modelo de negocio y plan de marketing de la start-up «LinKple»* (Trabajo de Fin de Grado), Repositorio de la Universidad Pontificia Comillas.
MARTÍNEZ CALVO, Javier (2020): «El derecho de rectificación ante informaciones falsas o inexactas, con especial mención a las publicadas en Internet/The reply right against false or inexact information, with particular reference to information published on the Internet», *Revista de Derecho civil,* 7(4), 137-181.
MERCHÁN MURILLO, Antonio (2021): «Identidad digital Blockchain e Inteligencia Artificial: aspectos jurídicos de presente y futuro a debate», *Ius et Scientia,* 7(1), 183-203.
— (2022): «La identidad digital en la contratación electrónica: una mirada desde el derecho internacional privado», *Actualidad jurídica iberoamericana,* 16, 1386-1411.
MUÑOZ MORUNO, Laia (2020): Identidad digital soberana (Trabajo de Fin de Grado), *Repositorio de la Universitat Politècnica de Catalunya*.
PACHECO JIMÉNEZ, María Nieves (2019): «De la tecnología blockchain a la economía del token», *Derecho PUCP,* 83, 61-87.
PUIG PASCUAL, Álex (2018): «Identidad digital sobre «Blockchain» a nivel nacional», *ICADE Revista de la Facultad de Derecho,* 101, 1-5.
RODRÍGUEZ ILLARAZ, Clara y Ricardo Javier PALOMO ZURDO (2023): «Transparencia, ciberseguridad e identidad digital en el entorno 5G», *Revista española de la transparencia,*18, 359-380.
RUILOBA CASTILLA, Juan Carlos (2006): «La actuación policial frente a los déficits de seguridad de Internet, *IDP: revista de Internet, derecho y política = revista d'Internet, dret i política*, 2, 52-62.
SALVADOR GARCÍA, Miriam (2024): «Derecho de acceso universal a internet. En F. J. Santamaría Ramos (dir.), *Derechos Digitales*, Valencia: Tirant Lo Blanch.
SANTAMARÍA RAMOS, Francisco Javier (2015): «Identidad y reputación digital Visión española de un fenómeno global», *Ambiente Jurídico*, 17, 11-43.
— (2024): «Datos abiertos y reutilización de la información: una mirada europea», *Revista Canaria De Administración Pública*, Extra 0, 37-64.
SANTOS-CABALEIRO, Pablo (2022): *Análisis y prototipado de Identidad Digital Descentralizada basada en Blockchain* (Trabajo de Fin de Grado), Repositorio Univeridade Da Coruña.
SUNDARARAJAN, Sujha (2021): «Microsoft, Hyperledger, UN Join Blockchain Identity Initiative», *Coindesk* [en línea] <https://www.coindesk.com/markets/2018/01/23/microsoft-hyperledger-un-join-blockchain-identity-initiative/>. [Consulta: 23/05/2024.]
YAHARI NAVARRO, Benjamín (2017): «Blockchain y sus aplicaciones», *Universidad Católica Nuestra Señora de La Asunción* [en línea] < https://bit.ly/2rc0iZ9. Fecha de último acceso: 25/08/2018>. [Consulta: 23/03/2024.]

Legislación

Directiva 1999/93/CE del Parlamento Europeo y del Consejo, de 13 de diciembre de 1999, por la que se establece un marco comunitario para la firma electrónica.
Reglamento (UE) Nº. 910/2014 del Parlamento Europeo y del Consejo de 23 de julio de 2014 relativo a la identificación electrónica y los servicios de confianza para las transacciones electrónicas en el mercado interior y por la que se deroga la Directiva 1999/93/CE.

Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se modifica el Reglamento (UE) n.º 910/2014 en lo que respecta al establecimiento de un Marco para una Identidad Digital Europea.
Ley 6/2020, de 11 de noviembre, reguladora de determinados aspectos de los servicios electrónicos de confianza.
Ley 20/2011, de 21 de julio, del Registro Civil.
Decreto-ley 14/2019, de 31 de octubre, por el que se adoptan medidas urgentes por razones de seguridad pública en materia de administración digital, contratación del sector público y telecomunicaciones.
Resolución de 6 de julio de 2023, de la Secretaría de Estado de Digitalización e Inteligencia Artificial, por la que se publica el Acuerdo del Consejo de Ministros de 27 de junio de 2023, por el que se determinan los supuestos de validez de sistemas de identificación y firma electrónica en la Administración del Estado cuya verificación se lleve a cabo por medio de un sistema de tecnología de registro distribuido.
Norma UNE 71307-1 titulada «Tecnologías Habilitadoras Digitales. Modelo de Gestión de Identidades Descentralizadas sobre Blockchain y otras Tecnologías de Registros Distribuidos. Parte 1: Marco de referencia.

Fecha de recepción: 16 de julio de 2024.
Fecha de aceptación: 20 de octubre de 2024.

SOBERANÍA, PROPIEDAD E ISLAS: MEDITACIONES JURÍDICAS SOBRE FRAGMENTOS DEL IMPERIO ESTADOUNIDENSE

SOVEREIGNTY, PROPERTY, AND ISLANDS: LEGAL MEDITATIONS ON PARTS OF THE AMERICAN EMPIRE

Jesús Pérez Caballero

Profesor-investigador, «Investigadores por México CONAHCYT», El Colegio de la Frontera Norte (México)

RESUMEN

El estudio del estatus jurídico de algunos territorios de Estados Unidos de América, en específico las islas (no incorporadas/organizadas y no incorporadas/no organizadas, según la terminología política estadounidense), permite entender ciertas cuestiones filosófico-jurídicas, entre ellas los límites de la soberanía, la naturaleza de la propiedad privada o la idea de «tierra de nadie» (*terra nullius*), así como las ponderaciones que la primera potencia mundial realiza entre sus intereses y los principios del Derecho internacional público. La taxonomía de este artículo, además, plantea la eventual existencia de un «Derecho imperial», entendido como conjunto de reglas de Derecho público que clasifican y jerarquizan territorios. De este modo, se superan interpretaciones economicistas que equiparan «imperialismo» e «imperio» y se pone en relación el ámbito jurídico con campos como el político, el militar o el geoestratégico.

PALABRAS CLAVE

Derecho internacional Público, soberanía, isla, propiedad privada, Estados Unidos de América, imperio.

ABSTRACT

The study of the legal status of some territories of the United States of America, specifically the islands (unincorporated/organized and unincorporated/unorganized, according to American political terminology), allows us to understand certain legal philosophical issues, such as the limits of sovereignty, the nature of private property or the idea of «no man's land» (*terra nullius*), as well as the considerations that the world leading power makes between its interests and the principles of International Law. The taxonomy of this article also raises the possibility of the existence of an «Imperial Law» as a set of public law rules that classify and prioritize territories. This would overcome economistic interpretations that equate «imperialism» to «empire», enabling to relate the legal field to others, such as politics, the military or the geostrategic.

KEYWORDS

International Law, sovereignty, island, private property, United States of America, empire.

DOI: https://doi.org/10.36151/TD.2024.115

SOBERANÍA, PROPIEDAD E ISLAS: MEDITACIONES JURÍDICAS SOBRE FRAGMENTOS DEL IMPERIO ESTADOUNIDENSE

Jesús Pérez Caballero
Profesor-investigador, «Investigadores por México CONAHCYT»,
El Colegio de la Frontera Norte (México)

1. INTRODUCCIÓN

La discusión sobre los imperios políticos —entendidos como entidades supraestatales y expansionistas— y la relación política y jurídica entre las partes que los componen persiste en la actualidad (Wheatley, 2016; Kumar, 2010). En lengua española contamos con numerosas aportaciones historiográficas sobre el imperio español, especialmente en Nueva España, que permiten reflexionar sobre las implicaciones de este tipo de estructuras políticas en los ordenamientos jurídicos que las han heredado (Tau, 2021; Bueno, 2000).

A pesar de que los avatares de los últimos años podrían sugerir la emergencia de orden jurídico internacional multipolar (agresión y conflicto armado internacional Ucrania-Rusia de 2022, conflicto armado internacional Israel-Gaza de 2023, etcétera), la asunción de EE. UU. como principal potencia contemporánea torna pertinente el estudio de la organización jurídica de las distintas entidades territoriales que componen ese país, entendido, también, como imperio.

Este artículo propone entender esos sujetos políticos coordinados por la soberanía estadounidense como parte de un «Derecho imperial estadounidense». No se trataría de asimilar el uso popular de la palabra «imperialismo» —término que, desde principios del siglo XX, se ha interpretado a la luz del repetidísimo adagio leninista «el imperialismo, fase superior del capitalismo»—, sino como un conjunto de reglas de Derecho público que clasifican y jerarquizan los territorios de su área de interés. Ello sin perjuicio de su relación histórica con las expansiones económicas de recursos, las ocupaciones militares alentadas o contestadas, o, en definitiva, con las previsiones del Derecho internacional en materia de ocupación o descolonización.

Dada la imposibilidad de abordar en detalle todos los rasgos de ese ordenamiento jurídico, el presente estudio expone y comenta, en primer término, una serie de características de ese «Derecho imperial estadounidense» a partir de la clasificación jurídico-administrativa de EE. UU. en su territorio para mostrar las distintas maneras de entender los sujetos políticos que, como potencia, debe coordinar. En el apartado pertinente se enlistarán los denominados «casos insulares» (*insular cases*), es decir, una serie de sentencias dictadas a principios del siglo XX por la Corte Suprema (CS) estadounidense (cuyas funciones son equivalentes, con matices, a las del Tribunal Supremo y el Tribunal Constitucional españoles) sobre las recientes posesiones que EE. UU. arrebató a España por conquista y que permitieron establecer distinciones en el modo de incorporar territorios a la plenitud de derecho de quienes allí viven. Seguidamente, el núcleo del artículo analiza las características de una serie de islas o atolones que ejemplifican el modo en que EE. UU. gestiona esos territorios como parte de su expansión marítima. Como se verá, se trata de entidades con distinta naturaleza, historias a veces confluyentes y otras disímiles, y, en cualquier caso, con regímenes de propiedad, modos de clasificación jurídica o conflictos sociopolíticos —incluso de soberanía disputada— que se han extendido hasta la actualidad.

Este trabajo trata de presentar técnicamente algunas de las discusiones jurídico-políticas clave del poder todavía hegemónico, lo que permite analizar el funcionamiento de las leyes imperiales en nuestro siglo. Se trata, por tanto, de una reflexión jurídica de utilidad para entender las relaciones entre los ordenamientos jurídicos nacionales y el Derecho internacional. También puede interesar a los lectores iberoamericanos no solo por la evidente influencia en sus países de la principal potencia mundial, sino también porque ofrece la oportunidad de reflexionar sobre el funcionamiento específico de ideas que son partes del debate político, entre ellas las de federalismo o secesión. Se trata, en muchas ocasiones, de situaciones límite en las que lo jurídico se amalgama con otros aspectos (militares, ambientales, tecnológicos), circunstancia que posibilita la discusión sobre algunos términos

jurídicos (soberanía, propiedad, *terra nullius*, posesión, etcétera) y engarza este artículo, también, con reflexiones universales sobre filosofía del Derecho.

2. TAXONOMÍA IMPERIAL Y DERECHO INTERNACIONAL PÚBLICO

2.1. MATICES PARA UN «DERECHO IMPERIAL ESTADOUNIDENSE»

Un vistazo a la geografía de EE. UU. muestra varias evidencias relevantes para el análisis que me propongo realizar. En primer lugar, hay un núcleo federal soberano. Son los EE. UU. como Estado parte del orden internacional vigente, simbolizado en términos físicos por la sede federal de Washington D.C., y en términos simbólico por las instituciones políticas, de seguridad y justicia, así como por el monopolio en cuestiones como, por ejemplo, la celebración de tratados internacionales o la declaración de una guerra.

En este espacio geográfico, la soberanía está, a su vez, delegada —pero condicionada— en los cincuenta estados de la unión: los estados fundadores de 1776 (las «trece colonias») y los que se han ido incorporando a la federación lo largo de los dos siglos y medio siguientes. Aun así, la completitud geográfica que se observa en cualquier mapa de EE. UU. es engañosa. Lo que se ve, en realidad, son retazos ganados al tiempo y al espacio a través de cálculos y acuerdos políticos, y avances militares geoestratégicos exitosos que han posibilitado una unidad forjada retrospectivamente y enmarcada en una teleología que narra su historia como un «destino manifiesto». En este sentido, los estados de EE. UU. tienen distintos rasgos, unas diferencias claramente asociadas a la historia no lineal del país. Por ejemplo, hay cuatro «mancomunidades» o *Commonwealth*: Massachusetts, Pensilvania, Virginia (que formaban parte de las trece colonias fundadoras) y Kentucky. Desde 1959, hay un estado insular, el oceánico Hawái (el último que fue incorporado a la federación). Algunos fueron parte de otro imperio (el 55 % del territorio fue arrebatado a México en el siglo XIX); otros, además, fueron independientes (Vermont o Texas); en determinados casos (por ejemplo, California) se planteó una división entre norte y sur, del mismo modo que la península mexicana está dividida en las entidades federativas de Baja California y Baja California Sur. Las épicas expansiones hacia el Oeste narradas en los wésterns cinematográficos acabaron irradiando una estatalidad que en alguna ucronía podría haber creado Estados que nunca existieron. Finalmente, otros territorios fueron comprados (Alaska a Rusia, el 30 de marzo de 1867; Florida a España, el 22 de febrero de 1819; Luisiana a Francia, el 30 de abril de 1803). Durante la administración de Donald Trump el presidente estadounidense (2017-2021 y desde 2024), propuso comprar Groenlandia a Dinamarca (o, en su caso, realizar un trueque por Puerto Rico).

En un radio exterior a la masa continental en la que se encuentran todos los estados federados —salvo Hawái—, hay unos territorios que, a ojos de un no experto, pasarían también por Estados, sea federados o independientes. Se trata, sin embargo, de territorios que poseen una *estatalidad orbitante*, si se me permite la metáfora. En último término, gravitan en torno a las decisiones de EE. UU. (concretamente, del Congreso federal). Es decir,

podría defenderse su estatalidad material (su viabilidad como Estados en el ordenamiento internacional) si esta se quiere distinguir de la *estatidad* de los estados federados. Cada uno de estos territorios tiene un estatus distinto que depende de la naturaleza de su incorporación o de otros factores que engranan lo jurídico y lo político de modo tal que en ocasiones parece que se utilizan terminologías *ad hoc* para determinarlo. Precisamente, las variaciones en el estatus de Samoa, Guam, las Islas Marshall o las Islas Vírgenes Estadounidenses tiene consecuencias incluso en el Derecho civil. Lo veremos al analizar determinadas cuestiones relacionadas con el derecho de propiedad en las que es relevante la distinción entre nativos y no nativos.

Paradójicamente, un rasgo común a todos estos territorios insulares es que su estatus político es menor al de los estados federados. Tómese el ejemplo de Puerto Rico. Se trata, oficialmente, de un «Estado libre asociado» (*Commonwealth of Puerto Rico*), término enrevesadamente nebuloso debido a su polisemia. Tan «mancomunidad» (esa es la traducción de *Commonwealth*) es Puerto Rico, al igual que las Islas Marianas del Norte (Mancomunidad de las Islas Marianas del Norte) y algunos estados federados mencionados más arriba. Incluso hay mancomunidades de municipios (por utilizar un ejemplo pedestre, la Mancomunitat de municipis de La Safor, mancomunidad de municipios de La Safor, una comarca valenciana, etcétera). De hecho, el término «mancomunidad» debe hacer equilibrios entre el convenio bilateral y los acuerdos *ad hoc*, ya que «[…] no es un estatus político autorizado por la Constitución de los EE. UU.», que solamente contempla cuatro estatus políticos: federación, estado, territorio y tribu (Statham, 2004: 257).

