INTELIGENCIA ARTIFICIAL Y DERECHOS HUMANOS: UNA DISTOPÍA POSIBLE[1]

Miguel L. Lacruz Mantecón
Profesor Titular de Derecho civil
Universidad de Zaragoza

Sumario. I. Los sistemas inteligentes de reconocimiento y los derechos fundamentales. II. Control de los sistemas de IA y su incidencia en los derechos fundamentales. 2.1 El paso de la regulación ética a la normativa: La Ley de la IA. 2.2 La clasificación de la LIA de los sistemas de riesgo para los derechos de las personas. 2.2.1 Sistemas prohibidos (con excepciones). 2.2.2 Sistemas de alto riesgo. 2.2.3 Diferencia entre ambos tipos de sistemas. III. Las aplicaciones de la IA y el posible daño a los derechos fundamentales. 3.1 Los sistemas de datos biométricos e identificativos, en general. 3.2 Aplicaciones de los datos biométricos. 3.2.1 Sistemas de identificación y localización de personas. 3.2.2 El perfilado predictivo. 3.2.3 El reconocimiento de emociones y la detección del pensamiento. 3.3 Las técnicas subliminales y el *nudge*. 3.4 Los sistemas de calificación o crédito social. 3.5 La falsificación de la realidad. IV. Epílogo: Doomers contra boomers. Bibliografía por autores.

I. Los sistemas inteligentes de reconocimiento y los derechos fundamentales

El día 6 de diciembre de 2022, la activista Isabel Vaughan-Spruce fue arrestada en Birmingham frente a un centro de abortos. El motivo de su arresto

1 Trabajo escrito al amparo del Proyecto de investigación «Derecho e inteligencia artificial: nuevos horizontes jurídicos de la personalidad y la responsabilidad robóticas», IP. Margarita Castilla, (PID2019-108669RB-100 / AEI / 10.13039 / 501100011033).

fue que su actitud, en silencio y con la mirada baja, podía interpretarse, y de hecho así se hizo por la policía inglesa, como si estuviera rezando mentalmente, y rezar es algo que no se permite en la proximidad de las clínicas abortivas. Como relata la web *Infovaticana*[2], las imágenes de video de su arresto muestran a un oficial de policía que le pregunta si estaba rezando, a lo que ella responde: «Podría estar rezando mentalmente».

Pues bien, este perspicaz policía inglés cuenta en la actualidad con un auxiliar muy eficaz que podría incluso confirmar sus sospechas de que la detenida había incurrido en este rezo mental, y me refiero a los sistemas de Inteligencia artificial, o IA. En este trabajo veremos cómo gracias a la tecnología de la IA ya existen sistemas de control y seguimiento individualizado de las actividades privadas, en particular en China, con sus sistema de crédito social, mientras que otros sistemas inteligentes nos examinan al desembarcar en los aeropuertos, o predicen nuestras posibilidades de pagar un crédito, o de reincidir penalmente; incluso se están desarrollando sistemas que ya permiten leer los pensamientos humanos, traduciéndolos a imágenes o palabras.

Es un lugar común entre los científicos o en general entre los comunicadores el anunciar un importante cambio social como consecuencia de la nueva tecnología inteligente, hablándose de una tecnología disruptiva, y de una auténtica revolución similar a la industrial del siglo XIX. Como nos dice el profesor de Oxford, Luciano **Floridi**[3], estamos ante un nuevo capítulo de la historia de la humanidad: «La revolución digital sólo sucederá una vez, y esa vez es ahora. Hemos pasado una página en la historia humana y un nuevo capítulo ha comenzado». Pero quizá no se ha insistido en que estamos ante una tecnología tremendamente invasiva, muy «fisgona», que facilita y reproduce un vicio humano muy común, como es la curiosidad por las vidas ajenas, y otro defecto humano también bastante extendido, como es la pasión por la manipulación y control de los demás. A esto se une, como indican **Pollicino** y **Paolucci**[4], el hecho de que las nuevas ciudades son ciudades inteligentes, cuyo funcionamiento implica interconexión mediante «…la combinación del Internet de las Cosas (IoT), el *Big data*, la computación ubicua y la nube. Todos estos elementos son los fundamentos de la arquitectura sobre las que descansa la (ideal) ciudadanía inteligente, y son los encargados de hacerlo más abierta, optimizable y, sobre todo, controlable». Aquí hay que poner el acento en el control de la ciudadanía: la IA aporta increíbles novedades como

2 https://infovaticana.com/2023/02/06/retiran-los-cargos-contra-la-inglesa-que-rezaba-en-silencio-frente-a-un-abortorio/

3 Floridi, Luciano, *The Ethics of Artificial intelligence*, Oxford University Press, Oxford, 2023, p. xi.

4 Pollicino, Oreste y Paolucci, Federica, «Digital constitutionalism to the test of the smart identity», *Journal of E-Learning and Knowledge Society,* Vol. 18, No. 3 (2022), pp. 8-21. P. 9. DOI: https://doi.org/10.20368/1971-8829/113581

técnica para la vigilancia policial de las calles, para la identificación de personas, para la investigación criminal, y asimismo para la influencia ideológica en las personas y el control de la ciudadanía.

El resultado puede ser una sociedad en la que no exista privacidad, en la que todo sea visible. El profesor de la Universidad George Washington, Jonathan **Turley**[5] denomina *sociedades-pecera* a aquéllas en las que no existe ningún tipo de intimidad, y señala que este tipo de distopía o antiutopía ha sido una creación literaria y cinematográfica de éxito, así en títulos como *1984*, *Fahrenheit 451*, *Minority Report* o *Total Recall*. Pues bien, la IA puede hacer de estas pesadillas una realidad.

Como vamos a ver, la tecnología inteligente implica unos riesgos evidentes para los derechos humanos, y señala al respecto **Megías Quirós**[6] que la limitación de los riesgos derivados de la IA va a ser el objetivo que inspire las más recientes elaboraciones regulatorias en la materia, como son la *Recomendación sobre la Ética de la Inteligencia Artificial* de la UNESCO de noviembre 2021[7] (y el subsiguiente *Report of the International Bioethics Committee on the ethical issues of neurotechnology*, de 15 de diciembre de dicho año), y la europea *Ley de la Inteligencia Artificial*, de 21 de abril 2021[8]. En la doctrina, esta idea de la peligrosidad de la IA es compartida por bastantes autores, como Martin **Ebers**[9], que nos dice que los sistemas de IA pueden dañar de manera impredecible la vida, la salud y la propiedad de las personas, así como «… afectar a los valores fundamentales en los que se basan las sociedades occidentales, dando lugar a violaciones de los derechos fundamentales de las personas, incluidos los derechos a la dignidad humana y a la autodeterminación, a la privacidad y a la protección de los datos personales, a la libertad de expresión y de reunión, a la no discriminación o al derecho a la tutela judicial efectiva y a un juez imparcial, así como a la pro-

5 Turley, Jonathan, «Anonymity, Obscurity and Technology: Reconsidering Privacy in the Age of Biometrics», *Boston University Law Review*, Vol. 100:2179, 2020, En https://www.bu.edu/bulawreview/files/2021/01/TURLEY.pdf

6 Megías Quirós, José Justo, «Derechos humanos e Inteligencia artificial», *Revista DIKAIOSYNE*, N.° 37, número especial sobre DDHH, en coedición con el Observatorio de Derechos Humanos de la Universidad de Los Andes. Mérida - Venezuela. Enero, 2022, p. 140.

7 Conferencia General de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), reunida en París del 9 al 24 de noviembre de 2021, en su 41.ª reunión.

8 Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen Normas Armonizadas en Materia de Inteligencia Artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión, COM/2021/206 final.

9 Ebers, Martin, «El futuro marco jurídico europeo de la inteligencia artificial», *Revista General de Legislación y Jurisprudencia*, 2023, número 2, pág. 254.

tección de los consumidores». **Pollicino** junto a **de Gregorio**[10] advierten que la tecnología inteligente puede ser lesiva *per se* para los derechos de las personas, como consecuencia de su falta de transparencia. Es decir, tiene repercusiones negativas sobre derechos fundamentales de las personas, como el derecho a la libre determinación, libertad de expresión y privacidad. Pero además «la difusión de la toma de decisiones automatizada también desafía a los sistemas democráticos por su impacto en el discurso público y la imposibilidad de comprender las decisiones que se realizan mediante sistemas automatizados que afectan a los derechos y libertades individuales». Otros, como **Matefi** y **Darius**[11] afirman que, aunque la IA nos trae innegables beneficios en muchos ámbitos, también en el del Derecho, pese a ello hay que asumir que «…derechos fundamentales como el derecho a la intimidad, a la dignidad, a la libertad de expresión, a la libertad de movimiento y a la seguridad de las personas, y muchos otros derechos de la personalidad o fundamentales reconocidos internacionalmente, son infringidos por el uso inapropiado de la IA».

Estos peligros, por otra parte, derivan precisamente de la perfección del funcionamiento de los sistemas inteligentes, que como nos señala el mismo **Megías Quirós**[12], al tiempo que traen muchos beneficios, plantean evidentes riesgos para los derechos humanos, «en especial los derechos a la integridad física, la vida privada, la igualdad, la libertad de expresión y de reunión, la tutela judicial efectiva, etc., riesgos que se agudizan cuando afectan a grupos vulnerables» (menores, inmigrantes, personas con discapacidad, edad avanzada, etc.). En este sentido, recoge **De Asís Roig**[13] las palabras de Michelle Bachelet, Alta Comisionada de las Naciones Unidas para los Derechos Humanos, en su discurso «Derechos humanos en la era digital ¿Pueden marcar la diferencia?» de 17 de octubre de 2019, cuando dijo: «Es esencial que en esta era digital prestemos especial atención a los derechos humanos… La revolución digital plantea un considerable problema de derechos humanos a escala mundial. Sus beneficios indudables no anulan sus riesgos evidentes».

Es precisamente para minimizar los riesgos que afectan a estos derechos para lo que la Unión europea ha seguido dos modelos de regulación distintos, uno ético y otro jurídico, elaborando primero textos de fijación de directrices éticas, y luego de propuestas normativas, que culminan en la

10 Pollicino, Oreste y De Gregorio, Giovanni, «Constitutional Law in the Algorithmic Society», *Constitutional Challenges in the Algorithmic Society*, Cambridge University Press, 2021, pp. 3-24. DOI: https://doi.org/10.1017/9781108914857.002, p. 4.

11 Matefi, Roxana, y Darius, Cupu, «Artificial intelligence and its impact on personality rights», en *Tratado de Inteligencia artificial y Derecho en el nuevo milenio*, Olejnik, Santiago de Chile, 2022, págs. 75-76.

12 Megías Quirós, «Derechos humanos e Inteligencia artificial», cit., p. 143.

13 De Asís Roig, Rafael Francisco, «Ética, Tecnología y Derechos», en *Inteligencia Artificial y Filosofía del Derecho*, Ediciones Laborum, Murcia, 2022, p. 35.

ya citada Ley de la Inteligencia artificial[14]. Este texto, como señala **Megías Quirós**[15], concreta los derechos reconocidos en la Carta de Derechos Fundamentales de la UE que pueden verse afectados por el uso de sistemas inteligentes. Serían los siguientes: El derecho a la dignidad humana (art. 1) y a la vida privada y familiar y la protección de datos de carácter personal (arts. 7 y 8). El derecho a la no discriminación y la igualdad entre hombres y mujeres (arts. 21 y 23). El derecho a la libertad de expresión y reunión (arts. 11 y 12), el derecho a la tutela judicial efectiva y a un juez imparcial, a la presunción de inocencia y los derechos de la defensa (arts. 47 y 48). Asimismo, los derechos de determinados grupos, como el de las condiciones justas y equitativas del trabajador (art. 31), o el derecho a la protección del consumidor (art. 28), los derechos del niño (art. 24) y la integración de las personas discapacitadas (art. 26).

La preocupación por la irrupción de los sistemas inteligentes, y su posible efecto negativo para la protección de los derechos de las personas, se detecta en otros textos europeos, destacando **Coca Payeras**[16] la *Resolución del Parlamento Europeo, de 6 de octubre de 2021, sobre la inteligencia artificial en el Derecho penal y su utilización por las autoridades policiales y judiciales en asuntos penales*. Este texto impone la necesidad de garantizar el respeto a los derechos y libertades fundamentales consagrados en la Carta de derechos fundamentales, así como la necesidad de que la tecnología de IA se desarrolle de manera que sitúe a las personas en su centro. Además, alerta sobre la utilización de herramientas de IA por las autoridades judiciales para la toma de decisiones sobre prisión preventiva, o para dictar sentencias, calcular las probabilidades de reincidencia y determinar la libertad condicional o resolver litigios en línea.

Esta Resolución expresa también gran preocupación por el uso por parte de las fuerzas policiales y servicios de inteligencia de bases de datos de reconocimiento facial, como la base *Clearview AI*, una base de datos de más de 3000 millones de imágenes que se han recopilado de redes sociales y otros lugares de internet. Asimismo, se preocupa por el uso de la inteligencia artificial en el control de fronteras, como el proyecto europeo *iBorderCtrl*, un sistema inteligente de detección de mentiras que, como dice **Coca Payeras** *(Ibid.)*, «elabora perfiles de los viajeros a partir de una entrevista automatizada por ordenador realizada a través de la cámara web del viajero antes del viaje y un análisis de 38 microgestos basado en la inteligencia

14 *Propuesta de Reglamento del Parlamento Europeo y del Consejo por el Que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión*, de 21.4.2021, COM(2021) 206 final.

15 Megías Quirós, «Derechos humanos e Inteligencia artificial», cit., p. 158.

16 Coca Payeras, Miguel, «Las iniciativas de la Unión europea sobre inteligencia artificial: de la persona electrónica, al difícil equilibrio entre la necesidad de impulsarla y evitar sus riesgos», *Revista de Derecho Civil,* vol. X, núm. 2, especial (junio, 2023), p. 24. En http://nreg.es/ojs/index.php/RDC

artificial, probado en Hungría, Letonia y Grecia». Ante estos sistemas de reconocimiento o biométricos, la Resolución pide a la Comisión que, por medios legislativos y no legislativos, y si es necesario a través de procedimientos de infracción, prohíba el tratamiento de datos biométricos, incluidas las imágenes faciales, mediante vigilancia masiva en espacios públicos con fines coercitivos (epígrafe 31).

II. Control de los sistemas de IA y su incidencia en los derechos fundamentales

2.1. El paso de la regulación ética a la normativa: La Ley de la IA

Centrándonos en el ámbito europeo, la vía por la que se opta es más la del establecimiento de principios éticos y la defensa de los derechos de la persona, que la creación de nuevos neuroderechos. Resume **Megías Quirós**[17] la evolución de la normativa europea sobre IA y protección de los derechos humanos señalando que los primeros textos europeos sobre cuestiones éticas de la IA siguen un camino similar al de la UNESCO, fijando inicialmente algunos principios éticos objetivos en tema de IA, pero luego incidiendo en la vía jurídico-normativa. El primer texto del Parlamento Europeo con contenido ético significativo fue la Carta sobre robótica, en la Resolución de 16 de febrero de 2017[18]. Le sigue la Resolución de 12 de febrero de 2019[19], sobre política industrial global europea en materia de inteligencia artificial, en la que, «además de instar a la Comisión a revisar y adaptar la legislación europea a la nueva realidad desde una perspectiva ética, concretaba nuevos principios para complementar la legislación e insistía en la aprobación "de una carta ética de buenas prácticas para la IA"». Ésta viene a ser la carta *Directrices éticas para una IA fiable (Ethics Guidelines for Trustworthy Artificial Intelligence)*, del Grupo independiente de expertos de alto nivel sobre IA, de 8 de abril de 2019. Como señalan **García Inda, González Ordovás,** y **Muñoz Soro**[20],

17 Megías Quirós, José Justo, «Derechos humanos e Inteligencia artificial», *Revista DIKAIOSYNE*, N.° 37, número especial sobre DDHH, en coedición con el Observatorio de Derechos Humanos de la Universidad de Los Andes. Mérida - Venezuela. Enero, 2022, pp. 155 y ss.

18 *Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica* (2015/2103(INL)).