A su vez, el imperio se prolonga hacia el Caribe y el Pacífico en otro anillo insular formado por territorios (islas, atolones o arrecifes) a los que se atribuye un estatus específico en función de la taxonomía de incorporado/no incorporado y organizado/no organizado. Estos tienen, también, estatus diferentes, sobre los que se reflexiona extensamente más abajo. Además, han cambiado a lo largo del tiempo, ya que se han utilizado como monedas de cambio políticas, económicas o militares.

Finalmente, hay otras categorías jurídico-políticas de ordenación territorial que, si bien no son objeto de análisis en este artículo, serán mencionadas porque forman parte del «Derecho imperial estadounidense». Se trata de territorios con los que, por utilizar una metáfora, la soberanía estadounidense negocia, o a los que somete, en virtud de su fortaleza. Más que a cuestiones coyunturales, la negociación o sometimiento responde, en estos casos, al hecho de que EE. UU. no puede incorporarlos políticamente —porque son partes de Estados soberanos con capacidad de resistir o aliados—, o porque, incorporados jurídicamente a EE. UU., no pueden desdibujarse en la genérica ciudadanía estadounidense sin que estos perciban un daño (histórico, cultural). En este último grupo las tribus o naciones étnicas, que tienen una autonomía específica y a las que se reconocen derechos relacionados con la explotación de recursos o la organización de sus partes, pero que no llegan ser soberana (por ejemplo, la *Comanche Nation* o nación comanche, entre otras muchas). Como se verá más adelante, la denominación de «tribu reconocida federalmente» es clave para ejercer determinados derechos de los que otros grupos étnicos nacionales de EE. UU. carecen.

Por otra parte, es preciso hacer referencia al conjunto de bases militares que posee EE. UU. en los cinco continentes —y en las que están emplazados sus materiales de guerra o sus tropas—, que deben regirse conforme a los acuerdos establecidos con las autoridades soberanas. En este rubro se incluyen también las *contra-bases*, también llamadas *black sites* (sitio o centro negro, centro clandestino de detención), término que alude a aquellos lugares donde operaba la «guerra contra el terrorismo» mediante las «rendiciones extraordinarias» y otras actuaciones ilícitas llevadas a cabo con el consentimiento expreso o tácito del país soberano que albergaba ese tipo de centros. También incluye zonas inherentemente ambiguas que, desde el punto de vista jurídico, se hallan en puntos ciegos del Derecho internacional: además de un enclave como Guantánamo, cabe mencionar también los lugares situados en aguas internacionales, como las cárceles flotantes de la isla británica de Diego García (Pigrau, 2016; Guardans, 2009; Maris, 1967).

En resumen, según las fuentes jurídicas administrativas estadounidenses, todas estas entidades pueden clasificarse partiendo de los binomios incorporado/no incorporado y organizado/no organizado (Departamento de Interior de EE. UU., s.f. *a*). Aunque en este artículo veremos que estas diferenciaciones tienen excepciones y demandan matices, una primera aproximación a las categorías más problemáticas es la que entiende como «incorporadas» a aquellas entidades que alcanzan la categoría de estado federado, además de aquellas que adquieren su estatus directamente de la soberanía (a modo de prolongaciones de esta, como el mar territorial, o, como sucede con el Atolón de Palmira, por haber sido parte de una entidad incorporada). Por su parte, una entidad «no organizada», es aquella que no cuenta con un gobierno o que depende de otra entidad administrativa mayor debido a que determinadas circunstancias materiales (la insularidad, el tamaño o la falta de población) o formales (la inacción de la entidad soberana) les impiden disponer de un gobierno funcional. Así, se observan las siguientes categorías:

i) Incorporado/organizado. El medio centenar de estados federados que forman EE. UU.

ii) Incorporado/no organizado. Es el Atolón de Palmira por haber sido parte de Hawái. Antes de ser estados tuvieron este estatus California, Dakota o Nevada, sobre todo por la falta de población. Además, se encuadran en esta categoría el mar territorial o franja de mar adyacente al territorio (12 millas náuticas o 22.2 km), así como los buques con bandera estadounidense, tanto militares como la marina mercante civil, derivados del vínculo del pabellón.

iii) No incorporado/organizado. Puerto Rico y las Islas Vírgenes (situadas en el mar Caribe) y Guam y las Islas Marianas del Norte (ubicadas en el océano Pacífico)

iv) No incorporado/no organizado. Por razones materiales, se enmarcan en esta categoría las islas, los atolones y los arrecifes a los que se hará referencia en el apartado 3. Por una razón formal (inacción del Congreso federal estadounidense), se incluye la Samoa americana. Este estatus jurídico-político supone que los habitantes de estos territorios son titulares de determinados derechos fundamentales, pero no de todos los derechos constitucionales.

De esta clasificación, aun esquemática, pueden extraerse varias cuestiones útiles para la reflexión jurídica. Seguidamente analizaremos la principal: ¿por qué existen esas diferencias organizativas?

2.2. LOS «CASOS INSULARES», UNA DISCUSIÓN UNIVERSAL

A lo largo de la historia de EE. UU. —particularmente, entre los siglos XVIII y XIX—, el país recurrió a diferentes estrategias para incorporar territorios. Al menos en retrospectiva, durante los primeros años de formación de la federación se colegía cierta teleología de incorporación y, en un estado final, el territorio en cuestión pasaba a ser un estado federado. Esta dinámica cambió a finales del siglo XIX y principios del XX, con la herencia de territorios ganados en las conquistas de 1898 de los restos del Imperio español. A partir de este momento, los «casos insulares» motivaron que la dimensión jurídico-constitucional irrumpiera en el debate militar. La expresión «casos insulares» hace referencia a la media docena de sentencias de la CS federal que dieron respuesta a la discusión sobre si los individuos que eran parte de territorios sin el estatus de estado federado tenían los mismos derechos que los que vivían en estos (Littlefield, 1901: 170). El núcleo de esos casos es de 1902-1922. Todas estas controversias presentan rasgos comunes: el objeto (temas comerciales, aunque con incidencia constitucional), y la ausencia de unanimidad de la decisión; salvo en *Huus*, los demás casos se resolvieron por una mayoría de cinco votos (contra cuatro) emitidos por los magistrados (Torruella, 2007: 300).

La cuestión central de este debate, que entronca con discusones jurídico-políticas universales, es si «la Constitución sigue a la bandera», es decir, si la conquista de un territorio lleva aparejada la inclusión de este al país conquistador como parte de pleno derecho. Es obvio que la primera respuesta es negativa a la luz de la rama del Derecho internacional humanitario que regula el Derecho de ocupación y los derechos y obligaciones de la potencia ocupante (Bermejo y López-Jacoiste, 2019). Ha transcurrido más de un siglo desde la inclusión en la institucionalidad estadounidense de estas discusiones y, dado que continúan generando polémica, hay razones para que una *juridificación del limbo* sea una afirmación proporcional, a pesar de sus tintes hiperbólicos.

En uno de los casos insulares, *Dowries vs. Bidwell* (1901), relativo a Puerto Rico, se acuñó la categoría de «territorio no incorporado» (recuérdense los binomios mencionados *supra*), entendido como aquel que, a pesar de que no es anexado con el fin de que pase a ser un estado parte (federado), pertenece a EE. UU. Sus habitantes, entonces, no son extranjeros, pero tampoco ciudadanos estadounidenses (Kaplan, 2005: 841). Esa sentencia fue básica para configurar el estatus jurídico político subalterno mediante la denominada, no sin discusión, «cláusula territorial». En el preámbulo de una propuesta de ley sobre descolonización de Puerto Rico (Ley para la descolonización inmediata de Puerto Rico, 7/2017, 2 de febrero de 2017), pueden leerse críticas acerbas a dicho término. Si la expresión «cláusula territorial» es considerada un burdo eufemismo de «cláusula colonial», las críticas a la sentencia denuncian que la resolución de la CS «fabrica» las categorías de territorio incorporado y no incorporado, y establece «un crucigrama jurídico» que legitima la discrimina-

ción racial y territorial. Asimismo, recuerdan la sentencia de la CS federal de *Puerto Rico vs. Sánchez Valle*, en la que la CS determinó que Puerto Rico no posee «soberanía primigenia», a diferencia de los estados de la Unión y las reservas indígenas. Ciertamente, es axiomático que, cuando se adjetiva a un Estado (libre, asociado, etcétera), es porque hay algo que lastra la naturaleza material del Estado, su capacidad de desenvolverse como tal entre sus supuestos pares políticos soberanos. Esto es, cuando se utilizan perífrasis como soberanía económica, cultural y similares se enmascaran las carencias de soberanía política. En el fondo, dice Torruella, el tinglado jurídico y jurisprudencial que ha acompañado a lo largo del siglo XX a los «casos insulares» pretendía «[…] quitar a Puerto Rico de la cobertura del artículo 73(e) de la Carta de las Naciones Unidas», parte del Capítulo XI (Declaración relativa a territorios no autónomos), para no tener que enviar reportes sobre el territorio portorriqueño (Torruella, 2018: 85). Se pretendía, así, evitar la propaganda negativa en un foro internacional, es decir, impedir que los representantes de algunos Estados equiparasen a EE. UU. a una nación todavía colonialista.

A pesar de fallarse en unos *hic et nunc* concretos (y en un marco axiológico indudablemente racista: el de los EE. UU. de inicios del siglo XX), la doctrina jurídica de los «casos insulares» se ha extendido a otros temas actuales, lo que robustece la tesis, sostenida en este apartado, de la universalidad de sus planteamientos, que puede contribuir a la reflexión sobre situaciones similares. Así, tal renacer doctrinal es azuzado en determinados debates contemporáneos que serán reconocibles por cualquier lector. Algunos, de inicios del siglo XXI, como la legislación de excepción en Guantánamo, reactivada tras el 11-S, y la preeminencia del gobierno estadounidense en la política ejecutiva militar. Otros, del primer cuarto de este siglo, como la discusión sobre los límites de los «derechos culturales» y la potestad del Congreso estadounidense (en la línea del artículo cuarto, tercera sección de la Constitución de 1787) para establecer distingos (incluso privilegios o «discriminación positiva») entre poblaciones adscritas a un «grupo cultural».

En un plano estrictamente jurídico, esta última cuestión se sustancia en una «cláusula cultural», esto es, en la «[…] oportunidad de reutilizar el marco [de los "casos insulares"] para proteger la cultura indígena de la imposición del escrutinio y la supervisión federales» (Capítulo tres, 2017: 1680). En *Tuaua vs. United States* (caso fallado el 5 de junio de 2015 por la Corte de Apelaciones estadounidense del circuito del Distrito de Columbia), se «[…] señala un conflicto entre la extensión de los derechos constitucionales individuales y la protección de la cultura territorial». La sentencia negó de nuevo que los samoanos americanos (recuérdese, un territorio no incorporado cuyos habitantes son nacionales estadounidenses, pero no ciudadanos) tuvieran la ciudadanía estadounidense por nacimiento. De hecho, las mismas autoridades samoanas estuvieron de acuerdo con esta afirmación. Adujeron que extender la ciudadanía a estos sujetos podría poner en peligra ciertas particularidades de la isla, que es, además, uno de los territorios no autónomos que por mandato de la ONU debería ser descolonizado, en el bien entendido que su estatus se define por la transitoriedad; por tanto, por el cuidado de procesos que no pueden equipararse a los Estados reconocidos por la comunidad internacional. Estas particularidades implican, por ejemplo, la protección contra «[…] los efectos que una sentencia de ese tipo tendrá en los

asuntos internos, cultura y autogobiernos samoanos». Sería un modo de oponerse a un adagio que señalase: *Si los samoanos pueden ser como un estadounidense más en EE. UU., entonces cualquier estadounidense —incluido un agente de bienes raíces— puede ser como un samoano en Samoa*.

Este potencial *laissez faire, laissez passer* de una disposición de recursos propiciada por la asimetría de los ciudadanos de la nación más poderosa del mundo se percibe claramente en los problemas que generan las compras y ventas de tierra. En *United states vs. Guam*, la CS describe la demanda del Departamento de Justicia de EE. UU. (28 de septiembre de 2017) contra una agencia de Guam (Comisión Chamorra de Fideicomiso de Tierras) que circunscribe a los nativos chamorros la adquisición de derechos de propiedad. La sentencia recurre al ejemplo de un hombre de origen africano casado con una chamorra. Construyeron juntos una casa, pero cuando la esposa falleció la Comisión Chamorra lo desalojó de la propiedad por razones étnicas, dado que no podía ser propietario por no ser biológicamente chamorro (Cuison, 2018: 127). Cuison sugiere que solo una *tribalización* (la cláusula de la «tribu reconocida federalmente») permite esta decisión, pero que no puede entenderse en un juego de suma cero a partir de la oposición entre lo político —no aceptable, por cuestionar la soberanía federal— y lo étnico — aceptable, si se presenta con el envoltorio cultural—. Cabe añadir que la aceptación oficial como «tribu» no es una cuestión de voluntad política, es decir, una suerte de un cabildeo para conferir relevancia a la tribu. Semejante perspectiva pasaría por alto otros aspectos históricos. Por ejemplo, ¿no tendrá que ver que tribus indias están en suelo estadounidense, y no se consideran ajenas a EE. UU. por no haber una pretensión de secesión o no poder plantearse —al menos de momento— en términos de la retórica de la «descolonización» en el marco del Derecho Internacional Público? ¿No se percibirá a estos como un problema propio, de identidad nacional, y a otros —por ejemplo, los habitantes de la Samoa americana o de las islas Marianas del Norte— como una cuestión de gestión de problemas históricos pasados, pero desde una perspectiva, innegociable por imperial, de geoestrategia? Debemos recordar, por añadidura, que ha habido traslaciones no automáticas en el propio ordenamiento jurídico estadounidense, analogías cuya pertinencia se ignoraba y giros políticos que deberían prevenir de cualquier respuesta lineal. Así, las previsiones que en principio se aplicaban a las tribus india, después se aplicaron a Estados con grupos raciales propios, como Hawái, y posteriormente a territorios, como la Mancomunidad de las Islas Marianas del Norte (Cuison, 2018: 142-143).

Podría aducirse que, según esta división competencial habría que *dar a lo federal lo que es de lo federal, y a lo local, lo que es de lo local,* por formular un adagio de raigambre evangélica que se deduce del actuar gubernamental. La tautología, palmaria, se explicita únicamente a efectos didácticos. Es decir, ¿qué es un «asunto puramente local»? Esta categoría no es clara y resulta difícil de deslindar de otros niveles (¿dónde empieza y termina la soberanía, si se quiere ver tan falazmente como una línea continua?). En los términos discursivos de la doctrina seguida aquí, «[…] todo el Derecho territorial es, en cierto modo, de carácter federal: el legislativo de Guam, por ejemplo, no puede aprobar leyes por su autoridad soberana independiente, sino que tiene el derecho a ello por delegación del Congreso». Por otra par-

te, las legislaciones locales «[…] pueden parecer, funcionar y sentirse como constituciones. En el caso de Guam, por ejemplo, su Ley Orgánica contiene una declaración de derechos y estructura el gobierno guameño» (Capítulo cuatro, 2017: 1713). Algunas cuestiones clave, entre ellas el manejo de la deuda de crisis en un territorio no incorporado, han sido decididos por la CS federal, no por su equivalente guameño (*Limtiaco vs. Camacho*), ello a pesar de que una sentencia reciente de la CS que «[…] invalidaba el plan de la isla para reestructurar sus obligaciones ha exacerbado claramente la crisis de deuda de Puerto Rico» (*ibid.*: 1724), lo que demuestra el sesgo que anida en la separación federal/local a la que se recurre al abordar el estatus polémico de estos territorios.

En resumen, en este apartado se ha explicado que los «casos insulares» no son únicamente jurídicos, aristas exclusivamente a resolver por jueces, sino que siempre conllevan cuestiones políticas. Parece adecuado que el Congreso tenga la última palabra. Lo que es más discutible es que la decisión se funde en una doctrina aplicada a un tiempo y lugar muy específicos —casi diríase, críticos para el orden interestatal en el que nos encontramos en el siglo XXI—, a saber, los EE. UU. de finales del siglo XIX y principios del XX. En todo caso, con la explicación presentada se han suscitado cuestiones relativas al modo en que el entramado jurídico-político que sostiene el imperio estadounidense debe establecer distintos grados de ordenación política administrativa y equilibrarlos con los derechos de los habitantes en esas entidades.

3. SOBRE ISLAS E IMPERIO: IDEAS LÍMITE DE LA SOBERANÍA

La figura jurídica de los «territorios no incorporados y no organizados» (*unincorporated/unorganized*) en el ordenamiento estadounidense es el hilo conductor que, en este apartado, permite reflexionar sobre otros términos jurídicos a un nivel básico, nuclear. Si bien algunos textos internacionales, entre ellos la Carta de las Naciones Unidas (26 de junio de 1945), parten del consenso sobre la prohibición de la adquisición de territorios mediante el uso de la fuerza establecida por el Derecho internacional, es posible observar en EE. UU. una geografía imperial eslabonada que se remonta a estados primigenios del Derecho que permiten visualizar ideas límite sobre la soberanía, la posesión o la propiedad privada, así como una flexibilidad para el despliegue territorial como potencia ocupante (Rittich, 2018). Por eso, se propone un análisis a partir de la comprensión del ordenamiento jurídico como un continuo con otros aspectos materiales —por ejemplo, la posibilidad fáctica de controlar efectivamente un territorio — que, en esencia, se opone a las teorías iusnaturalistas que subyacen a muchas de las argumentaciones jurídicas, no solo a las estadounidenses.

El *Libro de los hechos mundiales* (*The World Factbook*), publicado por la Agencia Central de Inteligencia, describe una serie de territorios federales no incorporados/no organizados (salvo el Atolón de Palmira, parcialmente privado y, como se explicará más adelante, «incorporado»). A su vez, estos lugares se reúnen en conglomerados de protección de la naturaleza: el denominado Monumento nacional marino de las islas remotas del Pacífico (*Pacific Remote Islands Marine National Monument*), que forman parte del Refugio nacio-

nal de vida salvaje del Pacífico (*United States Pacific Island Wildlife Refuges).* Dicho Monumento es controlado por el Departamento de Interior estadounidense en consulta con el Departamento de Comercio, salvo en lo relativo a la Isla Wake, que pertenece al Departamento de Defensa (The White House President George W. Bush, 2009: 4). En 2014, el presidente federal Barack Obama expandió el monumento para que comprendiese las áreas del anterior monumento desde «[…] las aguas y las tierras sumergidas de las islas Jarvis y Wake y el atolón Johnston», que estaban en el límite del monumento expandido «hasta el límite hacia el mar de la ZEE [Zona Económica Exclusiva]» estadounidense (*Pacific Remote Islands Marine National Monument Expansion*, 2014: 58647).

Los territorios analizados en este apartado son ocho (seis en la Polinesia, uno en la Micronesia y otro más en el Caribe), y han sido seleccionados porque su situación explica sobradamente la tesis planteada. Algunos siguen formando parte de EE. UU., y otros pasaron a otro Estados o tuvieron un devenir particular que será explicado en el rubro correspondiente. Tanto su pervivencia como parte de la taxonomía imperial como los cambios que han experimentado a este respecto servirán para entender la evolución del Derecho nacional,en relación con el internacional, así como el modo en que se adaptan las leyes públicas a las oleadas de la historia.

3.1. ISLAS BAKER Y HOWARD

EE. UU. tomó posesión de ellas en 1857 como reserva de guano, un fertilizante compuesto por nitrógeno, fósforo y potasio muy preciado como abono. A lo largo del siglo XIX, esta materia prima fue explotada principalmente por empresas estadounidenses y británicas. Ello dio lugar a giros estratégicos —a veces abruptos— en la relación entre el Derecho internacional Público y el de los Estados, así como a un renacer de las teorías sobre ocupación de territorios deshabitados y su relación con la insularidad, en correlación, usualmente, con su aislamiento.

Aunque el mejor guano (del quechua *wánu,* abono) se encontraba en las costas del Perú, EE. UU. colocó la materia prima guanera como eje de una expansión marítima moderna; de hecho, las «[…] islas guaneras fueron […] [sus] primeras adquisiciones territoriales más allá del continente […]» (Pierre, 2014: 1). Entre los movimientos que propiciaron el predominio a EE. UU. en esta carrera política y empresarial destaca la punta de lanza legislativa unilateral de la naciente potencia mundial. De acuerdo con la Ley de las islas de guano (Guano Islands Act, 1856), si un ciudadano estadounidense encontraba una isla no reclamada o deshabitada, y la ocupaba con el propósito de obtener guano, EE. UU. se adjudicaba la tenencia de su soberanía, si bien ello no implicaba (podía suceder, pero no necesariamente) que fuera para siempre. De hecho, la interpretación común era que esa situación se prolongaba hasta que terminaba de explotarse el guano y, una vez acabada esta tarea, volvían a ser susceptibles de ser redescubiertas y reocupadas. Es decir, recuperaban el estatus de *terra nullius* (*no man's land*, tierra de nadie).

Esta interpretación dio lugar a varios pronunciamientos doctrinales y jurisprudenciales. Por ejemplo, en una sentencia que ayuda a estructurar la argumentación de este apartado,

los demandantes enumeran una serie de territorios para que EE. UU. los reconozca como *tierra de nadie* y, por tanto, susceptibles de apropiación privada. Según la sentencia *Kingman Reef Atoll Development L.L.C. vs. United States* (The United States Court of Federal Claims, 2014: 8, 18 y 50), estos territorios eran las Islas Swan, Isla Swains, el Archipiélago Spitsbergen, la Isla Jan Mayen, las Islas Los Monjes y las Islas Aves. Mientras, otros espacios catalogados de igual modo, entre ellos los denominados «Cayos del Guano» (Bancos de Bajo Nuevo, Serranilla, Quitasueño, Serrana y Cayos del Roncador), fueron devueltos a Colombia en 1972 (Sandner, 1988: 106-107).

EE. UU. se opuso parcialmente a la doctrina de la tierra de nadie, por lo que, en el orden internacional, el estatus de cada uno de estos territorios discutidos ha terminado siendo distinto:

i) Las caribeñas Islas Swan (Cisne, Santanilla o, según la denominación colombina de 1502, Isla de las Pozas) conforman un archipiélago de tres islas (Cisne Grande, Cisne Pequeño y Cayo Pájaro Bobo) y dos bancos, es decir, atolones sumergidos (Banco Rosario y Banco Misteriosa, este último descubierto por el navegante Tomás Nicolás de Villa en abril de 1787) que se devolvieron a Honduras en 1972. Estuvieron en posesión estadounidense durante más de un siglo (134 años), y allí se instaló un aeropuerto militar, además de estaciones de radio, telecomunicaciones, radar y propaganda en el marco de la Guerra Fría. Recuérdese, al respecto, la tutela de los contras nicaragüenses para desestabilizar a ese país centroamericano y las operaciones de la CIA en la región que abarcaban esas islas (Sandner, 1988: 108).

ii) Respecto a las Islas Swains, se trata de un atolón coralino situado a 359 km de Pago Pago, capital de la Samoa americana. Cercano históricamente al archipiélago de Tokelau (que depende administrativamente de Nueva Zelanda), el gobierno tokelauense reivindica, incluso, en un borrador constitucional de independencia —sin vigor— las Swains, a las que denomina «Olosega» (Agencia Central de Inteligencia, 2024).

Tiempo atrás atrás, quienes en el siglo XIX pretendieron el título de propiedad sobre estas islas (la familia Jennings, que aún hoy poseen derechos) se dirigieron a las instituciones estadounidenses. Estas convalidaron el título imperfecto o rudimentario y tomaron la decisión soberana de considerarla privadas, a diferencia de lo que ocurrió en el caso del arrecife Kingman, que al menos desde 1934 es propiedad pública. Anecdóticamente, un ciudadano estadounidense llegó a este atolón en 1856 (el mismo año en el que se aprobó la Ley sobre las islas guaneras) y lo reivindicó para EE. UU. hasta que en 1925 pasó a ser parte de la Samoa americana (U.S. Census Bureau, 1994: 14). Todavía hoy está habitado por una población que, de acuerdo con los censos estadounidenses, consta de una quincena de individuos —¿alguno de la familia Jennings? ¿En qué condiciones?—; envían un representante al congreso samoano, pero no tiene derecho a voto (U.S. Census Bureau, 1994: 17). La elección de representante «se decide por reunión pública en la isla de Swains» (Agencia Central de Inteligencia, 2024).

iii) El Archipiélago Spitsbergen, «[…] un racimo desolado de islas áridas» situado en el Océano Ártico (Ederington, 1997: 284), tuvo, *de facto,* la condición de tierra de nadie hasta que se descubrió carbón en 1900. Entonces, varias compañías dedicadas a explotar esa materia prima —algunas de ellas, estadounidenses— lo reclamaron. Veinte años después, se adjudicó la soberanía a Noruega con la condición de que este país reconociera los derechos adquiridos por ocupantes (individuos y empresas) mediante un tratado suscrito por los nueve Estados con intereses en esas islas: además de EEUU (vía la Artic Coal Company), Dinamarca, Francia, Gran Bretaña, Italia, Japón, Noruega, Holanda y Suecia. En 1924 se unió un décimo Estado, Rusia. Una fórmula similar se utilizó cuando Gran Bretaña cedió Oregón a EEUU (15 de juio de 1846), si bien el respeto a las posesiones de ciudadanos británicos nada tiene que ver con el hecho de que Oregón fuese, en ese momento, tierra de nadie (Ederington, 1997: 284-285).

iv) Respecto a las islas guaneras de Los Monjes y de Aves (a las que EE. UU. sí catalogó explícitamente como tierra de nadie), Venezuela reclamó ese «conjunto de rocas estériles» cuando se inició la explotación de guano, y expulsó a las compañías estadounidenses que explotaban ese bien, aunque años después las indemnizó tras la firma de un tratado bilateral entre ese país y EE. UU. (Ederington, 1997: 288; *The United States Court of Federal Claims*, 2014: 75).

v) Finalmente, la decisión de un tribunal noruego sobre la Isla Jan Mayen (poseída, tras ser catalogada como tierra de nadie, privadamente) determinó que el país europeo debía indemnizar a su propietario privado (Jacobsen vs. Norwegian Government). Aun así, es difícil sostener que esta solución es aplicable al resto de lugares similares. La razón es que el propietario siguió un curso legal para perfeccionar su título, algo que, por ejemplo, no hicieron los supuestos propietarios en otras disputas (*The United States Court of Federal Claims*, 2014: 76). Ederington también menciona este supuesto —de hecho, la sentencia incorpora algunos de sus postulados— como ejemplo paradigmático de un propietario que, supuestamente, estaría por encima de la soberanía, planteamiento recurrentemente esgrimido por quienes postulan derechos de propiedad basados en el iusnaturalismo, al estilo Robinsón Crusoe. Sin embargo, tales elucubraciones parten de la comprensión de esta materia como una tabla rasa susceptible de apropiación por átomos individualistas y desconectados del contexto político en el que, necesariamente, ocupan la tierra, la organizan y la conservan. Aunque la sentencia hizo algunas concesiones a esos argumentos, señaló que, en términos generales, son los Estados, no los individuos, los que pueden adquirir territorios (¡no propiedades!) de nadie o tolerar su adquisición, e hizo una mención específica a las tierras submarinas y los mares, que, se presupone, pertenecen al Estado salvo habilitación expresa de cambio de propiedad. Es más, la resolución judicial señala que la *traslatio imperii* —un término que tiene tal relevancia que se utiliza en debates históricos de más largo alcance sobre la soberanía, la estatidad y la estatalidad (Vélez, 2013)— funcionó siguiendo esta hilación concatenada: lo tuvo la corona de Inglaterra sobre territorio actual de EE. UU., después pasó a las

colonias, de estas a los estados y finalmente al Gobierno federal, al Estado en sí (*The United States Court of Federal Claims*, 2014: 51-54). Este es un argumento contra las traslaciones individuales, contradictorias en sí mismas, dado que les es consustancial a la figura la subdivisión *ad infinitum* del territorio.

El tribunal estadounidense señala que, si un territorio desbabitado *(uninhabited country)* es descubierto por varios individuos que no tienen ningún vínculo con ningún Estado ni le deben lealtad (hipótesis de laboratorio: ¿serían apátridas que renuncian a su nacionalidad para poder ser nacionales de un nuevo país?), pasa a tener un título en común sobre el territorio y se lo dividirán con base en este axioma: «El título de toda la tierra está en toda la sociedad» *(The title of the whole land is in the whole society).* Ederington señala que este planteamiento es abiertamente contradictorio, dada la dificultad de concebir la existencia de individuos particulars que, esgrimiendo sus intereses privados, oponen lo logrado a un mundo anegado de Estados (1997: 278). Es, en fin, una hipótesis de laboratorio, como la de los terroristas que atacan a un Estado desde una *terra nullius* porque es la única manera de justificar una hipotética costumbre internacional de legítima defensa preventiva (Pastor Ridruejo, 2004: 109).

Más matizada es la postura de Charles Hyde, que señala dos casos viables de *terra nullius*. Por una parte, una tierra tan alejada que, a pesar de que no pertenece a individuos privados, el Estado delega o tolera que la posean (quizá esto fue lo que sucedió en el arrecife de Kingman). Por otro lado, las propiedades situadas en territorios que no están estatalizados (tribales, por ejemplo, o lugares de paso de tribus nómadas) y para los que, se entiende, funcionaría la figura de la tierra de nadie a efectos de la protección del Estado a esos colonos, ello a pesar de la contradicción que supone reforzar la propiedad de los colonos y que, al mismo tiempo, el país conquistador suprima las potestades públicas (y puede que las privadas) de los indígenas, un proceder que fue convalidado por la Conferencia de Berlín (1884-1885) y que tenía precedentes en los protectorados de Alemania sobre las Islas Marshall y de Gran Bretaña sobre las Gilbert, entre otros casos (Hyde *apud* Ederington, 1997: 280).

De hecho, esta última parece ser la vía a través de la que EE. UU. sugirió a Noruega proteger los derechos de propiedad de la multinacional Polarfront, propiedad de un ciudadano de ese país y que construyó un par de granjas de zorros cuando la isla disputada era no *man's land* (Ederington, 1997: 282-284). Quizás la contradicción pueda vadearse considerando que la protección no prejuzga un mejor título de propiedad, sino que, simplemente, protege un estado de hecho revestido de la protección de derechos civiles. En todo caso, se constata de nuevo que la cadena de la traslación de la soberanía es todo menos eslabonada.

La Segunda Guerra Mundial abortó un intento de colonización desregulada de las islas Baker, Howard y Jarvis (que serán analizadas en el siguiente apartado). En este caso, un «[…] pequeño número de colonos voluntarios hawaianos […] que ocuparon las islas de 1935 a 1942 para ayudar a asegurar el reclamo territorial de Estados Unidos» (*Pacific Remote Islands Marine National Monument Expansion*, 2014: 5864). Desde 1936, son administradas por el Departamento del Interior de EE. UU., y en 1974 pasaron a ser parte del refugio de vida salvaje, tras la erradicación de los gatos, ejemplo recurrente y feroz de fauna

invasiva. Así, entre 1937 y 1965, en el marco de un intento de colonización de Nueva Nantucket (*New Nantucket*, emplazamiento sugestivo para quienes hayan leído la novela *Moby Dick*, fue uno de los anteriores nombres de la isla Baker, según recoge la Enciclopedia Británica), acabaron con las aves indígenas; en la Isla Howland, los felinos acabaron con las ratas negras (1854-1937), y aquellos terminaron con la fauna local hasta su erradicación medio siglo después (Agencia Central de Inteligencia, 2024).

A pesar de la escasa población de las islas Baker y Howard, están registradas en el censo de modo separado, si bien hasta 1940 aparecían como parte de Hawái (U.S. Census Bureau, 1994: 4). En todo caso, las islas Baker y Howard han conservado un vínculo cultural y administrativo con el estado número cincuenta (Rauzon, 1985: 4).