19 *Resolución del Parlamento Europeo, de 12 de febrero de 2019, sobre una política industrial global europea en materia de inteligencia artificial y robótica (2018/2088(INI))*

20 García Inda, Andrés, González Ordovás, María José y Muñoz Soro, José Félix, «IA y Filosofía del Derecho: derechos, normas y sesgos», en Proyecto IASAC Unizar, Ciencias sociales y jurídicas, http://unidigitaliasac.unizar.es/ficha/la-ai-vista-desde-la-filosofia-del-derecho

el documento basa la fiabilidad de la IA en el cumplimiento de tres requisitos: que sea lícita, ética y robusta. Y añade «cuatro grandes principios éticos que pueden servir de referencia para pensar la relación entre el Derecho y la IA: el principio de respeto de la autonomía humana, el de prevención del daño, el de equidad y el de explicabilidad. O, dicho con otras palabras, la IA debe ser respetuosa, benigna, equitativa y transparente (o explicable)».

Por nuestra parte, en España tenemos la *Carta de derechos digitales*, adoptada en julio de 2021 por el Gobierno, que se inscribe en el contexto de la Estrategia Española Nacional de Inteligencia Artificial de 2020, pero que carece de efectos normativos.

En la doctrina, señala Francisco **de Asís Roig**[21], se estima que la evitación del daño a los derechos humanos pasa por una defensa de los derechos fundamentales con base jurídica, eventualmente complementada con el reconocimiento de nuevos *neuroderechos* (con la precaución de no crear una inflación de los mismos que los devalúe), y asimismo con nuevas normas y textos internacionales. Resalta este autor que la Oficina del Alto Comisionado de Naciones Unidas para los Derechos Humanos, en su informe sobre *El derecho a la privacidad en la era digital* (2021), proclamaba que los Estados deben establecer mecanismos de supervisión y reparación relacionados con la privacidad. Y también la UNESCO, a través de su Comité Internacional de Bioética, emitió un informe el 15 de diciembre de 2021 sobre *Cuestiones Éticas de la Neurotecnología*, en el que se encuentran sugerencias como las siguientes: *a) Agregar protocolos a los tratados internacionales, como la Declaración Universal de los Derechos Humanos, para abordar los desafíos que plantean las neurotecnologías. b) Reforzar la Declaración Universal de los Derechos Humanos, considerando que la neurotecnología desafía los derechos humanos existentes y que se requerirán nuevas garantías en función de las posibilidades de vulneración. c) Elaborar una Nueva Declaración Universal de Derechos Humanos y Neurotecnología*. En esta línea, el Congreso chileno aprueba el 12 de abril 2021 una reforma constitucional reconociendo el derecho a la integridad neuronal.

Volviendo al ámbito europeo, a partir de 2020 la UE reconoce la insuficiencia del marco ético para una protección eficaz de los derechos humanos frente a la IA, —se trata de la *Resolución de 20 de octubre sobre aspectos éticos de la IA, robótica y tecnologías conexas*— y se opta por una defensa propiamente jurídica, proponiendo a la Comisión la aprobación de dos Reglamentos de aplicación en los Estados de la UE. Como nos dice **Megías Quirós**[22], «El primer Reglamento propone un marco regulador de la IA con la conversión de principios éticos en obligaciones jurídicas, porque "los principios éticos comunes solo son eficaces cuando están también asentados en

21 De Asís Roig, «Ética, Tecnología y Derechos», cit., p. 38.

22 Megías Quirós, «Derechos humanos e Inteligencia artificial», *Revista DIKAIOSYNE*, cit., pp. 155 y ss.

Derecho" y porque las orientaciones éticas son un buen punto de partida, pero no garantizan que los desarrolladores, desplegadores y usuarios actúen de manera justa ni aseguran la protección eficaz de las personas». El segundo Reglamento es la *Resolución 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial*, que propone el establecimiento de un régimen de responsabilidad civil, objetiva y subjetiva, para que pueda ser reclamado cualquier daño causado por la IA.

Sin embargo, el paso definitivo hacia una regulación de una IA que va más allá de la ética lo da en 2021 la *Propuesta de Reglamento del Parlamento europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial*[23], la llamada *Ley de Inteligencia Artificial* (LIA)*,* que establecería un marco jurídico mediante normas claras para garantizar una IA fiable, segura y respetuosa con los derechos fundamentales. Destaca en este sentido **Coca Payeras**[24]**,** que entre los objetivos de la Ley de la IA según su Exposición de motivos, están los de «garantizar que los sistemas de IA introducidos y usados en el mercado de la UE sean seguros y respeten la legislación vigente en materia de derechos fundamentales y valores de la Unión», y «mejorar la gobernanza y la aplicación efectiva de la legislación vigente en materia de derechos fundamentales y los requisitos de seguridad aplicables a los sistemas de IA». En definitiva, no bastan los principios éticos para la salvaguarda de los derechos fundamentales, y por eso la Exposición de motivos de la LIA dice que, en consecuencia, «… las normas relativas a la IA …deben estar centradas en las personas, a fin de que la población tenga la seguridad de que la tecnología se usa de un modo seguro y en consonancia con la ley, lo que también implica respetar los derechos fundamentales».

Los últimos textos europeos siguen también esta línea, en particular la *Declaración conjunta del Parlamento Europeo, el Consejo y la Comisión de 23 de enero de 2023 sobre los Derechos y Principios Digitales para la Década Digital*[25]. Como dice **Coca Payeras**[26], en su Preámbulo esta Declaración insiste en su primer capítulo en que las personas constituyen el núcleo de la transformación digital de la Unión Europea, y que esta tecnología debe servir y beneficiar a todos los europeos para que cumplan sus aspiraciones*, …en*

23 Bruselas, 21.4.2021, COM(2021) 206 final, 2021/0106(COD)

24 Coca Payeras, Miguel, «Las iniciativas de la Unión europea sobre inteligencia artificial: de la persona electrónica, al difícil equilibrio entre la necesidad de impulsarla y evitar sus riesgos», *Revista de Derecho Civil,* vol. X, núm. 2, especial (junio, 2023), p. 24. En http://nreg.es/ojs/index.php/RDC

25 *Diario Oficial de la Unión Europea*, 23.1.2023, C 23/1.

26 Coca Payeras, «Las iniciativas de la Unión europea sobre inteligencia artificial… », cit., p. 37.

total seguridad y respetando plenamente sus derechos fundamentales. Nos comprometemos a: … b) adoptar las medidas necesarias para que los valores de la UE y los derechos de los ciudadanos reconocidos por el Derecho de la Unión se respeten tanto en línea como fuera de línea; c) fomentar y garantizar una acción responsable y diligente por parte de todos los agentes digitales, públicos y privados, en el entorno digital; d) promover activamente esta visión de la transformación digital, también en nuestras relaciones internacionales. Como vemos, una reiteración continua de la necesidad de respeto de los derechos de las personas (y, por tanto, un reconocimiento del riesgo que plantean estas nuevas técnicas para estos derechos).

2.2. La clasificación de la LIA de los sistemas de riesgo para los derechos de las personas

2.2.1. Sistemas prohibidos (con excepciones)

En la citada Ley de la IA, Propuesta de Reglamento europeo de abril de 2021, se distingue inicialmente en su articulado, tras una serie de conceptos y definiciones, una serie de *Prácticas de Inteligencia Artificial Prohibidas*, según reza el Título II de esta Ley. Veamos cuáles son estas aplicaciones técnicas de la IA que son prohibidas por conllevar riesgos inaceptables para los derechos fundamentales, señalando así el artículo 5 LIA que quedan prohibidas las siguientes *prácticas* de inteligencia artificial:

– Las técnicas subliminales *que trasciendan la conciencia de una persona para alterar de manera sustancial su comportamiento* provocando perjuicios físicos o psicológicos a esa persona o a otra. Es decir, técnicas de persuasión subliminal. Asimismo, las *técnicas de persuasión* que aprovechen vulnerabilidades de grupos específicos de personas debido a su edad o discapacidad física o mental *para alterar de manera sustancial el comportamiento de una persona que pertenezca a dicho grupo* de un modo que provoque o sea probable que provoque perjuicios físicos o psicológicos a esa persona o a otra. En definitiva, técnicas de persuasión y alteración subliminal de comportamientos.

– La utilización de sistemas de IA por las autoridades públicas para evaluar o clasificar a personas físicas *atendiendo a su conducta social o a características personales o de su personalidad conocidas o predichas*, de forma que la clasificación social resultante dé lugar a situaciones de trato perjudicial o desfavorable en contextos sociales sin relación con el que generó los datos de clasificación, o que sea desproporcionado con respecto al comportamiento social clasificado. Estamos ante técnicas de perfilado de sujetos y calificación social de los mismos.