3.2. ISLA JARVIS

Descubierta por los británicos en 1821 (pero mantenida deshabitada), EE. UU. la anexionó en 1858 para explotar el ansiado guano. El agotamiento de esta materia prima motivó su abandono en 1879. Reino Unido la anexionó en 1889, pero no realizó ninguna acción material de ocupación, por lo que EE. UU. la ocupó y en 1935 la reclamó. Abandonada en la Segunda Guerra Mundial, pasó a ser reserva de vida salvaje en la misma fecha que las anteriores, si bien con problemas de fauna invasiva (los incansables gatos, que buscaban erradicar la invasión previa de las aún más incansables ratas). Los felinos fueron introducidos en un intento de colonización (la de 1935), y no pudieron ser erradicados sino hasta la década de los noventa (Rauzon, 1985). Así, también en la Isla Jarvis se cumplió el axioma siguiente: «Los gatos, a lo largo de la historia, han contribuido a la desaparición de al menos el 14 % de las especies de vertebrados que se han extinguido en las islas» (Belenguer, 2014: 115).

3.3. ATOLÓN DE JOHNSTON

Según el *Diccionario de la lengua española*, un atolón es una isla madrepórica —es decir, formada por restos de corales— que tiene forma de anillo y cuyo centro es una laguna interior; suele comunicar con el mar por pasos estrechos. Su origen es volcánico; se trata, específicamente, de un volcán erosionado que forma la base de la laguna.

El atolón de Johnston fue anexionado tanto por EE. UU. como por el Reino de Hawái en 1858; en esos momentos, lo que hoy es el estado federado número cincuenta era un protoestado apoyado por Reino Unido tras las unificaciones tribales verificadas en el siglo XIX y que EEUU protegió, para acabar incorporándolo en 1893. Hasta finales de 1880, EE. UU. explotó los depósitos de guano del atolón.

El atolón pasó a ser refugio de especies en 1926 y en las décadas de los 50 y 60 del siglo veinte su superficie se utilizó para realizar pruebas nucleares. De hecho, la Agencia de Defensa Nuclear (Defense Nuclear Agency) lo administra desde 1973, año en el que llegó a un acuerdo con la Fuerza Aérea estadounidense para hacerse cargo del territorio. Hasta 2005 continuó utilizándose para almacenar y destruir armas químicas. Mientras se le da

otra utilidad, «[…] el mar defensivo naval de tres millas a su alrededor permanece bajo la jurisdicción y el control administrativo de la Fuerza Aérea de los Estados Unidos» (Agencia Central de Inteligencia, 2024).

Respecto a los problemas con la naturaleza, habituales en estos lugares refractarios a la asepsia jurídica, cabe señalar que «[…] un enjambre de hormigas anoplolepis […] invadió la isla en 2010 dañando la vida silvestre nativa». Según esta fuente, todavía continúan las tareas de erradicación. De las cuatro islas del atolón, solo la isla Johnston está habitada, y su población se cuenta separadamente desde 1950, aunque en el censo de la década anterior —tal y como sucedió con las Baker, Howard y Jarvis, y, como veremos, con Midway, Atolón de Palmira— estaba registrada como parte de Hawái (U.S. Census Bureau, 1994: 4).

3.4. ARRECIFE DE KINGMAN

Anexionado por EE. UU. en 1922 —una afirmación que, como vamos a ver inmediatamente, amerita muchos matices—, «[…] su laguna protegida sirvió como estación de paso para los hidroaviones que volaban de Hawái a la Samoa Americana» a finales de 1930. A principios del siglo XXI, se estableció la protección de las aguas que rodean el arrecife (hasta doce millas náuticas), que está continuamente inundado. Una empresa pesquera presentó una demanda en la que aseguraba poseer el atolón y solicitaba una indemnización (de aproximadamente 54 millones de dólares) al gobierno de EE. UU. por haberle arrebatado la propiedad el 18 de enero de 2001 para convertirla en un refugio de vida salvaje. De hecho, la empresa demandante alegó que la familia propietaria del atolón Palmira, los Fullard-Leo (una familia de Honolulú, contratistas de techado, *roofing contractors),* también poseía ese arrecife desde 1922, y que algunos de ellos fueron parte del consejo de administración de la pesquera (The United States Court of Federal Claims, 2014: 2).

En el texto de esta demanda, a la que se ha aludido en varias partes de este artículo, el actor señala que la familia y su representante ostentaban la posesión pública e ininterrumpida del arrecife, argumento que seguramente remitía a principios como los que rigen la usucapión o incluso a los «usos y costumbres» para consolidar derechos en un ámbito local. Así, actuaban como propietarios en el atolón Palmira, desalojaban a quienes pescaban ilegalmente, realizaban inspecciones áereas de vigilancia (dando lugar a que los miembros de la familia se relacionasen con el arrecife desde los aires, sin haber puesto el pie allá en años), etcétera. Incluso con el consentimiento de EE. UU., en una declaración jurada el representante inmobiliario de esta familia señaló que viajó en avión para observar si las áreas eran adecuadas para que algunos residentes desplazados del atolón Bikini (en las Islas Marshall) fueran realojados allá y en el atolón de Palmira. Este mismo representante afirmó que quienes querían visitar el lugar eran referidos desde la base naval de Pearl Harbour (The United States Court of Federal Claims, 2014: 4-7).

Todos estos indicios deben tenerse en cuenta para resolver la cuestión de quién ejerce el dominio en disputa. No obstante, debe advertirse negativamente contra cualquier linealidad (del tipo «la propiedad era, irrevocablemente, privada» o «EE. UU. siempre poseyó ese lugar, a pesar de las apariencias») para no incurrir en dicotomías falaces. Es decir, al igual

que las autoridades estadounidenses acompañaban a los dueños privados cuando estos ejercían su dominio, también había actos que mostraban inequívocamente la voluntad (aquí casi fusionada con la capacidad) para fijar la propiedad en los términos que el Estado quisiera. Esa asimetría es la clave: el hecho de que alguien disfute de derechos de propiedad privada no significa que estos sean *legibus solutus*, ya que es imposible que la titularidad de ese derecho no se haya disuelto en el marco jurídico de algún Estado —que, por siglos, sabemos que sí puede ser *legibus solutus*—. Todo ello a pesar de que en la fecha clave de 1922,una compañía vinculada a esa familia tomase posesión del arrecife, en nombre de EE. UU, con todo el simbolismo y la parafernalia al respecto que demandan los ejemplos reales o ficticios de las apropiaciones de territorios. Estos los que se evocan al releer Las Leyes de Platón, las reflexiones teóricas sobre el nomos o leyes de la tierra (Schmitt, 2005) o las deducciones de lo que sería la Nueva España a partir de las primeras poblaciones reales de los territorios (Vargas y Machuca, 1599: Libro cuarto). En seguimiento a este tipo de formalidades —indiciarias, más que probatorias—, el representante de la familia recogió en 1926 tales restos y los despositó en los Archivos de la Comisión del Territorio de Hawái con el conocimiento de autoridades estadounidenses (The United States Court of Federal Claims, 2014: 13). Cuando dos ciudadanos de Hawái tomaron posesión del atolón de Palmira en nombre del rey de la isla (The United States Court of Federal Claims, 2014: 40), enterraron una botella en un guiño al mito del náufrago o, si se me permite la maledicencia, a la embriaguez (tal que *hybris*) de pretender encerrar en un objeto las demandas de soberanía.

No es cuestión de negar de plano esos derechos, como hace Case, con el sesgo evidente de la parte interesada, cuando busca argumentos para señalar que los Fullard-Leo no poseían el arrecife (Case, 2001). Cuestión distinta era el consenso o la inercia de la burocracia estadounidense —al menos, cuando había que tomar una decisión sobre compras, sobrevuelos, visitas, etcétera —, que toleraba esa propiedad, pero reservándose siempre un derecho de alzada, de veto, por así decirlo, dependiendo de la forma de los actos y su prolongación. No hay otra certeza del poder decisorio estadounidense —reseñable, no por menos obvio— que sus afirmaciones *inexorablemente ondulantes*, si se me permite la expresión, recogidas en el dictamen de la FWS Kingman Reef Ownership Status and Federal Jurisdictions (30 de marzo del 2000)[1]. Allí reconoce que, a lo sumo, esa familia posee un título imperfecto o rudimentario (*imperfect or inchoate rights*) que necesita de su perfección o robustecimiento por una acción expresa del Gobierno de EEUU, y ello a pesar de la «creencia de buena fe» de los Fullard-Leo de que ellos poseían ese territorio. Una buena fe que se diluyó cuando no respondieron a la orden ejecutiva de Franklin D. Roosevelt (Executive Order núm. 6935, 29 de diciembre de 1934) en virtud de la cual el arrecife se convertía en suelo público, silencio que, mças allá de algunas acciones puntuales, se prolongó entre 1938 y 2000 (The United States Court of Federal Claims, 2014: 66-69). En cambio, el Estado, en ejercicio de su soberanía, decidió que el arrecife pasara al Departamento del Interior estadounidense, que lo reconvirtió en la reserva a la que se ha hecho referencia arriba.

En esencia, a pesar de algunos argumentos en contra (en 1922, EE. UU. no la consideraba una isla de guano, lo que tuvo consecuencias para su posesión, como se confirmó en 1933), de la orden ejecutiva de Roosevelt se deduce que las autoridades imperiales esta-

dounidenses empezaron a ver su valor estratégico militar en la ruta del Pacífico… Pero sin conocer bien su génesis histórica y legal. Por su parte, la familia Fullard-Leo quería poseer las islas como bienes inmuebles —su administrador era un empresario de bienes raíces—, es decir, tener la propiedad, pero no su soberanía ni, por tanto, su jurisdicción. En el conflicto con la Secretaría de Marina que tuvo lugar en los años treinta del siglo XX (década prebélica) predominó esa ambigüedad, aunque la Armada estadounidense la catalogó —erróneamente— como isla guanera y señaló que era de EE. UU. (The United States Court of Federal Claims, 2014: 15-17 y 21-22).

Según The Nature Conservacy (TNC), organización que a finales de los noventa compró el atolón de Palmira a esa familia y que está presente en otros territorios, el representante legal de la misma también les ofreció el arrecife Kingman, pero lo rechazaron porque no era de su interés. En cualquier caso, sí es pertinente subrayar que la compra de atolón Palmira (se detalla *infra*) tuvo un impacto en el arrecife Kingman, ya que supuso que el Departamento estadounidense de Pesca y Vida Salvaje (U.S. Fish and Wildlife Service, FWS, adscrito al Departamento del Interior) se hiciese con este como parte de un área de protección alentada por el TNC, precisamente en conjunto con ese área gubernamental (The United States Court of Federal Claims, 2014: 26).

3.5. ATOLÓN DE PALMIRA

Compuesto por más de cincuenta islotes, el Reino de Hawái lo reclamó (1862) y EE. UU. lo incorporó a las islas hawaianas en su anexión (1898). Sin embargo, la Ley de estatidad hawaiana (Hawaii Statehood Act, 1959) no lo incluyó; de hecho, la tantas veces citada familia Fullard-Leo pagó impuestos por la propiedad inmobiliaria al gobierno de Hawái entre 1922-1959. Además, la principal diferencia respecto al caso del arrecife de Kingman fue que el gobierno hawaiano nunca lo consideró una propiedad pública que debía mantener, e incluso convalidó reiteradamente la propiedad privada de sus poseedores (The United States Court of Federal Claims, 2014: 11 y 73-74). Su población fue contabilizada como parte del área censal hawaiana hasta 1940, y hasta el censo de 1970 no apareció como área censal independiente (U.S. Census Bureau, 1994: 5). Además, su ubicación, a medio camino entre Hawái y la Samoa americana, en el Pacífico, lo convierte en un entorno especialmente aislado. De hecho, en la puesta en marcha de su habitabilidad fue clave un individuo con experiencia previa en un territorio tan hostil como Alaska (Case, 2001: 294).

Actualmente, el lugar se ha convertido en un refugio para fauna salvaje y es, parcialmente, propiedad de TNC, según remarca la Oficina de Asuntos Insulares, del Departamento de Interior (s.f., *c*). Llama la atención que el atolón contaba con selva; no obstante, «[...] el cultivo de copra, las ratas negras y la deforestación de los militares estadounidenses durante la IIGM lo redujeron a un par de parcelas». Paulatinamente recuperado (las ratas fueron erradicacadas en 2011), «[…] está en marcha un importante proyecto de restauración de la selva nativa», según se lee en la web de la TNC.

La presencia de esa organización privada, TNC, no es novedosa en el territorio. Por lo que parece, el interés científico viene de larga data (1842, 1874), y se concretó en numerosas expediciones estadounideses. El año 1913 fue relevante para el atolón, dado que el juez Henry E. Cooper, «[...] en ese momento [desde 1912] dueño de Palmira», invitó a un científico a realizar un inventario (Wegmann, 2019). El hecho de que la posesión privada fuera ostentada por un juez —a pesar de ello, el atolón fue clave para derrocar a la última reina de Hawái— ilustra la interrelación entre lo público y lo privado no solo en EE. UU., sino en cualquier estatalidad naciente, como muestra el Derecho novohispano o como, más atrás, puso en evidencia el Derecho de la antigua Roma. A la anexión por esos particulares (1862) siguió la adquisición de la propiedad del atolón por el rey de Hawái y, cuando el reino pasó a república (1894) y esta a ser territorio de EE. UU. (1898), también lo habría sido el atolón (The United States Court of Federal Claims, 2014: 41). Según el Departamento de Interior de EEUU (s.f. *c*), en 1922 Cooper lo vendió a los Fullard-Leo por 15 millones de dólares, excepción hecha de dos islas.

Por tanto, fue la familia Fullard-Leo la que en 1999 vendió el atolón a TNC. Prefirió esa opción a venderla para que fuera utilizado como depósito de desechos de combustible nuclear, a que se construyera un casino o a que se usara para pesca masiva. El precio, 37 millones de dólares, estuvo lejos de los 47 que solicitaba la familia inicialmente, y, según la web del TNC, e incluyó 680 acres de tierra (2.75 km² distribuidos en 25 islotes del medio centenar en que consiste el atolón) y 15 512 de arrecifes (62.77 km²). Hay un islote, la isla Cooper (la más grande del atolón), que todavía pertenece a los descendientes del juez, como puede leerse en la web Palmyra Atoll Digital.

Los Fullard-Leo habían poseído el lugar durante ochenta años (en 1922 lo compraron por 70 000 dólares). Hubo disputas por la propiedad, e incluso la CS estadounidense estuvo implicada, y las controversias se resolvieron a favor de la familia, amparada por el hecho de que el atolón hubiera sido reiteradamente excluido de la propiedad pública hawaiana —ni siquiera en la Ley de estatidad de 1959 la incluía—, ello a pesar de que dirante la Segunda Guerra Mundial las fuerzas armadas estadounidenses le dieron un uso militar.

La cuestión de la venta tiene más aristas —al fin y al cabo, EE. UU. tutela el lugar y es la fuerza a la que se pediría cualquier ayuda—, ya que se llegó a hablar de las presiones ejercidas por algunos congresistas o por el Gobierno estadounidense para evitar que se vendiera y se convirtiese en un lugar para los «piratas de los residuos nucleares», así como de presiones para orientar la compra. A su vez, TNC vendió una parte (por unos ocho millones de dólares) al Departamento del Interior estadounidense para que fuera un refugio de fauna salvaje (Gomes, 1998; Purdum, 2000).

El resto del atolón es propiedad del Gobierno federal. El atolón de Palmira constituye un ejemplo práctico de cómo sondear la dialéctica isla/roca, nacida de este silogismo: «Todas las rocas [...] son islas, pero no todas las islas son rocas», uno de los ejes del Derecho internacional del mar (Moscoso de la Cuba, 2013: 257). Desde 2001, se gestionan federalmente las lagunas y las aguas territoriales (las mencionadas doce millas náuticas alrededor del atolón) como parte de un área nacional protegida. Case detallaba el proceso de cabildeo para obtener el atolón, que empezó ya en 1991, y las aristas legales que implicaba: «Pro-

piedad de las tierras sumergidas, extensión del mar territorial, leyes de pesca, registro de títulos, peligros físicos, regulación gubernamental, la licencia de pesca que actualmente grava el atolón, [...] operaciones de conservación [ecologistas]» (Case, 2001: 292). En todo caso, es especialmente clave para el estatus del atolón el hecho de que, cuando la familia Fullard-Leo litigó, como se ha señalado anteriormente, contra la ocupación militar en la Segunda Guerra Mundial, la CS falló a favor de la familia con un presunto *deus ex machina* de título de la propiedad, una «presunción de concesión perdida» (*lost grant*) en virtud de la cual, «[...]dado que todos habían considerado privado el atolón durante 80 años, el rey [de Hawái] debió haber otorgado la empresarios originales una patente de la tierra, que luego se perdió» (Case, 2001: 293).

En realidad, la CS infiere que el título existía porque la persona que lo habría otorgado —el rey de Hawái— la podía transmitirlo *de facto* a los particulares. Por lo que, dada su personalidad jurídica, deduce el título *de iure*, y sostiene que quien podría oponerse a él (el gobierno de EE. UU.) nunca lo hizo ni por los hechos ni en Derecho (la sentencia dictada así lo prueba). *Lost grant* que, por cierto, no aplica al arrecife Kingsman materialmente —puesto que nunca fue parte de Hawái— ni formalmente, pues la posesión ininterrumpida no se prolongó durante veinte años (The United States Court of Federal Claims, 2014: 41-43 y 80-81).

En cualquier caso, ¿se trata como dice Case, del «[...] único territorio en EE. UU. de propiedad totalmente privada y que no es parte de ningún Estado»? (2001: 294). Cierto, siempre que no se entienda que este solo hecho lo hace soberano, algo con lo que la propia representante del TNC estaría de acuerdo. Considerarlo territorio soberano equivaldría, si se me permite la paradoja, a *creer que se puede erigir un edificio a base de puntos ciegos*. Si acaso, se ha logrado que un estado en precario, transitorio, vire hacia un cambio de propiedad que se difiere, y que, además, es susceptible de compraventa, al no haber sido reclamado por un Estado desaparecido, el Reino de Hawái (que ahora es parte de otro Estado) y gestionado por EE. UU. de un modo directo en algunas partes, mientras que en otras delega esa potestad. ¿Acaso alguien —ni siquiera un jurista— podría pensar que, si hubiera una crisis —como lo fue la Segunda Guerra Mundiaal— que obligase a poner en marcha todos los recursos reales o potenciales de EE. UU., ese Estado no retomaría la soberanía?

Finalmente, debe hacerse un matiz importante. Hawái (y, por tanto, el atolón de Palmira como parte del archipiélago) fue un territorio no incorporado entre el periodo que media entre el 12 de agosto de 1898 y el 30 de abril de 1900); posteriormente, se les aplicaron las leyes estadounidenses y pasaron a ser territorio incorporado. Cuando, a mediados del siglo pasado, Hawái pasó a ser un estado de EE. UU., se excluyó el atolón, aunque este mantuvo el estatus de territorio incorporado. Sería, entonces, el único de esta naturaleza entre las catorce áreas insulares, por lo que, de acuerdo con la terminología administrativa estadounidese, tan formalista, se trataría de un «Territorio», en mayúsculas, mientras que al resto se los denominaría «territorios» en minúscula (Departamento de Interior de, s.f. *c*).

3.6. ISLAS/ATOLÓN MIDWAY

Las islas Midway (aunque también se las conoce como atolón) forman parte del monumento marino Papahanaumokuakea y del refugio de vida salvaje hawaianos. EE. UU. tomó posesión de ellas en 1867. Habían sido avistadas por Miguel Zapiaín Valladares en 1799, que las llamó islas Patrocinio (Landín, 1992: 35). En 1903 contaba con residentes que se instalaron allí aprovechando el tendido del cable transpacífico. Las islas Midway fueron una pieza clave en la Segunda Guerra Mundial, dado que entre los años 1935 y 1947 se utilizaron «[…] como parada de reabastecimiento de combustible para vuelos transpacíficos». Hasta 1993 fueron una estación naval (Agencia Central de Inteligencia, 2024). En los censos que se elaboraron desde la primera década del siglo XX hasta los años 40, los habitantes se computaron como parte de Hawái, pero desde 1950 se cuentan separadamente (U.S. Census Bureau, 1994: 5).

3.7. ISLA DE NAVAZA

La isla caribeña de Navaza fue descubierta por marineros que acompañaban a Cristóbal Colón en 1504 (Agencia Central de Inteligencia, 2024). La denominaron «Navaza» por la palabra española «nava», esto es: «Tierra sin árboles y llana, a veces pantanosa, situada generalmente entre montañas» (*Diccionario de la lengua española*). A consecuencia de la Ley del Guano (1857), «[...] un representante de la empresa Baltimore Fertilizer tomó posesión de la isla en nombre de los EE. UU.». Después de que esas operaciones fueran encomendadas a la empresa Fosfatos de Navaza en 1889, diez años más tarde cesaron todas las operaciones empresariales y la isla volvió a estar abandonada (en 1976, incluso dejó de operar el faro administrado por la guarda costera estadounidense). Desde finales de 1999, el U.S. Fish and Wildlife Service lo administra como refugio natural (Departamento de Interior, s.f. *b*)

Desde el punto de vista geoestratégico, la isla está ubicada en el canal de Jamaica, al este de esta isla, al sur de Cuba —el punto terrestre más cercano es la bahía de Guantánamo, a 160 km— y al oeste de Haití. Fabio Spadi abre la perspectiva de la discusión filosófico-jurídica al recordar que en 1858, un año después de la ocupación estadounidense, el entonces dizque imperio haitiano (Second Empire d'Haïti, regido por el autoproclamado emperador Faustino I) envió un par de barcos de guerra para someterla. Ya en 1855 habían intentado invadir, infructuosamente, la República Dominicana, es decir, la otra parte de la isla La Española, que había oscilado entre una breve secesión como Haití español, la anexión al proyecto bolivariano de la Gran Colombia y, finalmente, la constitución como el país actual. Las acciones del Gobierno haitiano pretendían imponer a quienes buscasen guano en Navaza que le pidieran permiso como supuesta autoridad soberana. Ante ello, un buque estadounidense se dirigió a Puerto Príncipe, capital haitiana, y anunció que seguiría en Navaza para proteger a sus conciudadanos.

A su vez, desde inicios del siglo XX, y a lo largo de las décadas, Haití no solamente incorporó a sus constituciones la isla de Navaza como territorio propio, sino que realizó acciones personales y reales que evidenciaban su continua reivindicación (Spadi, 2001: 115). Esta pretensión se mantiene en la actualidad: el artículo 8 a) de la vigente Consti-

tución haitiana declara que la isla forma parte del terrritorio nacional, y las autoridades le han asignado una bandera no oficial. Además, el Gobierno ha construido una iglesia para los pescadores haitianos que faenan en sus aguas (Pierre, 2014: 38), si bien la isla sigue deshabitada. A pesar de ello, Spadi argumenta que esa reivindicación tiene, al menos, un par de puntos débiles en su flanco jurídico: la ahistoricidad y la carencia de derechos materiales. Respecto al primer punto, cuando se proclamó la independencia de Haití, Navaza no era considerada parte del país, a diferencia de otras islas (Spadi, 2001: 117). Respecto al segundo, son indudables tanto la asimetría político militar entre Haití y EE. UU. como los problemas recurrentes de gobernabilidad del primer país que dificultan el control material de su propio territorio reconocido oficialmente.

Además, Spadi se apoya en el análisis los «títulos de propiedad» para sostener que la isla no forma parte de Haití. Se trata, en esencia, líneas básicas y realistas relativas al modo en que se interpreta el orden jurídico internacional para poder reclamar un territorio, de ahí la necesidad de comentarlos contrastándolos con otras situaciones:

i) *Unidad territorial*. Se trata de una cuestión histórica, aunque también puede relacionarse con la continuidad geográfica (Spadi, 2001: 119). Al parecer, la excolonia Haití (Estado desde 1804) no había reclamado la isla como parte de su territorio. No obstante, los sucesivos gobiernos haitianos no han reconocido que pertenece a EE. UU.. A su vez, EE. UU. solamente la quiso reclamar en virtud de la ley guanera, si bien ciertos aspectos geoestratégicos concretos refuerzan su interés por la isla. Entre estos, cabe citar la apertura del canal de Panamá en 1914; la cercanía con Guantánamo (enclave actualmente revalorizado por su funcionalidad en las «guerras» contra el narcotráfico, el terrorismo o la inmigración ilegal). A estos factores geoestratégicos hay que otros de índole general, entre ellos el dominio imperial talasocrático sobre el Caribe. Se trata, sin embargo, de vindicaciones genéricas, de lineamientos que en muchas ocasiones chocan con argumentos que los desacreditan; por ejemplo, la falta de intención (*animus*) que mostraron las autoridades de Haití cuando, en el seno de la Comisión del Canal del Istmo, institución estadounidense que supervisó la construcción del canal de Panamá, rechazaron la soberanía o la posesión de la isla de Navaza (Pierre, 2014: 34). Eso sí, como se ha visto, al igual que la soberanía es distinta de la posesión, hay otras formas de influencia (propiedad, jurisdicción) que entran en liza.

ii) *Contigüidad*. La contigüidad hace referencia a los derechos —hipotéticos— de los que es titular del Estado más cercano al territorio en disputa (Spadi, 2001: 120). Pero esta reivindicación devuelve la discusión a otro punto muerto: El Estado más cercano sería Haití, que, efectivamente, está a 64.6 km del cabo Dame Marie. Sin embargo, le siguen a poca distancia Jamaica y Cuba, isla, esta última, en la que se encuentra el enclave guantanamero controlado por EE. UU. Por otra parte, si se aduce el ejemplo hipotético de las islas canarias españolas o la persistencia Estados enclavados —Estado rodeado por completo de otro—, debe asumirse que la contigüidad no es un factor plenamente habilitante para reclamar la soberanía. Del mismo modo, el hecho de permitir, como se hace, a los pescaderos haitianos que

faenen en las aguas navacesas no acerca un ápice a la tesis de la contigüidad, dado que esto nada tiene que ver con la adquisición de derechos políticos, a no ser que se acepte la tesis de que la soberanía de un país es vulnerada por el simple hecho de que se permita pescar en sus aguas nacionales. Aunque para Pierre este un argumento definitivo, ello se debe a sus presupuestos decoloniales, que derivan en el economicismo (Pierre, 2014: 39-40). Los argumentos económicos se aceptan con reticencia en el Derecho internacional del mar, y aunque estimemos de acuerdo en que son un factor clave para no desestimar aspectos de *Realpolitik*, es necesario que se ponderen con una perspectiva más amplia, ligada a los recursos naturales del territorio (Pozo, 1997).

iii) Tierra de nadie. Haití, como antes España y Francia, adoleció de «falta de acción» sobre la isla de Navaza, por lo que, cuando se aprobó la ley guanera, la ínsula fue calificada como *terra nullius*. Quizás EE. UU. podría haber aceptado que Navaza era haitiana, como hizo con la isla de Alta Vela para la República Dominicana, si hubiera aparecido en la Constitución haitiana antes de la disputa de mediados del siglo XIX. Sin embargo, hasta la aprobación de la Constitución de 1874 (art. 2, refrendado en los textos fundamentales de 1879 (art. 1) y 1987 (art. 8 a)), una veintena de años después de la disputa, ninguna carta magna de Haití había declaró explícitamente que la isla pertenecía al país. En las ocho constituciones previas, esa pretensión se basó en una referencia genérica a las «islas adyacentes» (Pierre, 2014: 29 y 69-70; Spadi, 2001: 120).

En contraste, la doctrina que brinda soporte a las tesis haitianas señala que el descubrimiento español aplicaba a la isla La Española y a los territorios adyacentes, entre ellos Navaza, por lo que, cuando en el siglo XVII España cedió a Francia la isla mayor, también cedió la menor en virtud del principio *uti possidetis* o de herencia de la organización territorial previa a la independencia (Pierre, 2014: 28).

iv) Realpolitik. Como se ha explicado, la ley guanera habilitaba a EE. UU. a anexionar islas para la explotación. Esta autotutela partía de la premisa —ideal— de que toda isla anexionada por tal ley era *no man's land*; la misma práctica señala que las islas se abandonarían cuando se hubiera agotado la materia prima que justificaba esa labor empresarial. Sin embargo, EE. UU. sigue manteniendo su presencia en la isla, dando lugar a «[…] un gobierno que posee la isla sin querer su soberanía territorial [EE. UU.] y otro que no ejerce ninguna autoridad sobre ella, pero reclama su soberanía en su título original [Haití]» (Spadi, 2001: 123). Frente a esa *realpolitik*, podría alegarse que la devolución de Navaza es un acto de equidad que se ha dado en otros países con islas similares, entre ellos México, Canadá, Honduras, Kiribati, Tuvalu o Colombia (Pierre, 2014: 36). De hecho, para robustecer este argumento podría aducirse —así lo hace Sandner— la importancia de la Tercera Conferencia Internacional sobre Derecho Marítimo (1972-1982), que cambió «[...] el mapa político-administrativo, sobre todo en espacios con exagerada fragmentación territorial» (Sandner, 1988: 103).

v) *Teoría del rango de visión*. Spadi —que, en cualquier caso, reclama la firma de un pacto bilateral entre EE. UU. y Haití, y no acciones unilaterales— sostiene que Navaza, «[...] en condiciones claras, es visible desde el extremo suroccidental de Haití [...] El contacto visual es el marco dentro del cual el interés humano y las relaciones son general y probablemente estimuladas» (Spadi, 2001: 125). Sin embargo, este argumento falla en un par de puntos. Por una parte, sería imposible recurrir a ese ese argumento si más de un Estado tuviese en su rango de visión el territorio. ¿Se impondría el que lo ve más cerca en virtud del argumento de contigüidad que hemos analizado en el apartado *ii)*)? Por otra parte, la visión no es, en sí misma, un criterio, dirimente, pues es obvio que EE. UU. es la nación que dispone de los recursos más sofisticados para ver a larga distancia (drones, satélites). Con respecto a territorios tan disputados, islas que han sido objeto de reivindicaciones extemporáneas y que en su día fueron tierra de nadie, parajes remotos y deshabitados, ¿por qué debería prevalecer la visión física y no, por ejemplo, una clarividencia tecnológica o la visión a través de cuerpos opacos que posibilitan las nuevas tecnologías?

3.8. ISLA DEL MAÍZ Y PEQUEÑA ISLA DEL MAÍZ (*CORN ISLANDS*)

Durante 99 años, estas islas fueron concedidas a EE. UU. En perspectiva, son un ejemplo del modo en que entre Centroamérica y el Caribe la tierra pierde solidez, se fragmenta y cobran importancia las franjas, las ciudades libres, las islas como punta de lanza para el desarrollo de actividades en las que, aún hoy, la frontera entre lo legal y lo ilegal se desdibuja. Así puede considerarse la Costa de los Mosquitos (La Mosquitia) —alusión al pueblo indígena—, ahora anexada a Nicaragua, al igual que las islas del Maíz, actualmente adscritas a la región autónoma Costa Caribe Sur. El contrabando que florece en estos enclaves se percibe como una fuente de inestabilidad de la América española, dado que son punta de lanza de potencias extranjeras y de piratas amparados por estas. No obstante, también «[...] los asentamientos coloniales ingleses eran un enlace entre la parte hispanoamericana de Centroamérica y el Caribe, y esta última región significaba, a su vez, la comunicación con Europa y con los depósitos ingleses, franceses u holandeses en el Caribe» (Potthast-Jutkeit, 1998: 515). Parte de ello se ve en otra actividad ligada al tráfico ilícito: el narcotráfico. En cualquier caso, estas islas fueron parte del censo estadounidense en las décadas de los 50 y 60 del siglo XX; el acuerdo de cesión de 1914 culminó el 15 de abril de 1971, año en el que, como se ha dicho, se devolvió el control pleno de las islas a Nicaragua (U.S. Census Bureau, 1994: 3).