– El uso de *sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público* salvo que sea estrictamente necesario para la búsqueda selectiva de víctimas concretas de un delito, incluidos menores desaparecidos; la prevención de una amenaza específica, importante e inminente para la vida o la seguridad física de las personas físicas, como un atentado terrorista. En particular, la identificación de personas relacionadas con delitos especialmente importantes como pertenencia a organización delictiva, terrorismo, trata de seres humanos, explotación sexual de niños y pornografía infantil, tráfico de drogas o de armas[27]. Una técnica de identificación biométrica se traduce en la vigilancia espacial de los sujetos en tiempo real, y como veremos, se puede trasladar sus resultados, una vez identificado el sujeto, al ámbito de la calificación social del mismo.

Estos sistemas de identificación biométrica son especialmente agresivos para los derechos fundamentales, por lo que además de esta prohibición, se establece en la Ley de la IA una serie de cautelas para su utilización, como son el tener en cuenta sus consecuencias para los derechos y las libertades de las personas implicadas, la adopción de salvaguardias para su uso, establecimiento de limitaciones temporales, geográficas y personales y, además *autorización previa por parte de una autoridad judicial o una autoridad administrativa independiente del Estado miembro donde vaya a utilizarse dicho sistema*.

2.2.2. Sistemas de alto riesgo

No terminan aquí las amenazas de la IA a los derechos fundamentales, aparte de estos sistemas prohibidos o de riesgo inaceptable, la Ley de la IA también refiere una serie de sistemas que se admiten pero que son calificados como de *alto riesgo*. Se trata de sistemas de IA referidos en los Anexos II y III de esta Ley que se aplican a tareas como las relacionadas con la salud, funcionamiento del tráfico rodado y el suministro de agua, gas, calefacción y electricidad, o sistemas de selección para el acceso a instituciones educativas y acceso al empleo. Asimismo, se incluyen los sistemas para la calificación crediticia o solvencia, o los sistemas empleados por las autoridades públicas para la aplicación de la ley, o la administración de justicia y la gestión de la migración, el asilo y el control fronterizo.

Pero como puede comprobarse, aquí no se está haciendo referencia a una aplicación técnica de la IA en particular, sino a la utilización de los sistemas inteligentes en funciones de control de ámbitos sensibles (datos médicos, regulación del tráfico, control de suministro de agua y energía) en los que se

27 Se trata de los delitos recogidos, como nos dice el art. 5.1.iii de esta Ley de la IA, en la Decisión Marco 2002/584/JAI del Consejo, *para el que la normativa en vigor en el Estado miembro implicado imponga una pena o una medida de seguridad privativas de libertad cuya duración máxima sea al menos de tres años*…

precisa una vigilancia especial para evitar daños a las personas. Es decir, que el concepto de *alto riesgo* va aquí referido a determinados ámbitos de utilización de la IA, mientras que el concepto de *prohibición* o riesgo inaceptable se aplica directamente a técnicas concretas de IA que actúan inmediatamente sobre las personas.

2.2.3. Diferencia entre ambos tipos de sistemas

A partir de lo anterior, veo que es posible diferenciar, por un lado, sistemas inteligentes que son, por el riesgo que producen para los derechos fundamentales, lesivos *per se* para estos derechos, y cuya utilización o está prohibida o sólo es posible con carácter excepcional y con fuertes medidas de control; y por otro lado, están los ámbitos en los que el mal funcionamiento de sistemas inteligentes conlleva un riesgo de afectación a los derechos fundamentales, ya por gestionar infraestructuras o instalaciones básicas para la vida en sociedad (suministros de agua, electricidad, líneas de datos), ya por tratarse de actividades que implican intervención pública o que tienen carácter especialmente delicado (administración de justicia, educación, empleo, instituciones financieras), en las que puede ser fácil que se produzcan lesiones a los derechos de las personas por un funcionamiento sesgado o por una avería del sistema.

Esto es lo que diferencia la lesión en uno y otro caso: en el supuesto de los sistemas inteligentes de riesgo inaceptable o prohibidos, los daños se producen por el buen funcionamiento del sistema, que consigue resultados que van más allá de lo que se podría esperar mediante la intervención exclusivamente humana; en cambio, en los sistemas inteligentes que actúan en ámbitos sensibles o de alto riesgo, los daños derivan del mal funcionamiento del sistema, que padece errores o disfunciones (sesgos, averías, bucles o caídas del sistema) que producen lesión a los derechos de las personas.

Veamos a continuación de qué riesgos estamos hablando, y cuáles son estas técnicas de IA peligrosas para los derechos fundamentales.

III. Las aplicaciones de la IA y el posible daño a los derechos fundamentales

3.1. Los sistemas de datos biométricos e identificativos, en general

A continuación, veremos los distintos sistemas inteligentes que hoy se aplican a tareas de vigilancia en calles y comercios, tareas de identificación y control social de la ciudadanía, comenzando por un breve examen de los datos biométricos que sirven de base para su funcionamiento.

Los sistemas de reconocimiento de personas por sus datos biométricos son invasivos por sí mismos, e impactan directamente en la protección de datos personales prescrita en el artículo 16 del Tratado de Funcionamiento de la Unión europea. Se trata de la utilización de sistemas inteligentes que, a través de videocámaras, reconocen los rasgos faciales de las personas y a partir de éstos, la identidad de cada persona en particular. A esto se puede añadir que, mediante diversos periféricos (medidores de tensión arterial, cámaras termográficas, scanner de huellas, micrófonos), se obtienen otros datos corporales como huellas dactilares, temperatura corporal, frecuencia cardiaca, sudoración o respiración del sujeto, tono de voz, movimiento de las pupilas, datos que permiten ampliar la exploración a informes sobre el estado de salud, emociones o incluso pensamientos de la persona, como luego veremos.

La Ley de la IA de 2021 procede a definir los «Datos biométricos» en el art. 3, n.º 33) como *…los datos personales obtenidos a partir de un tratamiento técnico específico, relativos a las características físicas, fisiológicas o conductuales de una persona física que permitan o confirmen la identificación única de dicha persona, como imágenes faciales o datos dactiloscópicos*[28]. Distingue **Cotino Hueso**[29] entre identificadores «fuertes», más utilizados por las tecnologías identificadoras de primera generación (huellas dactilares, ADN, estructura del iris, rostros, voz) e identificadores «débiles» que cada vez cobran más protagonismo (formas de andar, patrones de vasos sanguíneos, patrones de pulsación de teclas etc.): «Con la nueva generación de tecnologías se va más allá de la finalidad de identificación y se habla de "biometría del comportamiento" para el perfilado, reconocimiento de emociones o categorización de personas. Es por ello que, frente a los datos biométricos ligados únicamente a la identificación, se propone el concepto más amplio inclusivo de "datos basados en la biometría"». Por ejemplo, el dato de la pulsación de teclas puede llevarnos al diagnóstico del Parkinson en el sujeto, como investiga en el MIT un equipo del que forman parte los españoles Álvaro Sánchez-Ferro y Carlos Sánchez Mendoza[30].

A continuación, pasa el art. 3 LIA a definir los distintos sistemas de datos biométricos, definiciones que nos proporcionan una muestra de las diferentes aplicaciones de esta técnica: *34) Sistema de reconocimiento de emocio-*

28 Repite la definición del art. 3 de la Directiva (UE) 2016/680 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales: *13) «datos biométricos»: datos personales obtenidos a partir de un tratamiento técnico específico, relativos a las características físicas, fisiológicas o de conducta de una persona física que permitan o confirmen la identificación única de dicha persona, como imágenes faciales o datos dactiloscópicos*.

29 Cotino Hueso, «Reconocimiento facial automatizado y Sistemas de identificación biométrica bajo la regulación superpuesta de Inteligencia artificial y protección de datos», *Derecho Público de la inteligencia artificial,* Coordinadores Balaguer Callejón, Francisco, Cotino Hueso, Lorenzo, Fundación Manuel Giménez Abad, Zaragoza, 2023, p. 348.