4. RECAPITULACIÓN FILOSÓFICO JURÍDICA: GEOGRAFÍA DEL MANTO INSULAR DESDE UNA PERSPECTIVA MARINA

i) La reflexión sobre el modo en que se organiza territorialmente la principal potencia del presente permite argumentar a partir de la existencia de un «Derecho imperial

estadounidense» no necesariamente entendido en el sentido economicista. La perspectiva de análisis se centra, más bien, en la forma a través de la cual el Derecho nacional de una potencia capaz de ejercer como imperio organiza territorios, ya sea continentales o insulares, y los encuadra en un ordenamiento jurídico interestatal —con los correspondientes intereses soberanos en alianza o conflicto—, así como en las distintas ramas del Derecho internacional público.

ii) Un análisis caso por caso sobre el estatus jurídico de las islas mencionadas —de un modo más panorámico, aquellas a las que aplican los «casos insulares» y, más ampliamente, las que son territorios no incorporados/no organizados— proporciona una visión panorámica de distintos temas que sobrepasan cualquier reflexión acotada al país objeto de este artículo, puesto que entroncan con cuestiones que cualquier Estado soberano podría plantearse, entre ellas los límites de la soberanía, la naturaleza del derecho a la propiedad privada, el estatus de los territorios conquistados y mantenidos en el tiempo o la relación entre federación, estados federados y territorios de otra índole.

iii) Los «casos insulares» constituyen un tema polémico en el que pugnan el principio de incorporación de pleno derecho a EE. UU. y el de soberanía de los territorios que podrían ser Estados conforme a lo establecido por la comunidad internacional. Tanto la generalidad de la Convención de Montevideo (1933) como muchos otros argumentos de peso hacen que territorios —por ejemplo, Puerto Rico— puedan considerarse, al menos en teoría, como Estados homologables al resto de Estados de la comunidad internacional. Este hecho evidencia que, en el fondo, el problema es doble. El primero, predominantemente político —aunque se ramifique en otros ámbitos—, está relacionado con lo que EE. UU. considera pertinente para mantener su rol hegemónico mundial. El segundo es jurídico (aunque no puede aislarse de otras consideraciones), en la medida en que pone sobre la mesa la cuestión relativa al aseguramiento de una transitoriedad que otorgue a los ciudadanos nativos el pleno control sobre recursos territoriales. Cuando se parte de una situación de rezago en las propiedades inmobiliarias o cuando entes como el ejército estadounidense desequilibran los cálculos de esta balanza, entonces la resolución de estos problemas se torna aún más enrevesada, y obliga a matizar continuamente para desdeñar generalizaciones incorrectas.

iv) Los territorios no organizados/no incorporados son, en ocasiones, hipótesis de laboratorio vivientes. Con esta paradoja pretende afirmarse que, por un lado, el análisis de esta temática se centra en temas que parecen alejados de los intereses del lector actual, dado que están relacionados con situaciones anteriores a los tratados de Westfalia (1648), entre ellas el significado de tomar (ocupar) un territorio, las circunstancias en las que este se considera *terra nullius* o la relación entre aquellos pioneros privados que, posteriormente, se enfrentan a la realidad de que la institucionalidad oficial recalifica un terreno visto en términos de patrimonio con base en lineamientos jurídicos de soberanía. Pero los asuntos añejos se unen a otros recurrentes como la libre navegación o comunicación y la explotación los recursos

naturales, e incluso con temas del siglo XX y XXI, entre ellos la transformación de excolonias en Estados que reivindican una igualdad en términos del Derecho internacional público o la naturaleza de principios como el de «autodeterminación».

v) Como hipótesis conclusiva, puede plantearse que el análisis precedente pone de manifiesto la existencia de una *geografía del manto insular*. Es decir, EE. UU. establece como modelo expansivo la potenciación de las islas como punta de lanza para penetrar en zonas hostiles o el uso de enclaves rodeados de tierra y no de agua para fortalecerlos. Obviamente, esta ejecutoria no es privativa de EE. UU., dado que el precedente talasocrático de Gran Bretaña es evidente (Schmitt, 2005); sin embargo, EE. UU. ha imprimido a este modo de actuar una idiosincrasia peculiar que es fruto de la utilización de la tecnología. Esta singularidad también es favorecida por su *perspectiva marina*, es decir, por el modo a través del cual que la tecnología de la principal potencia mundial *levanta* —si se me permite la metáfora— el lecho marino (en la línea de esta metaforización, también puede descorrer el cielo, explorar el espacio, u horadar la escala cuántica) transitando con libertad de una parte a otra del mundo para abastecer esos territorios o reactivar sus derechos políticos (sus capacidades, más bien) de acuerdo con factores ya coyunturales, ya estructurales.

vi) A la geografía del manto insular —como filosofía política— y la perspectiva marina —en cuanto militar— las acompaña un ordenamiento jurídico que posibilita esas acciones y las consolida. Es el discurso sobre las armas y las letras, de tan larga tradición. Algunos fragmentos de las formas jurídicas a través de las que se despliega en los territorios y los categoriza caracterizan el «Derecho imperial» de nuestra época.

NOTAS

1. «El arrecife de Kingman es una posesión estadounidense no organizada y no incorporada. [...]

Esta [los Fullard-Leo] es la misma familia propietaria del atolón de Palmira, y cuya propiedad fue confirmada por la decisión de la Suprema Corte, Estados Unidos v. Fullard-Leo, 331 U.S. 256 (1947); el caso no se refería al arrecife Kingman.

[...]

Las personas que deseen visitar el arrecife de Kingman deben obtener el permiso de la familia Fullard-Leo. Sin embargo, el Gobierno de los Estados Unidos no reconoce el derecho imputado de la familia a actuar de esta manera.

[...]

Los reclamos de propiedad privada hechos por la familia Fullard-Leo, aunque no están fundamentados, puede que necesiten ser resueltos» (The United States Court of Federal Claims, 2014: 33). La traducción del inglés es mía.

BIBLIOGRAFÍA

AGENCIA CENTRAL DE INTELIGENCIA (2024): «The World Factbook» [en línea] <https://www.cia.gov/the-world-factbook/>. [Consulta: 04/03/2024.]

BELENGUER BARRIONUEVO, Roque (2014): «El medio terrestre tabarquino: el gran olvidado», en G. Canales Martínez, J.M. Pérez Burgos y F. Lozano Quijada (coords.), *Nueva Tabarca, un desafío multidisciplinar, Alicante: Instituto Alicantino de Cultura Juan Gil-Albert, 109-124.*

BERMEJO, Romualdo y Eugenia LÓPEZ-JACOISTE (2019): «El régimen jurídico de la ocupación a la luz de algunos casos recientes, especialmente el de la franja de Gaza», *Revista UNISCI / UNISCI Journal*, 51, 255-297.

BUENO MARTÍNEZ, Gustavo (2000): *España frente a Europa*, Barcelona: Alba Editorial.

CAPÍTULO tercero (2017): «American Samoa and the citizenship clause: A study in *Insular Cases* revisionism», *Harvard Law Review*, 130(6), 1680-1703.

CAPÍTULO cuarto (2017): «Guam and the case for federal deference», *Harvard Law Review*, 130(6), 1704-1727.

CASE, Suzanne (2001): «Palmira Atoll», *Hastings Environmental Law Journal*, 7(3), 291-298.

CUISON VILLAZOR, Rose (2018): «Problematizing the protection of culture and the insular cases», *Harvard Law Rev*iew, 131(6), 127-152.

DEPARTAMENTO DE INTERIOR DE EEUU (s.f.a.): «Definitions of Insular Area Political Organizations» [en línea] <https://www.doi.gov/oia/islands/politicatypes>. [Consulta: 04/03/2024.]

— (s.f.b.): «Navassa Island» [en línea] <https://www.doi.gov/oia/islands/navassa>. [Consulta: 04/03/2024.]

— (s.f.c.): «Palmyra Atoll» [en línea] <https://www.doi.gov/oia/islands/palmyraatoll>. [Consulta: 04/03/2024.]

EDERINGTON, L. Benjamin (1997): «Property as a Natural Institution: The Separation of Property from Sovereignty in International Law», *American University International Law Review*, 13, (2), 263-331.

GOMES, Andrew (1998, 12 de abril): «Palmira-Isle of Mistery-For Sale», *Pacific Business News* [en línea] <https://www.bizjournals.com/pacific/stories/1998/04/13/story3.html>. [Consulta: 04/03/2024.]

GUARDANS CAMBÓ, Ignasi (2009): «Entregas extraordinarios, torturas y vuelos de la CIA», en A.R. Cuerda Riezu y F. Jiménez García (eds.), *Nuevos desafíos del Derecho Penal Internacional: terrorismo, crímenes internacionales y derechos fundamentales*, Madrid: Tecnos, 47-72.

KAPLAN, Amy (2005): «Where is Guantamano?», *American Quarterly*, 57(3), 831-858.

KUMAR, Krishan (2010): «Nation-states as empires, empires as nation-states: Two principles, one practice?», *Theory and Society*, 39(2), 119-143.

LANDÍN CARRASCO, Amancio (1992): «Los hallazgos españoles en el Pacífico», *Revista Española del Pacífico,* 2(2), 13-37.

LITTLEFIELD, Charles E. (1991): «The Insular Cases», *Harvard Law Review*, 15(3), 169-190.

MARIS, Gary L. (1967): «International Law and Guantanamo», *The Journal of Politics*, 29(2), 261-286.

MOSCOSO DE LA CUBA, Pablo (2013): «Las islas y sus efectos para la delimitación marítima en la jurisprudencia de la Corte Internacional de Justicia y del Tribunal Internacional del Derecho del Mar», *THEMIS. Revista de Derecho, 63, 255-274.*

PASTOR RIDRUEJO, José A. (2004): «¿Ha sido legal el uso de la fuerza en Afganistán?», en C. Ramón Chornet (ed.), *Los retos humanitarios del siglo XXI*, Valencia: Tirant lo Blanch, 95-109.

PIERRE, Michaelle (2014): «Haiti's claim over Navassa Island: a case study», tesis de maestría, World Maritime University.

PIGRAU SOLÉ, Antoni (2016): «El caso de la isla de Diego García: Territorio sin Derecho Internacional, personas sin derechos», *Revista Electrónica de Estudios Internacionales*, 31, 1-36.

POTTHAST-JUTKEIT, Barbara (1998): «Centroamérica y el contrabado por la Costa de los Mosquitos en el siglo XVIII», *Mesoamérica*, 36, 499-516.

POZO SERRANO, María Pilar. (1997): «¿Hacia una solución «económicamente» equitativa? Reflexiones sobre la evolución de la jurisprudencia internacional en materia de delimitación marina», *Anuario Español de Derecho Internacional*, 13, 407-466.
PURDUM, Todd S. (2000, 4 de mayo): «A Nature Group Contracts to Buy and Preserve One of the Last Pristine Pacific Atolls», *The New York Times* [en línea], <https://www.nytimes.com/2000/05/04/us/nature-group-contracts-buy-preserve-one-last-pristine-pacific-atolls.html>.[Consulta: 04/03/2024.]
RAUZON, Mark J. (1985): «Feral cats on Jarvis Island: their effects and their eradication», *Atoll Research Bulletin*, 282, 1-30.
RITTICH, Kerry (2018): «Occupied Iraq: Imperial Convergences?», *Leiden Journal of International Law*, 31(3), 479-508.
SANDNER, Gerhard (1988): «Límites e interrelaciones geopolíticas en el Caribe», *Análisis político*, 3, 103-112.
SCHMITT, Carl (2005): *El nomos de la tierra en el Derecho de Gentes del «Ius publicum europaeum»*, Buenos Aires: Struhart y Cía.
SPADI, Fabio (2001): «Navassa: Legal Nightmares in a Biological Heaven?», *IBRU Boundary and Security Bulletin,* Autumn 115-130.
STATHAM, Robert E. (2004): «The Commonwealth of the Northern Marianas Islands: Power vs the Rule of Constitutional Law», *The Journal of Pacific History*, 39(2), 255-257.
TAU ANZOÁTEGUI, Víctor (2021): *Casuismo y sistema: Indagación histórica sobre el espíritu del Derecho Indiano, Madrid: Universidad Carlos III de Madrid – Athenaica Ediciones Universitarias.*
THE UNITED STATES COURT OF FEDERAL CLAIMS (2014): «Kingman Reef Atoll Development L.L.C. v. United States» núm. 06-828L [en línea] <https://ecf.cofc.uscourts.gov/cgi-bin/show_public_doc?2006cv0828-139-0>.[Consulta: 04/03/2024.]
THE WHITE HOUSE PRESIDENT GEORGE W. BUSH (2009): «Establishment of the Pacific Remote Islands Marine National Monument. A Proclamation by the President of the United States of America» [en línea] <https://web.archive.org/web/20090114105644/https://www.fws.gov/pacific/news/2009/Monuments/pacificremoteislands.pdf >. [Consulta: 4/03/2024.]
TORRUELLA, Juan R. (2007): «The Insular Cases: The Establishment of a Regime of Political Apartheid», *University of Pennsylvania Journal of International Law*, 29(2), 283-347.
— (2018): «Why Puerto Rico does not need further experimentation with its future: A reply to the notion of "territorial federalism"», *Harvard Law Rev*iew, 131(3), 65-104.
U.S. CENSUS BUREAU (1994): «Puerto Rico and the Outlying Areas», en *Geographic Areas Reference Manual*, EE. UU.: US Department of Commerce. Economics and Statistics Administration, 1-42.
VARGAS Y MACHUCA, Bernardo de (1892): *Milicia y descripción de las Indias*, vol. II, Madrid: Librería de Victoriano Suárez.
VÉLEZ, Fabio (2013): «"Translatio imperii". Del mar de Dante al océano de Vázquez de Menchaca», *Revista de Estudios Histórico-Jurídicos*, 25, 783-797.
WEGMANN, Alex (2019): «175 Years of Science», *The Nature Conservancy* [en línea] <https://www.nature.org/en-us/about-us/where-we-work/united-states/hawaii/stories-in-hawaii/palmyra-175-years-of-science/>.[Consulta: 04/03/2024.]
WHEATLEY, Natasha Grace (2016): «Law, Time, and Sovereignty in Central Europe: Imperial Constitutions, Historical Rights, and the Afterlives of Empire», Tesis de doctorado, Universidad de Columbia.

Fecha de recepción: 19 de julio de 2024.
Fecha de aceptación: 20 de octubre de 2024.