30 neuroQWERTY. Massachusetts Institute of Technology, https://neuroqwerty.mit.edu/

nes: un sistema de IA destinado a detectar o deducir las emociones o las intenciones de personas físicas a partir de sus datos biométricos. 35) Sistema de categorización biométrica: un sistema de IA destinado a asignar a personas físicas a categorías concretas, como un sexo, edad, color de pelo, color de ojos, tatuajes, origen étnico u orientación sexual o política, en función de sus datos biométricos. 36) Sistema de identificación biométrica remota: un sistema de IA destinado a identificar a personas físicas a distancia comparando sus datos biométricos con los que figuran en una base de datos de referencia, y sin que el usuario del sistema de IA sepa de antemano si la persona en cuestión se encontrará en dicha base de datos y podrá ser identificada. A su vez, estos sistemas de identificación remota pueden funcionar *en tiempo real*, es decir procediendo a una identificación instantánea, o *en diferido*.

Apunta **Ebers**[31] a que el art. 5.1.d) de la LIA prohíbe los sistemas de IA de identificación biométrica remota (*biometric identification system* o BIS) «en tiempo real» en espacios de acceso público (por ejemplo, los sistemas de reconocimiento facial utilizados para identificar a una persona en la calle), pero que este enfoque de prohibir únicamente los sistemas de identificación biométrica utilizados para la aplicación de la ley es demasiado restrictivo, y sus excepciones demasiado amplias. Critica también que el reconocimiento automatizado de características como el sexo, la sexualidad o el origen étnico, lo que se conoce como *biometric categorization systems* o BCS, así como el reconocimiento automatizado de las emociones (*emotion recognition system* o ERS)[32], no están prohibidos por la LIA.

Por su parte, el Gobierno estadounidense elabora en octubre de 2022 un texto de directrices denominado *AI Bill Of Rights - Making Automated Systems Work For The American People*[33], en el que se tiene en cuenta la identificación biométrica. Y autores como Margaret Hu[34] advierten que estos sistemas biométricos, que ya están siendo usados para fines de vigilancia y control de fronteras, seguridad e inmigración, son de «alto riesgo», y pueden colisionar con derechos constitucionales fundamentales y derechos humanos. En definitiva, se trata de un desafía normativo-constitucional de primer orden.

---

31 Ebers, «El futuro marco jurídico europeo de la inteligencia artificial», cit., p. 267.

32 Un sistema inteligente de reconocimiento de emociones trata de detectar diferentes emociones del sujeto mediante la información procedente de las expresiones faciales, el movimiento corporal, los gestos y el lenguaje.

33 *AI Bill Of Rights - Making Automated Systems Work For The American People,* October 2022. En https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf

34 Hu, Margaret, «Biometrics and an AI Bill of Rights» ((2022), *William & Mary Law School Scholarship Repository*, Faculty Publications, Summer 2022. En https://scholarship.law.wm.edu/facpubs/2078

El análisis de estos datos biométricos permite conseguir toda una serie de resultados que vamos a ver a continuación, y su incidencia sobre los derechos fundamentales de las personas puede ser radical. Es cierto que tiene esta técnica una serie de utilidades prácticas, como la de utilizar la «huella facial» para desbloquear un teléfono móvil, o acceder a cuentas bancarias, pero su utilización principal es policial, como veremos. Da cuenta **Lucasiewicz**[35] de una curiosa utilización del reconocimiento facial, la de encontrar en los bancos de gametos a los donantes más semejantes a los receptores del material genético, para lograr que los hijos se les parezcan.

Pero pese a estos usos poco conflictivos, y como señala **Cotino Hueso**[36], la tecnología biométrica tiene el potencial de impactar prácticamente en todos los derechos fundamentales de las personas: «Entre otros, la Agencia de la Unión Europea para los Derechos Fundamentales (FRA) menciona los siguientes derechos: «la dignidad humana, al respeto de la vida privada, la protección datos personales, la no discriminación, los derechos del niño y de los mayores, los derechos de las personas con discapacidad, la libertad de reunión y asociación, la libertad de expresión, el derecho a una buena administración, y el derecho a un recurso efectivo ante la ley y a un juicio justo». Y esto en las distintas aplicaciones de la recolección de datos biométricos, que pasamos a examinar.

3.2. Aplicaciones de los datos biométricos

3.2.1. Sistemas de identificación y localización de personas

Los sistemas biométricos llevan tiempo siendo utilizados para asuntos como desbloquear el teléfono móvil, pero su aplicación estrella consiste en la vigilancia policial, identificación de personas y seguimiento de las mismas. La incidencia de estas técnicas sobre los derechos fundamentales es brutal. Como señala **Cotino Hueso**[37], nada impide hoy a la policía detectar a un ciudadano en la vía pública o en una manifestación, identificarlo y analizar si está en alguna base de datos específica, georreferenciarlo y reconstruir sus

35 Lucasiewicz, Rafal, «Facial recognition. Matching in gamete donation using AI», Tratado de Inteligencia artificial y Derecho en el nuevo milenio, Ediciones Olejnik, Santiago de Chile, 2022, pp. 391 y ss.

36 Cotino Hueso, Lorenzo, «Sistemas de inteligencia artificial con reconocimiento facial y datos biométricos. Mejor regular bien que prohibir mal», *El Cronista del Estado Social y Democrático de Derecho*, N.º. 100 (Septiembre-Octubre), 2022 (Ejemplar dedicado a: Inteligencia artificial y derecho), p. 71.

37 Cotino Hueso, «Sistemas de inteligencia artificial con reconocimiento facial y datos biométricos…», cit., p. 72.

recorridos, interacciones con otras personas y evaluar su comportamiento. Cabe añadir que, si además procesamos los datos de su teléfono móvil, la información puede completarse con todos los perfiles que quepa extraer de dicho dispositivo.

Pero estas tecnologías suponen un ataque a la intimidad personal, y ya hay sentencias que afirman la lesividad de las mismas, como la del Tribunal de Apelaciones del Noveno Circuito de los Estados Unidos, en *Patel v. Facebook, Inc.*[38], que condenó a *Facebook* en 2019 por la recopilación no consensuada de archivos faciales de los usuarios. El tribunal concluyó que la creación de una base de datos de rostros e identidades constituía una conducta contraria a la *Illinois Biometric Information Privacy Act (BIPA)*.

Dada su potencial peligrosidad, el tratamiento jurídico que reciben estos sistemas en la Ley de la IA europea no es suficientemente protector de los derechos fundamentales. Como señala **Cotino Hueso**[39], las prohibiciones y restricciones de los arts. 5 y 6 de la indicada Ley no impiden realmente la vulneración de derechos, pues los «sistemas de identificación biométrica» que en principio estarán prohibidos son los que funcionan «en tiempo real», en espacios de acceso público y con fines policiales. Luego a contrario, «…no estarían prohibidos los reconocimientos faciales que no funcionen a partir de imágenes en tiempo real, algo que ha sido especialmente criticado… Tampoco estarían prohibidos en los lugares que no sean de acceso al público, como locales de empresas y fábricas, oficinas y lugares de trabajo, las prisiones, zonas de control fronterizo y espacios en línea (Considerando 9 AIA[40]). Cabe recordar que, por exclusión, también quedan fuera de la prohibición las finalidades de defensa e inteligencia. Asimismo, los usos de estos sistemas en el ámbito de "migración o de asilo" parecen admitirse (art. 5.4 AIA, versión Presidencia checa 2022)». En el contexto privado, no hay que excluir las finalidades de seguridad privada en establecimientos de acceso público (supermercados, transporte, estadios, escuelas, y no estarían prohibidas las finalidades privadas de marketing, comercio u otras.

Desde el punto de vista de la normativa de protección de datos, el tratamiento de los datos obtenidos mediante técnicas biométricas está sometido al régimen general de protección de datos de la ley española, Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPD), en su art. 9, al tratarse de una *categoría especial* de datos. Aparecen específicamente mencionados los datos biométricos como categoría especial también en el art. 9 del Reglamento (UE) 2016/679

38 Patel v. Facebook, Inc., No. 18-15982 (9th Cir. 2019)

39 Cotino Hueso, «Sistemas de inteligencia artificial con reconocimiento facial y datos biométricos…», cit., pp. 69 y 70.

40 AIA es el equivalente en inglés a la LIA: *Artificial Intelligence Act*.

del Parlamento Europeo y del Consejo, de 27 de abril de 2016[41] (RGPD), y se les suman las garantías de los tratamientos automatizados del art. 22 de dicho Reglamento. Sin embargo, el mismo art. 9 RGPD permite excepcionalmente el tratamiento de estos datos especiales cuando *g) el tratamiento es necesario por razones de un interés público esencial, sobre la base del Derecho de la Unión o de los Estados miembros, que debe ser proporcional al objetivo perseguido, respetar en lo esencial el derecho a la protección de datos y establecer medidas adecuadas y específicas para proteger los intereses y derechos fundamentales del interesado*.