NORMAS DE EDICIÓN DE LA REVISTA

http://teoriayderecho.tirant.com

I. CRITERIOS GENERALES PARA LA PRESENTACIÓN DE ARTÍCULOS A LA REVISTA *TEORÍA & DERECHO*

1. Los textos presentados para su publicación deberán ser trabajos originales e inéditos.
2. Los originales de los artículos de las dos primeras secciones («Debate» y «Estudios») tendrán una extensión de entre 8.000 y 14.000 palabras como máximo, un interlineado de 1.5 y letra de cuerpo 12. Deberán presentarse en soporte informático, preferentemente en Word (o formato revisable).
3. Los términos de presentación de originales finalizan el 1 de mayo (para el número que se publica en diciembre) y el 1 de noviembre (para el número que se publica en junio del año siguiente).
4. Al comienzo del texto han de indicarse el título del artículo, el nombre y apellidos del autor o autora, su cargo o profesión y la vinculación institucional con la que los autores o autoras desean aparecer identificados en el encabezamiento del artículo.
5. Los autores y autoras harán llegar con el artículo los datos suficientes para que la redacción de la Revista se pueda poner en contacto con ellos (dirección postal, teléfono y dirección electrónica).
6. Todos los artículos deberán ir acompañados de un resumen de 10 líneas o 1.000 caracteres en castellano y en inglés. También hay que incluir una lista de cinco o seis palabras clave o descriptores de la materia en castellano y en inglés. El título del artículo debe escribirse en español y en inglés.
7. Los artículos deben ir precedidos de un sumario de los diferentes apartados en que se estructuran. Los títulos y subtítulos de los textos seguirán el denominado «Sistema de numeración decimal de capítulos» y se escribirán solo con cifras arábigas. Las secciones se numerarán correlativamente a partir de 1; cada sección o epígrafe puede subdividirse en *n* partes (subtítulos o subepígrafes) desde 1 hasta *n*; cada subepígrafe, a su vez, puede dividirse en otras subsecciones o partes menores desde 1 hasta *n*, y así sucesivamente: (1., 1.1., 1.2., 1.2.1…) (4., 4.1, 4.1.1, 4.1.2…).
8. En el caso de los artículos publicados en la sección «Temas de Hoy», se podrán obviar algunos requisitos formales de la revista (resumen, informes externos, extensión, originalidad), pero nada impide que el original enviado los cumpla plenamente, en cuyo caso la Revista lo hará constar.
9. En las traducciones y textos de la sección titulada «Varia» no se exige el cumplimiento de los requisitos anteriores.

II. ADMISIÓN DE ORIGINALES Y SISTEMA DE EVALUACIÓN

Los originales publicados en las secciones «Debates» y «Estudios» serán sometidos al criterio de expertos. El sistema de arbitraje recurre a evaluadores externos y la Revista garantiza su anonimato. Cualquiera de los evaluadores puede hacer observaciones o sugerencias a los autores, siempre y cuando el trabajo haya sido aceptado. Los autores recibirán el resultado de la evaluación y, en su caso, se les concederá un período de tiempo suficiente para que puedan hacer las modificaciones propuestas por los evaluadores.

Los evaluadores tendrán en cuenta, especialmente: *i)* la idoneidad temática; *ii)* la calidad y el rigor de los argumentos que se presentan; *iii)* la adecuación de la estructura expositiva del texto; *iv)* la oportunidad y relevancia del artículo para la discusión de problemas en su área de investigación; y *v)* la validez de los datos y las fuentes bibliográficas.

Tras la evaluación realizada por los expertos, la publicación definitiva de los trabajos será sometida de nuevo a la consideración del Consejo Editorial de *Teoría & Derecho*, que se reserva la facultad de revisar y corregir los textos si estos presentan errores ortotipográficos, gramaticales o de estilo.

III. CORRESPONDENCIA CON LA REVISTA

Los originales deben ser enviados por correo electrónico a la dirección: teoria.derecho@uv.es. Los autores deben mandar dos archivos diferenciados en formato Word o compatible. El primero de ellos será el texto original que deberá estar completamente cegado y no contener ninguna referencia que permita al revisor/a conocer su autoría. En este mismo deberá aparecer el título, resumen y palabras clave en español e inglés. Asimismo, deberá suprimirse cualquier propiedad del archivo que haga referencia al autor o autora. El segundo documento debe contener título del texto, nombre y apellidos del autor, institución de origen, puesto que se ocupa en la actualidad, dirección de correo electrónico, así como los agradecimientos y referencias a proyectos de investigación en los que se enmarque la investigación del artículo.

Adjunto a los anteriores debe remitirse cumplimentado el formulario relativo a la gestión y protección de los datos personales por parte de la editorial de acuerdo con la normativa vigente. Dicho formulario puede encontrarse en la web de la revista: teoriayderecho.tirant.com.

También pueden remitirse por correo postal (copia impresa y en soporte digital) a la siguiente dirección: Revista *Teoría & Derecho*. Editorial Tirant lo Blanch, C/ Artes Gráficas 14, entresuelo, Valencia 46010.

IV. SISTEMA DE CITAS

1. Habida cuenta de la naturaleza de la revista, orientada fundamentalmente a la reflexión y al pensamiento teórico, el Consejo Editorial considera que la forma óptima de presentar las referencias es a través de una nota bibliográfica general situada al final del texto que refleje el estado de la cuestión, los textos más relevantes y la fundamentación de las tesis del autor.

2. Podrá utilizarse también el sistema fecha / autor (o «sistema Harvard») de citas, esto es, referencia en el texto y elenco bibliográfico al final del artículo.

Cuando se opte por este sistema, las citas se insertarán en el texto con un paréntesis que contenga: autor (coma), año de aparición de la obra y número(s) de página(s) precedido(s) de dos puntos. Ejemplo: (Vives, 2011:129).

Si se trata de una segunda (o posterior) edición de la obra, tal circunstancia se indicará con una voladita situada antes del año de publicación. Ejemplo: (Vives, ^{2}2011: 129).

– Al final del artículo se incluirá un elenco con las referencias completas de todas las obras mencionadas, según los criterios que se indican seguidamente.

En el caso de que se citen varias obras del mismo autor, se ordenarán cronológicamente. En la segunda y siguientes menciones, se sustituirán los apellidos y el nombre del autor por una doble raya (—) seguida de un espacio fijo y sin ninguna puntuación antes del paréntesis en el que se indica el año de publicación.

Si se citan varias obras del mismo autor y año, tanto en las referencias del texto principal como en el elenco bibliográfico se ordenarán alfabéticamente con letra minúscula, que se escribirá en cursiva (2011*a*).

Si la obra referenciada es de dos o más autores, únicamente se invertirá el orden del nombre y los apellidos del primer autor de la obra, pero no el de los autores sucesivos. Ejemplo: Pérez López, José, Marisa Fernández García y Javier Rodríguez Jiménez (2015): «Los delitos económicos», *Revista de Penología*, 23 (2), 45-64.

La fórmula anterior es preferible a la de consignar únicamente el nombre del primer autor seguido de las expresiones *et al.* (en cursiva) o «y otros», que también son admisibles.

Tanto en el texto principal como en la bibliografía se utilizarán las comillas angulares o latinas (« »), no las inglesas (“ ”). Si dentro de una oración o segmento de texto ya entrecomillado hay una expresión o un enunciado también entrecomillado, se utilizarán las comillas inglesas: (« “ ” »)

– En todo caso, las referencias completas del elenco bibliográfico deberán ser presentadas de la siguiente forma:

• Libros: apellido / s y nombre de pila completo (al menos el primero), año de publicación (entre paréntesis seguido de dos puntos), título del libro (en cursiva y con mayúscula sólo en la inicial de la primera palabra: la mayúscula en todas las palabras del título es una práctica anglosajona), lugar de edición (dos puntos), nombre del editor, (y, eventualmente, colección).

Ejemplo: Vives Antón, Tomás S. ([2]2011): *Fundamentos de derecho penal*, Valencia: Tirant lo Blanch.

• Capítulos de libro: apellido / s y nombre de pila completo (al menos el primero) del (o de los) autores, año de publicación (entre paréntesis seguido de dos puntos), título de la obra (entre comillas angulares), responsables subordinados (compiladores, editores, etc., precedidos de «en»; el nombre de pila de los responsables subordinados puede consignarse con la inicial antes del apellido), título del libro (ver arriba), lugar de edición (dos puntos), nombre del editor, (y, eventualmente, colección) y páginas (sin las grafías «pp.»).

Ejemplo: Pérez Pérez, María y Laura Martínez Martínez (2014): «Algunas acotaciones sobre los actuales modelos de teoría de la legislación», en P. Rodríguez Pérez y M. Ramos Ramos (comps.), *Nuevos modelos de teoría de la legislación*, Madrid: Teorema, 34-51.

• Artículos de revista científica: apellido / s y nombre de pila completo (al menos el primero) del (o de los) autores, año de publicación (entre paréntesis seguido de dos puntos), título del artículo (entre comillas angulares y con mayúscula solo en la primera palabra), nombre de la revista (en cursiva y mayúsculas), (eventualmente, volumen y número de la revista) y páginas (sin las grafías «pp.»).

Ejemplo: Cotterrell, Roger (2015): «The politics of jurisprudence revisited: a Swedish realist in historical context», *Ratio Juris*, 28 (1), 1-14.

• Recursos electrónicos: la inclusión de la referencia de los textos en formato electrónico en el elenco bibliográfico solo será obligatoria cuando el documento *únicamente* sea accesible en la red y no esté publicado en una revista o libro (incluidos los electrónicos). En este caso, la cita deberá presentarse del siguiente modo: apellido / s y nombre de pila completo (al menos el primero) del (o de los) autores, año de publicación (entre paréntesis seguido de dos puntos), título de la obra, indicación de soporte entre corchetes («[en línea]»), dirección completa y disponibilidad, que empezará siempre con el protocolo de transferencia de hipertexto (http) e irá entre antibambdas y con punto después de la antilambda de cierre (< >.), y fecha de consulta (entre corchetes y con punto final antes del corchete de cierre).

Ejemplo: Romero Carrascal, Susana (2008): «Archivos y delitos. La actuación de la Fiscalía de Patrimonio Histórico» [en línea], <http://www.arxivers.com/index.php/documents/formacio-1/jornades-d-estudi-i-debat-1/228-jed-080528-romero-1/file>. [Consulta: 12/06/2014.]

V. DERECHOS DE PROPIEDAD INTELECTUAL

Los autores y autoras ceden en exclusiva a la Revista los derechos de reproducción en revistas científicas de los textos admitidos. Asimismo, permiten al Consejo de la revista distribuir sus contenidos en todas aquellas bases de datos científicas en la que se encuentre indexada la Revista, siempre con el objetivo de asegurar una amplia distribución de los contenidos cedidos por los autores.

VI. ÍNDICES DE CALIDAD Y BASES DE DATOS

Cumple todos los criterios de calidad de revistas establecidos por el Comité 9 CNEAI (sexenios)
Journal Scholar Metrics. Revistas Españolas de Derecho
http://www.journal-scholar-metrics.infoec3.es/layout.php?id=rank&subject=law&start=100&order=h5_index&sort=DESC&related=T&country=es
ERIH PLUS (European Index for de Humanities and the Social Sciences) La revista Teoría y derecho ha sido incluida en este índice en febrero de 2017.
https://dbh.nsd.uib.no/publiseringskanaler/erihplus/periodical/info?id=488762
Google Scholar Metrics (2011-2015), Índice H de las revistas científicas españolas. La revista es la 81 de 148 posiciones.
CARHUS PLUS : valorada como A
RESH (Revistas españolas de ciencias sociales y humanidades).
Incluida en revistas jurídicas interdisciplinares. La mejor valorada por los expertos
http://epuc.cchs.csic.es/resh/indicadores
DICE (Difusión y calidad editorial de las revistas españolas de humanidades y ciencias sociales y jurídicas) http://dice.cindoc.csic.es/revista.php?rev=1888-3443
LATINDEX: cumple 33 criterios, sobre 33.
ISOC: Sumarios Derecho
CIRC: Calificada como B

I. GENERAL CRITERIA FOR SUBMISSIONS TO THE JOURNAL *TEORÍA & DERECHO*

1. All manuscripts submitted for consideration should be original and unpublished.
2. «Debate» and «Estudios» original texts should average between 8.000 and 14.000 words in length, spaced in 1.5, and 12p for the body type. They must be submitted electronically, preferably in Word (or other similar revisable format).
3. Submissions are required to be sent before May 1st. (for December's issue) and before November 1st. (for next year's June's issue).
4. At the beginning of the article, authors must indicate the title of the paper, full name of the author(s), position of the author(s), and the institution they come from.
5. Authors must send their contact information (specifically: address, telephone number and e-mail) within their submission form in order to allow the Journal to mantain regular contact with them.
6. Submissions must include an abstract of 10 lines or 1000 characters and a list of 5-6 key words in Spanish and English languages in any case. Titles must be written in Spanish and English.
7. Submissions should include a Summary, where titles of the different sections of the text are written. Submission structure must follow the so-called «Decimal Numbering System for Chapters and Subheading», and they should be written only with Arabic figures. Sections will be numbered consecutively starting on number 1. Each one can be divided in *n* sections, from 1 to *n*; each subsection can also be divided from 1 to *n*, and so on: (1., 1.1., 1.2., 1.2.1…) (4., 4.1., 4.1.1., 4.1.2…).
8. Submissions to the section «Temas de Hoy» may omit, fully or partially, the formal requeriments mentioned (abstracts, peer review, length and the requirement to be an unpublished work), althougth the manuscript submitted may accomplished some or all of them. In this case, the Journal *Teoría & Derecho* would mention it.
9. In the case of texts and translations submitted to the «Varia» section, there is no need to comply with the formal requirements mentioned above.

II. PEER REVIEW PROCESS AND REVIEW OF SUBMISSIONS

In order to select the texts to be published, all submissions will be reviewed through a peer review process. The assessment will be carried out by two reviewers. These reviewers, preferably external experts, will be anonymous and they will present their reconsiderations on admitted texts. In order to be published, submissions should obtain a positive assessment of both reviewers. In case of contradiction, the final decision will be taken by the Editorial Board. Authors will receive the reviewers' reconsiderations and they will be given an appropiate period of time to modify their texts according to the comments of the experts.

Reviewers will take into consideration the following: *i)* the thematic suitability; *ii)* the scientific quality and the competence of the arguments presented; *iii)* the appropriateness of the text's structure; *iv)* the opportunity and relevance of the submission in a given research area; and *v)* the acceptance of the data and bibliographical resources used.

Once the revision process undertaken by the experts has finished, the final text to be published will be approved by the Editorial Board of the Journal, which will be competent as well to make orthotypographics, gramatical and writting-style corrections, if necessary.

III. CORRESPONDENCE WITH THE JOURNAL

Manuscripts should be sent to this e-mail address: teoria.derecho@uv.es

Authors must send two different files in Word (or compatible format). First the manuscript which must be completely blind. This is, no reference may let the reviewer know the authorship. In this one, the title, the abstract and the keywords must be include in English and Spanish. Moreover, all the properties of the file referring the author/s must be also erased. The second file should contain: the title, name and surname/s of the author/s, institution, current position, email address, and the acknowledgments and references to the research projects or so on within which the research has been developed.

Attached to the former ones, the form related to the personal data protection and its treatment should be submitted to the editorial, according to the current regulations. This form can be downloaded at the website of the journal: teoriayderecho.tirant.com

Manuscripts can be also sent through postal mail to Revista *Teoría & Derecho*. Editorial Tirant lo Blanch, C/ Artes Gráficas 14, entresuelo, 46010 Valencia.

IV. QUOTES CITATION FORMAT

Only two citation systems can be used:

1. Taking into account the nature of the Journal, which is orientated to theoretical thought, it should be preferred to include the bibliographical references at the end of the article, through a bibliographic note which reflects the state of the question, the most relevant articles and the fundamentals of the author's thesis or reasoning.

2. The date / author citation system (or Harvard system) could also be used. In this case, a list of bibliography must be placed at the end of the document.

Doing this, citations should go in brackets, including author's surname, date of publication and the page / s. For instance, (Vives 2011:129).

If it is a second or later edition, this matter will be indicated with a superscript, written just before the year of publication. For instance, (Vives, ^{2}2011: 129).

– At the end of the article, a complete list of bibliography should be included according to the following criteria:

If several papers of the same author are cited, they should follow a chronological order. In second and subsequent mentions, authors' surnames and names will be substituted by a double line (—), followed by a space, and no punctuation before the brackets, with the year of publication.

If the articles or books are published by the same author in the same year, they will be ordered alphabetically with a lowercase letter written in italics (2001*a*).

If the paper referenced to is written by two or more authors, the order will be changed only in relation to the first author's name and surnames. For instance, Pérez López, José, Marisa Fernández García y Javier Rodríguez Jiménez (2015): «Los delitos económicos», *Revista de Penología*, 23 (2), 45-64.

The previous system will be preferible to the one in which the first author's name is followed by the expression *et al.* (italics) or «and others», although the latter is also allowed.