Apunta **Cotino Hueso**[42] que el uso de sistemas biométricos en el ámbito policial y penal quedará bajo la regulación especial de la Directiva 2016/680 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales en el ámbito penal (testigos, víctimas o encausados)[43]. En esta Directiva, el tratamiento de los datos biométricos como datos personales aparece en el art. 10: *El tratamiento de datos personales que revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, así como el tratamiento de datos genéticos, datos biométricos dirigidos a identificar de manera unívoca a una persona física… solo se permitirá cuando sea estrictamente necesario, con sujeción a las salvaguardias adecuadas para los derechos y libertades del interesado y únicamente cuando: a) lo autorice el Derecho de la Unión o del Estado miembro…* Y la necesidad estricta, y consiguiente autorización, nos la señala el art. 5 de la LIA, que permite el uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público para fines de aplicación de la ley con los objetivos siguientes: *i) la búsqueda selectiva de posibles víctimas concretas de un delito, incluidos menores desaparecidos; ii) la prevención de una amenaza específica, importante e inminente para la vida o la seguridad física de las personas físicas o de un atentado terrorista; iii) la detección, la localiza-*

41 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). «DOUE» núm. 119, de 4 de mayo de 2016. Artículo 9. Tratamiento de categorías especiales de datos personales: *1. Quedan prohibidos el tratamiento de datos personales que revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, y el tratamiento de datos genéticos, datos biométricos dirigidos a identificar de manera unívoca a una persona física, datos relativos a la salud o datos relativos a la vida sexual o la orientación sexual de una persona física…*

42 Cotino Hueso, «Sistemas de inteligencia artificial con reconocimiento facial y datos biométricos…», cit., p. 73.

43 De la misma fecha es el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de éstos, que sustituye al anterior Reglamento General de Protección de Datos de 1995.

ción, la identificación o el enjuiciamiento de la persona que ha cometido o se sospecha que ha cometido alguno de los delitos mencionados en el artículo 2, apartado 2, de la Decisión Marco 2002/584/JAI del Consejo 62, precepto éste que hace referencia a una larga serie de delitos[44].

Señala **Turley**[45] que, a diferencia de inventos anteriores, esta nueva técnica no es el resultado de un simple avance tecnológico como en las escuchas telefónicas o los dispositivos de escucha no intrusivos como los micrófonos direccionales. Además, está diseñada en gran parte para utilizar imágenes o datos que se coloquen por el propio sujeto a disposición de cualquiera colgándolas en Internet. Los datos de los que se nutre son variadísimos, como ya se ha visto: desde la imagen facial o del iris a sistemas que identifican a las personas por su manera de caminar, o por reconocimiento de su voz, sistemas de reconocimiento de pulsaciones de teclas, sistemas de reconocimiento de venas en las manos y muchos más. Como dato curioso, el ejército estadounidense utiliza una tecnología llamada vibrometría láser, que puede detectar con láseres infrarrojos la *firma cardíaca* única de personas con una precisión del 95 % y un alcance de 200 metros.

La utilización policial de estos sistemas ha sido discutida desde su inicio, como exponen **Pollicino** y **Paolucci**[46]: Frente a la utilización por la policía inglesa de sistemas de reconocimiento de asistentes a eventos públicos, entre mayo de 2017 y abril de 2019, un activista por los derechos humanos, el Sr. Bridges, reclamó ante los tribunales exigiendo el respeto a su derecho a la privacidad. Rechazada su demanda en primera instancia, el tribunal de apelación la admitió, declarando ilegal el uso de la tecnología biométrica por lesión a la privacidad y a la protección de datos de los particulares, con efectos negativos sobre la libertad de expresión y asociación[47]. Por su par-

---

44 Como son: -pertenencia a organización delictiva, - terrorismo, - trata de seres humanos, - explotación sexual de los niños y pornografía infantil, - tráfico ilícito de estupefacientes y sustancias psicotrópicas, - tráfico ilícito de armas, municiones y explosivos, - corrupción, - fraude … - blanqueo del producto del delito, - falsificación de moneda…, - delitos de alta tecnología, en particular delito informático, - delitos contra el medio ambiente, incluido el tráfico ilícito de especies…, - ayuda a la entrada y residencia en situación ilegal, - homicidio voluntario, agresión con lesiones graves, - tráfico ilícito de órganos y tejidos humanos, - secuestro, detención ilegal y toma de rehenes, - racismo y xenofobia, - robos organizados o a mano armada, - tráfico ilícito de bienes culturales, incluidas las antigüedades y las obras de arte, - estafa, - chantaje y extorsión de fondos, - violación de derechos de propiedad industrial y falsificación de mercancías, - falsificación de documentos administrativos y tráfico de documentos falsos, - falsificación de medios de pago, - tráfico ilícito de sustancias hormonales…, - tráfico ilícito de materiales radiactivos…, - tráfico de vehículos robados, - violación, - incendio voluntario, - delitos incluidos en la jurisdicción de la Corte Penal Internacional, - secuestro de aeronaves y buques, - sabotaje.

45 Turley, «Anonymity, Obscurity, And Technology: Reconsidering Privacy …», cit., pp. 2206 y 2207.

46 Pollicino y Paolucci, «Digital constitutionalism to the test of the smart identity», cit. p. 9.

47 Bridges v. South Wales Police, Case No: C1/2019/2670

te, relata Cotino Hueso[48] que la utilización de sistemas de reconocimiento para fines policiales de vigilancia o seguridad pública es ya un hecho incluso en las democracias europeas, dando cuenta de su uso en países como Inglaterra, Alemania, EE. UU.., Brasil, Países Bajos o Italia. En los Países Bajos, varios municipios utilizan reconocimiento facial durante los carnavales y otros grandes eventos y desde 2016 la policía holandesa utiliza el sistema de reconocimiento facial CATCH a través de las imágenes de los teléfonos inteligentes, las cámaras corporales y la nube. En Italia el Garante italiano de la protección de datos considero inadmisible el 16 de abril de 2021 el «Sistema Automatico di Riconoscimento Immagini» *SARI*, utilizado desde 2019. En Alemania, en Hamburgo con motivo de una reunión del G20 en 2017 se implantó un sistema de reconocimiento facial a partir de grabaciones para la detección e investigación de delitos. Se usa también esta identificación por empresas privadas, sobre todo almacenes y tiendas, y así en España, la cadena de supermercados *Mercadona*, recibió una fuerte sanción por implantar un sistema inteligente biométrico que controlaba si quienes accedían a algunos establecimientos estaban en sus listas de «personas con una orden de alejamiento o medida judicial análoga en vigor»[49].

3.2.2. El perfilado predictivo

El perfilado se define en el art. 4.4 del Reglamento (UE) 2016/679 de protección de datos (RGPD): *4) «elaboración de perfiles»: toda forma de tratamiento automatizado de datos personales consistente en utilizar datos personales para evaluar determinados aspectos personales de una persona física, en particular para analizar o predecir aspectos relativos al rendimiento profesional, situación económica, salud, preferencias personales, intereses, fiabilidad, comportamiento, ubicación o movimientos de dicha persona física.* Surge como técnica que parte de la obtención de datos del sujeto, para empezar el de su identidad personal y digital (su dirección IP, número de móvil, dirección de correo electrónico, avatar o *nickname*), y otros acerca de sus intereses vitales (políticos, deportivos, ideológicos o religiosos), nivel económico y de gasto, preferencias de consumo y de entretenimiento, etc. Su objeto es obtener una imagen del sujeto que permita abordarlo con mensajes de todo tipo, desde publicidad y ofertas de turismo o inmobiliarias a información política o económica, con el objeto de provocar decisiones de compra o actuaciones

48 Cotino Hueso, «Reconocimiento facial automatizado y Sistemas de identificación biométrica bajo la regulación superpuesta de Inteligencia artificial y protección de datos», *Derecho Público de la inteligencia artificial,* Coordinadores Balaguer Callejón, Francisco, Cotino Hueso, Lorenzo, Fundación Manuel Giménez Abad, Zaragoza, 2023, p. 348.

49 Resolución de la Agencia Española de Protección de Datos, procedimiento sancionador PS 120/2022

de cualquier tipo en dicho sujeto-objetivo. Todos tenemos la experiencia de los reclamos que llegan a nuestro móvil, comerciales o de ocio.

La incidencia del perfilado en los derechos fundamentales tiene lugar sobre todo en el derecho a la intimidad y privacidad: su funcionamiento implica la creación de bases de datos de características personales de conjuntos de sujetos muy amplios. Bien es verdad que en muchos casos el identificador del sujeto no es sino una dirección IP, pero a partir de ahí y relacionando este dato con otros, es fácil llegar a la identidad real del mismo. Desde una finalidad sobre todo mercantil y publicitaria, el perfilado evoluciona a un ámbito de predicción de la conducta futura del sujeto, partiendo de los datos obtenidos y de las conductas observadas en el pasado, y de datos de tipo emocional obtenidos indirectamente.

El perfilado implica la clasificación de las personas según las características que se determinen en el algoritmo. A esto se refiere el art. 3 de la LIA cuando define en su número 35 a un «Sistema de categorización biométrica» como …*sistema de IA destinado a asignar a personas físicas a categorías concretas, como un sexo, edad, color de pelo, color de ojos, tatuajes, origen étnico u orientación sexual o política…* Pero el perfilado va más allá, añadiendo datos económicos e intereses de consumo. O peor, datos médicos y neurológicos, pues como nos cuenta **González Tapia**[50], mediante inteligencia artificial se puede lograr una extracción indirecta de datos de las personas a través de patrones de comportamiento, detección de emociones, tono y modulación del lenguaje, y datos biométricos. Y pone como ejemplo «… la detección temprana de enfermedades que puede hacerse, entre otros medios, a través del patrón de tecleo en el móvil con relación al Parkinson, de deambulación con respecto al Alzheimer o patrones de atención… Así mismo, datos relativos a navegación por las redes sociales y el análisis de los mensajes arrojados en ellas, permiten detectar riesgos como comportamientos o ideaciones suicidas; datos obtenidos en el entorno laboral sobre las emociones sentidas por sus trabajadores, pueden informar sobre su motivación».

Además del marketing y la publicidad, otro ámbito en el que se desarrolla el perfilado es el de la averiguación de tendencias políticas, pues como señala **Garriga Domínguez**[51], el desarrollo actual de la tecnología de IA permite el perfilado ideológico de los usuarios de cualquier plataforma, tanto con finalidades de marketing como políticas: «… La aplicación de la IA en estos ámbitos permite el perfilado ideológico individual y, a través de las técnicas de focalización podrá elaborarse información política personalizada,

50 González Tapia, M.ª Isabel, «Protección penal de los neuroderechos: el uso directo de las Neurotecnologías sobre el ser humano», *Inteligencia Artificial y Filosofía del Derecho*, Fernando H. Llano Alonso, Director, Ediciones Laborum, Murcia, 2022, p. 328.

51 Garriga Domínguez, Ana, «Inteligencia artificial y el fenómeno de la desinformación: el papel del RGPD1 y las garantías recogidas en la propuesta de la ley de servicios digitales», en *Inteligencia Artificial y Filosofía del Derecho*, Ediciones Laborum, Murcia, 2022, p. 452.

El desarrollo de los procesos de segmentación de mercados ha evolucionado hacia una segmentación psicográfica avanzada, que se basa en un algoritmo que determina una serie de rasgos demográficos y de actitud que permite distinguir a cada individuo para cada segmento objetivo y que permite hacer predicciones precisas de la reacción de la audiencia objetiva … la cantidad y calidad de la información personal que se encuentra en las redes sociales, permite a los anunciantes mejorar el alcance e impacto de su publicidad … Obviamente, estas técnicas pueden utilizarse para vender un producto determinado, pero también para favorecer una determinada ideología».

Naturalmente, a partir de esta personalización, se abre la vía hacia la interpretación interesada, el *nudge* político y la información parcial o sesgada. En suma, la desinformación del sujeto, cuyas decisiones dejan de ser libres porque no han sido libremente formadas. Como ejemplo de esto, es bien conocido el escándalo de *Cambridge Analytica*, una consultora política británica, que obtuvo los datos de millones de usuarios de Facebook sin su consentimiento y mediante sistemas de IA los utilizó para crear perfiles psicológicos. Estos perfiles luego se utilizaron para ofrecer anuncios políticos personalizados en la campaña presidencial americana de 2016. Este campo del perfilado proporciona información para luego proceder a enviar *nudges* que influyan en la conducta de los sujetos, concepto éste del que luego nos ocuparemos.

También nos advierten de las consecuencias negativas del perfilado **G.ª Inda, González Ordovás,** y **Muñoz Soro**[52], que, citando a José María Lasalle, dicen: «…"se nos neutraliza en la capacidad decisoria jugando con la sutileza algorítmica que hace que las personas no actúen según su juicio sino que se optimice este en nuestro propio beneficio. (…) Se disuelve la espontaneidad de la conducta bajo el peso de una mezcolanza de datos que se manipulan desde fuera de nosotros para determinarnos". Dicho con otras palabras, la IA puede ser un recurso para mejorar las condiciones de vida y ampliar la libertad del ser humano, compensando sus limitaciones "naturales" y facilitando la toma de decisiones en contextos complejos; o puede convertirse en un instrumento de dominación, limitando o suplantando esa libertad y esa capacidad de decisión en función de intereses espurios». Como dicen las *Directrices éticas para una IA fiable* (parágrafo 50), la IA puede ser un medio para «subordinar, coaccionar, engañar, manipular, condicionar o dirigir a los seres humanos de manera injustificada».

El perfilado, como estudio probabilístico o actuarial del individuo y la posibilidad de que observe un determinado comportamiento, ha sido desarrollado sobre todo en el ámbito penal, donde **Solar Cayón**[53] advierte

52 García Inda, González Ordovás y Muñoz Soro, «IA y Filosofía del Derecho…», en Proyecto IASAC cit., p. 6.

53 Solar Cayón, José Ignacio, «Inteligencia artificial y justicia digital», en *Inteligencia Artificial y Filosofía del Derecho*, Ediciones Laborum, Murcia, 2022, p. 382.

que es habitualmente empleado en la mayoría de jurisdicciones estatales para informar las decisiones judiciales sobre medidas cautelares, en particular la concesión o no de libertad provisional. Señala que estos sistemas de evaluación de riesgos de reincidencia criminal se basan en modelos estadísticos «… generados automáticamente mediante técnicas de *machine learning* que, a partir del análisis de grandes volúmenes de datos correspondientes a casos pretéritos, son capaces de detectar una serie de correlaciones entre determinados factores personales y sociales y el riesgo de comisión de futuros delitos… indicadores relativos a las circunstancias personales del acusado (edad y sexo, nivel de estudios, contexto familiar, situación socio-laboral, consumo de drogas…), elementos socio-demográficos (lugar de residencia, contexto socioeconómico, relaciones sociales…) y su historial judicial (detenciones y delitos previos, historial de violencia, precedentes de incomparecencia ante el tribunal…)»[54]. Todos estos datos son factores predictores del riesgo de reincidencia futura, asignando a cada uno un valor en función del análisis de los casos pretéritos. Por su parte, **Juliá-Pijoan**[55] nos señala, como ejemplos de estos sistemas, el programa americano COMPAS *(Correctional Offender Management Profiling for Alternative Sanctions)* o el catalán RISCANVI, programas de gestión penitenciaria que en base a datos personales del sujeto detectan su posibilidad de reincidencia delictiva. Y advierte que el sesgo en el que pueden incurrir estos programas es el de la búsqueda de diferencias cerebrales a partir de un modelo de «normalidad cerebral» puramente arbitrario.

Naturalmente, el siguiente paso consistirá en pasar de la ayuda para la decisión administrativa a la decisión judicial automática, a los «jueces-robot», como nos cuenta **De Asís Pulido**[56] que ya funcionan en China, si bien todavía en funciones de apoyo y dejando la decisión final al juez humano, como el programa Xiao Zhi, en la Corte Suprema Popular de China: «… esta máquina organiza los eventos del proceso, analiza la presentación de los casos en lo relativo a su admisibilidad, resume los puntos en los que las partes están en desacuerdo, ayuda en la evaluación de las pruebas y crea propuestas de re-

54 También señala este autor que los sistemas de decisión automática ya se utilizan en España en los procedimientos administrativos sancionadores, así 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, tras la reforma operada por el Real Decreto 688/2021, de 3 de agosto, permite a la Inspección detectar incumplimientos basados en el análisis masivo de datos sin que se requiera la intervención directa de ningún funcionario.