In the main text and in the bibliography, Latin / Spanish quotation marks (« ») will be used, not English quotation marks (“ ”). English quotation marks can be used whenever an expression and / or part of a sentence is already quoted, within the major sentence (« “ ” »).

– In any case, bibliographical references must be quoted in the following way:

Books: Author's or Authors' Surname / s, Name (at least the first), year of publication (with brackets followed by colon), Title of the book (italics and capital letter only the initial of the first word, not all the words), place of edition (colon), editor's name, and, eventually, collection.

For instance: Vives Antón, Tomás S. ([2]2011): *Fundamentos de derecho penal*, Valencia: Tirant lo Blanch.

Book chapter: Author's or Authors' Surname / s, Name (at least the first), year of publication (with brackets followed by colon), Title of the chapter (double quotation marks), person in charge of the edition (editor, coordinator, compilator…, preceded by «in»; name of person in charge may be written with the initial before the surnames), title of the book (italics and capital letter only the first letter of the first word, not with all the words), place of edition (colon), editor's name, eventually, collection, and pages (without any graphical symbol).

For instance: Pérez Pérez, María y Laura Martínez Martínez (2014): «Algunas acotaciones sobre los actuales modelos de teoría de la legislación», en P. Rodríguez Pérez y M. Ramos Ramos (comps.), *Nuevos modelos de teoría de la legislación*, Madrid: Teorema, 34-51.

Article of Scientific Journals: Author's or Authors' Surname/s, Name (at least the first), year of publication (with brackets followed by colon), Title of the article (double quotation marks), title of the journal (italics and capital letters), eventually volumen and issue of the journal, and pages (without any graphical symbol).

For instance: Cotterrell, Roger (2015): «The politics of jurisprudence revisited: a Swedish realist in historical context», *Ratio Juris*, 28 (1), 1-14.

Electronic Resources: References to an electronic resources in the list of bibliography is not compulsory when the resource is only available online and it is not published in a journal or book (even electronic ones). If so, the reference should be as follows: Author's or Authors' Surname/s, Name (at least the first), year of publication (with brackets followed by colon), Title of the resource (double quotation marks), format between square brackets («[on line]»), link and availabilty, beginning always with the Hypertext Transfer Protocol (http) and it will be between broken brackets, followed by a stop after the end (< >.), and date of reference (with square brackets and full stop before the end square bracket).

For instance, Romero Carrascal, Susana (2008): «Archivos y delitos. La actuación de la Fiscalía de Patrimonio Histórico» [en línea], <http://www.arxivers.com/index.php/documents/formacio-1/jornades-d-estudi-i-debat-1/228-jed-080528-romero-1/file>. [Date of reference: 12/06/2014.]

V. COPYRIGHT NOTICE

The authors of submitted texts agree to assign their copyright and reproduction rights to the Journal. Therefore, the Journal will have exclusive rights to authorize the reproduction public display and / or distribution of the work. The authors authorize the Journal to make the work available and to share its content in scientific databases, in which the Journal is indexed, in order to make sure a greater citation of those contents granted by the authors.

VI. SCIENTIFIC RANKINGS AND DATABASES

The Journal meets the qualitiy criteria established by the "Comité 9 CNEAI" (six-year terms)
Journal Scholar Metrics / Spanish Scientific Legal Journals
http://www.journal-scholar-metrics.infoec3.es/layout.php?id=rank&subject=law&start=100&order=h5_index&sort=DESC&related=T&country=es

ERIH PLUS (European Reference Index for the Humanities and Social Sciences). The Journal Teoría & Derecho has been indexed in February 2017.
https://dbh.nsd.uib.no/publiseringskanaler/erihplus/periodical/info?id=488762
Google Scholar Metrics (2011-2015), H Index of Spanish Scientific Journals. The Journal is ranked 81 out of 148.
CARHUS PLUS : Position A
RESH (Spanish Journals of Social Sciences and Humanities). The journal is the highest ranked within the interdisciplinary scientific journals.
http://epuc.cchs.csic.es/resh/indicadores
DICE (Dissemination and editorial quality of the Spanish Scientific Journal of Humanities and Social and Legal Sciences)
http://dice.cindoc.csic.es/revista.php?rev=1888-3443
LATINDEX: it meets 33 criteria out of 33.
ISOC: Abstracts in Law
CIRC: Position B

NORMAS ÉTICAS Y DECLARACIÓN DE BUENAS PRÁCTICAS

http://teoriayderecho.tirant.com

NORMAS ÉTICAS Y DECLARACIÓN DE BUENAS PRÁCTICAS

Teoría & Derecho. Revista de Pensamiento Jurídico es una revista académica que publica trabajos científicos y asume un firme compromiso con el cumplimiento de los estándares éticos y buenas prácticas. Tirant lo Blanch, editor de la revista, y su Comité editorial son responsables de determinar y mantener las siguientes normas en el proceso de selección y aceptación de contribuciones remitidas, así como de asumir las responsabilidades que deriven del proceso de publicación. Asimismo, se comprometen a velar por la ética y la calidad científica y académica de la revista, y a que todas las partes involucradas acepten y respeten los prinicipios que siguen.

OBLIGACIONES DE LOS AUTORES

Los autores de los artículos deberán respetar las siguientes obligaciones, así como cumplir las instrucciones para los autores que pueden encontrarse en la página web de la Revista y que se incluyen en cada número.

Los autores se comprometen a presentar manuscritos que no estén sujetos a consideración o que hayan sido enviados previamente a otras publicaciones, y a someterlos a un proceso anónimo de revisión. La investigación debe haber sido desarrollada de acuerdo con estándares éticos e incluirse en la publicación, de modo que se aporte la suficiente información para permitir la réplica académica. Todos los autores deben mencionar las fuentes de financiación que les ha permitido llevar a cabo la investigación y alcanzar los resultados pertinentes, así como la posible influencia de estos en los mismos.

Los artículos remitidos deben ser trabajos completamente originales y, en el caso de que los autores hayan utilizado el trabajo y/o fragmentos de otros, estos deben ser citados de forma adecuada, según las indicaciones anteriormente referidas. Cualquier forma de plagio constituye una práctica contraria a estas buenas prácticas y no se aceptará bajo ninguna circunstancia. Los autores no deben plagiar, práctica por la que se entenderá una copia literal o casi literal, una paráfrasis del texto o de los resultados de la investigación de otros autores.

No es adecuado que un autor remita más de un trabajo en los que se describa una misma investigación, a menos que se trate de una remisión de un artículo rechazado para su publicación en otra revista.

En caso de conocer un error fundamental o una imprecisión relevante en los trabajos ya publicados, los autores deben comunicarlos a la revista.

OBLIGACIONES ÉTICAS DE LOS EDITORES

Los editores son responsables del cumplimiento de las obligaciones anteriormente enunciadas, deben velar por asegurar la calidad científica y académica de Teoría & Derecho, y son los principales responsables de la legalidad en materia de propiedad intelectual, derechos de autor, infracción y plagio.

El editor debe respetar la independencia intelectual de los autores y considerar todos los originales remitidos para su publicación, valorando cada una de las contribuciones de forma objetiva. También es el responsable de garantizar a autores y revisores que los manuscritos serán sometidos a un proceso de revisión anónimo. La decisión última y la responsabilidad sobre la aceptación o rechazo de un original recaerá en el Comité editorial.

Es obligación de los editores buscar revisores, que serán escogidos por razón de su prestigio, conocimientos y buen juicio, y valorarán la calidad científica de los trabajos remitidos a la revista. No obstante, los trabajos pueden ser rechazados para su publicación sin revisión externa si, a juicio de los editores, el trabajo es inapropiado para ser publicado en la revista. Esta decisión deberá basarse en la inadecuación del trabajo al contenido y la línea de la revista, la ausencia de actualidad o de suficiente interés, la

incorrección formal, o cualquiera otra razón relacionada con el ámbito de la revista. Los artículos, por tanto, se valorarán únicamente con base en su contenido, con independencia de intereses comerciales y sin discriminación de tipo étnico, racial, ideológico, religioso, de género o cualquier otra hacia el autor.

Los editores deben comunicar el resultado de los informes de valoración a los autores sin revelar ningún tipo de información acerca de la identidad de los revisores, pero informarán de los criterios utilizados por estos a la hora de evaluar el trabajo de cara a su publicación.

Los miembros del Comité editorial y cualquier otro editor no deben desvelar ningún tipo de información sobre un original sometido a consideración a nadie más que aquellos a quienes se solicite asesoramiento académico o científico. La confidencialidad debe ser protegida en todo caso durante el proceso de revisión (de los revisores para el autor, y viceversa). Los editores se abstendrán de participar de la evaluación y cualquier otro proceso editorial que incluya manuscritos en caso de tener algún conflicto de interés a causa de una posible relación competitiva, de colaboración o de otro tipo con cualquiera de los autores del manuscrito presentado.

La información no publicada, los argumentos o las interpretaciones contenidas en un original remitido a la revista no podrán ser utilizados en investigaciones desarrollada por los editores.

OBLIGACIONES DE LOS REVISORES

Los revisores tienen el deber de enjuiciar y valorar de forma objetiva la calidad de los originales, así como su originalidad, haciendo especial hincapié en el mantenimiento de la calidad científica y los estándares científicos. Los revisores ayudan a los editores en la toma de decisiones y pueden asistir a los autores a la mejora de los trabajos.

El revisor advertirá al editor de cualquier similitud sustantiva entre el manuscrito bajo revisión y cualquier artículo o manuscrito similar enviado o publicado en cualquier otra revista o editorial.

Los revisores deben actuar de forma adecuada, remitiendo su informe de valoración en el plazo establecido, aun de modo aproximado. Cuando un revisor propuesto considere no que no es la persona idónea o apta para valorar de la investigación presentada o que sepa que no será posible cumplir los plazos para la revisión, deberá notificar al editor y excusarse tan pronto como sea posible.

Los revisores deben considerar un original sometido a revisión como un documento confidencial. Nunca deberá ser mostrado o discutido con terceros, salvo en casos excepcionales en los que podrá consultarse a personas que puedan asesorar científica o académicamente; en estos supuestos, las identidades de las personas consultadas deben ser reveladas al editor.

CORRECIONES Y RETRACTACIÓN

En caso de reconocer algún dato o información publicado en la revista *Teoría & Derecho* como falso, erróneo, engañoso o fraudulento, los editores deberán informar a los autores de estas circunstancias y esperar una respuesta de los mismos antes de tomar una decisión editorial. En caso de no quedar satisfechos con la respuesta o si no hay ninguna, los edtiores podrán decidir retractarse de lo publicado previo acuerdo del Consejo editorial.

La revista *Teoría & Derecho* atenderá cualquier reclamación y queja sobre los contenidos publicados a través de la dirección teoria.derecho@uv.es. Los editores se comprometen a hacer un seguimiento y proceder a la revisión y, en su caso, retractación si estos son necesarios.

Cuando proceda una retractacion o una corrección de lo publicado en la revista, los editores se guiarán por los principios contenidos en las Guidelines for Retracting Articles del Committee on Publication Ethics (COPE).

ETHICAL GUIDELINES FOR JOURNAL PUBLICATION

Teoría & Derecho, Revista de Pensamiento Jurídico is an academic journal publishing scientific articles with a firm commitment with the meeting of these ethical standards and good practices. Tirant lo Blanch, as editor of the journal, and its Editorial Board have the responsibility to establish and maintain the guidelines to select and accept papers submitted to this journal, assuming those responsabilities arising from the publishing process. Also, they fully commit to ensure that all parties involved meet the ethics, scientific and scholar standards.

ETHICAL OBLIGATIONS OF AUTHORS

The authors are expected to adhere to the following ethical guidelines and to respect the instructions for authors, posted to the website of the Journal and included within each issue.

The authors commit to submit orginals which are not submitted to another journals, and to submit them to a double blind review process. The research conducted should be developed according to ethical standards, including sufficient information in the article to allow an academic response. The authors should refer the funding resources to develope the research and achieve the results, as well as the possible influence of these resources in the results.

The authors should ensure that they have written entirely original works, and if the authors have used the work and/or words of other, that this has been appropriately cited. Plagiarism in all its forms constitutes unethical and publishing behavior and is not acceptable. Authors should not engage in plagiarism —verbatim or near— verbatim copying, or very close paraphrasing, of text or results from another's work.

It is improper for an author to submit manuscripts describing essentially the same research to more than one journal of primary publication, unless it is a resubmission of a manuscript rejected for or withdrawn from publication.

When a fundamental error or a relevant inaccuracy within already published articles is known, the authors should advertise it to the editors.

ETHICAL OBLIGATIONS OF EDITORS

The Editor should be responsible for the accomplishment of these and above obligations, and shall ensure the scientific and academic quality of the journal Teoría & Derecho, including the respect of legal issues concerning submissions to this Journal, specially intellectual property rights.

The editor should respect the intellectual independence of authors and should consider manuscripts submitted for publication, judging each on its own merits. They are also in charge to ensure that authors and reviewers get involve in a double blind review process. The final decision and the responsibility for acceptance or rejection of a manuscript rests with the Editorial Board.

The editor is required to seek advice from reviewers, who will be chosen for their expertise and good judgment, as to the quality and reliability of manuscripts submitted for publication. However, manuscripts may be rejected without external review if considered by the editors to be inappropriate for the journal. Such rejections may be based on the failure of the manuscript to fit the scope of the journal, to be of current or sufficiently broad interest, to provide adequate depth of content, to be written properly, or other reasons related to the journal. Therefore, contributions may be reviewed solely according to their contents, regardless their commercial interest and without discrimination on the basis of racial, ideological, religious, gender, or any other feature of the author.

The editor should inform about the result of the assessment reports to the authors without disclosing any detail about the reviewers' identity, but the editor should provide the most detailed information about the criteria used by the revieweres to assess the submission.

Members of the Editorial Board and members of the editor's staff should not disclose any information about a manuscript under consideration to anyone other than those from whom professional advice is sought. Confidentiality shall be protected in any case during the peer-review process. The editor should refrain from participating in the review process in case of any conflict of interest due to a competitive or collaboration relation with the author/s of the submission.

Unpublished information, arguments, or interpretations disclosed in a submitted manuscript should not be used in an editor´s own research.

ETHICAL OBLIGATIONS OF MANUSCRIPT REVIEWERS

Reviewers have an obligation to do a fair share of reviewing. A reviewer of a manuscript should judge objectively the quality of the manuscript and its originality, with due regard to the maintenance of high scientific and theoretical standards. A reviewer should call to the editor's attention any substantial similarity between the manuscript under consideration and any published paper or any manuscript submitted concurrently to another journal.

The reviewer should inform the editor about any substantive similarity of the submission under review with any other submission or published article in any other journal or publication.

Reviewers should act promptly, submitting a report in a timely manner. A chosen reviewer who feels inadequately qualified to judge the research reported in a manuscript or knows that its prompt review will be impossible should notify the editor and excuse himself from the review process as soon as possible.

A reviewer should treat a manuscript sent for review as a confidential document. It should neither be shown to nor discussed with others except, in special cases, to persons from whom specific advice may be sought; in that event, the identities of those consulted should be disclosed to the editor.

CORRECTIONS AND RETRACTION

In case of misconduct, inaccuracy or any fraudulent, false or misleading information published in the journal *Teoría & Derecho*, the editor should consult the author/s giving them the oppotunity to respond to any allegations before taking an editorial decision. In case of insufficient or lack of response, the editor may decide to adopt a formal retraction or withdrawal of a publication fromt the journal, in conjunction with informing the head of the author or reviewer's department, Abstracting & Indexing services and the readership of the publication, with the agreement of the Editorial Board.

The journal *Teoría & Derecho* will take account of any claim and complaint about the published contents through the email teoria.derecho@uv.es. The editor is fully committed to follow them up and proceed to the review and, when applicable, to the retraction.

When a retraction or corrections proceed, the editor will follow the Guidelines for Retracting Articles del Committee on Publication Ethics (COPE).