55 Juliá-Pijoan, Miquel, «Una aproximación al perfilaje criminal desde la investigación neurocientífica», *FODERTICS 11.0 Derecho, entornos virtuales y tecnologías emergentes,* Comares, Granada, 2023, pp. 443 y ss.

56 González Tapia, M.ª Isabel, «Protección penal de los neuroderechos: el uso directo de las neurotecnologías sobre el ser humano», *Inteligencia Artificial y Filosofía del Derecho*, Ediciones Laborum, Murcia, 2022, p. 328.

soluciones judiciales (Chen/Li, 2020, 15)». Estos sistemas pueden funcionar tanto en modo decisorio como en predictivo, pero este tema excede de los límites de este breve trabajo. O incluso cabe dar un paso más y, como nos dicen los argentinos **Salvi y Nigri**[57], entrar en el terreno de la ciencia-ficción, configurando sistemas predictivos que se anticipen a la comisión del crimen, como se ve en la película *Minority Report*.

La posibilidad de perfilado ya la vaticinó Víctor **Drummond**[58], al considerar que las simples *cookies*, pese a su aspecto inofensivo, permitían un control de la navegación en Internet del sujeto y de sus perfiles y hábitos de consumo: «Esta problemática es propia de los sitios web que practican el comercio electrónico, ya que cuando se efectúa una compra o cualquier otro negocio en la red, el usuario deberá proporcionar voluntariamente una serie de datos personales. En ese momento, su nombre, dirección, número de tarjeta de crédito, entre otros datos, se integrarán automáticamente a una misma base de datos, junto con otras informaciones involuntariamente recogidas a través de las cookies... Este cruce de datos es un ejemplo auténtico de la transformación de datos en un principio irrelevantes en un perfil peligrosamente público del ciudadano».

Las cookies ya están superadas, el problema actual lo plantean las técnicas de detección de la huella digital, de las que nos habla **Fernando Pablo**[59]: Estas técnicas permiten el seguimiento de comportamiento en Internet de los usuarios finales, sin la protección de la Directiva 2002/58/CE, información que se puede recopilar con fines de identificación y seguimiento, utilizando técnicas que se describen como «toma de huellas digital del dispositivo» (generalmente sin el conocimiento del usuario), generando datos sobre sus actividades en línea o la ubicación física de su equipo. Como señala la Propuesta de Reglamento del Parlamento Europeo y del Consejo sobre el respeto de la vida privada y la protección de los datos personales en el sector de las comunicaciones electrónicas[60] en su Considerando 15, *…a medida que la tecnología avanza, aumentan también los medios técnicos para la interceptación. Dichos medios pueden abarcar desde la instalación de equipos que recopilan datos de los equipos terminales de las zonas seleccionadas, como los deno-*

57 Salvi, Nicolás y Nigri, Santiago, «Minority Report: the Road to a Deterministic Theory for the Philosophy of Criminal Law», June 28th, 2022, *Opinión Jurídica*, 21(46), Special Edition 2022, pp. 1-18, https://doi.org/10.22395/ojum.v21n46a2

58 Drummond, Víctor, *Internet, privacidad y datos personales*, Traducción y notas de Isabel Espín Alba, Reus, Madrid, 2004, pp. 117-118.

59 Fernando Pablo, Marcos M., «Capítulo I: Construyendo la dignidad digital de la persona en el entorno digital. De los datos de tráfico, a la plaza y mercado de los servicios de comunicaciones electrónicas», en *Desafíos éticos, jurídicos y tecnológicos del avance digital* / coord. por Alicia Rodríguez Sánchez, Pilar Talavera Cordero; José Luis Domínguez Álvarez (dir.), Daniel Terrón Santos (dir.) Iustel, Madrid, 2023, págs. 21-40. P. 33.

60 COM/2017/010 final - 2017/03 (COD

minados receptores de IMSI (identidad internacional de abonado móvil), hasta algunos programas y técnicas que, por ejemplo, efectúan un seguimiento subrepticio de los hábitos de navegación para crear perfiles de usuarios finales.

3.2.3. El reconocimiento de emociones y la detección del pensamiento

Ya hemos visto que uno de los campos en los que la captación de datos biométricos es muy efectiva es el de los datos sobre frecuencia cardiaca, respiración, expresión facial, movimiento de las pupilas, temperatura corporal, cuya combinación puede indicar la presencia de determinados estados emocionales. Por ejemplo, el software de la compañía española *Decoditive* permite determinar mediante una sencilla cámara si el sujeto está haciendo trampas al ajedrez o la eficacia real de los anuncios que está contemplando[61], simplemente examinando el campo visual y movimiento de las pupilas del sujeto.

La misma Ley de la IA hace expresa referencia a la utilización de los datos biométricos para la labor policial en los interrogatorios, y la averiguación de la veracidad o falsedad de las declaraciones de las personas. Se está refiriendo a datos que revelan estados emocionales, tales como frecuencia cardiaca, sudoración, movimiento de las pupilas, tono de voz o incluso frecuencia y fuerza de pulsación de las teclas del ordenador. Es en concreto en el Considerando 38 en el que se expone que son de alto riesgo los sistemas de IA para la aplicación de la ley, entre los que deben incluirse *…en particular, los sistemas de IA que las autoridades encargadas de la aplicación de la ley utilicen para realizar evaluaciones del riesgo individuales, los polígrafos y herramientas similares, o los sistemas utilizados para detectar el estado emocional de una persona física*. Y en su art. 3.34 la LIA define al «Sistema de reconocimiento de emociones» como *un sistema de IA destinado a detectar o deducir las emociones o las intenciones de personas físicas a partir de sus datos biométricos*.

Da cuenta **Cotino Hueso**[62] que estos sistemas permiten leer emociones, detectar la verdad de las manifestaciones, y predecir futuros comportamientos y que desde hace años se utilizan para el control de fronteras en EE. UU., con el denominado Agente Virtual Automatizado para la Evaluación de la

61 «El CEO de Decoditive, Joan Buch, recogió el premio en la categoría e-commerce de la Barcelona New Economy Week (BNEW 2021) de manos de la ministra Nadia Calviño». El Español, 21 octubre 2021, en https://www.elespanol.com/invertia/disruptores-innovadores/disruptores/startups/20211028/startup-cazaba-tramposos-ajedrez-atrapar-clientes-neuromarketing/622687953_0.html

62 Cotino Hueso, «Reconocimiento facial automatizado y Sistemas de identificación biométrica bajo la regulación superpuesta de Inteligencia artificial y protección de datos», cit., p. 348.

Verdad en Tiempo Real (AVATAR - *Automated Virtual Agent for Truth Assessments in Real Time*), que analiza el comportamiento no verbal y verbal de los viajeros. En Europa, la Comisión Europea financió el proyecto *Intelligent Portable Control System (iBorderCtrl*), con herramientas de detección del engaño y de evaluación de respuestas de solicitantes de visa de entrada, que generó una relativa reacción desde la sociedad civil, y que finalmente ha sido retirado. Las grandes compañías de aplicaciones informáticas han producido sistemas de reconocimiento biométrico y facial para identificar personas y detectar emociones y estados de ánimo, así en junio de 2022, Microsoft anuncia que retira sus sistemas *Azure Face.*

Y al lado de estas técnicas, tenemos la IA aplicada a una nueva rama de la neurología, la neurotecnología, que define **Moreu Carbonell**[63] de modo muy amplio como el conjunto de métodos e instrumentos que permiten una conexión directa de dispositivos técnicos con el cerebro y el sistema nervioso, con independencia de si se trata de técnicas de estimulación cerebral invasivas o no invasivas: «La neurotecnología es un concepto interdisciplinar en el que confluyen la inteligencia artificial, la informática y las neurociencias. Las investigaciones del cerebro pueden ya medir, registrar, alterar y manipular la actividad cerebral; es lo que se conoce como neuromodulación o alteración de la actividad cerebral por medio de la introducción de estímulos».

Hoy por hoy, la neurotecnología tiene un uso beneficioso para la humanidad, pero presenta riesgos jurídicos, éticos y morales, amenazando, como nos dice la autora, la privacidad y la seguridad de las personas, y la propia definición de persona. Para evitar estos males, propone la autora el desarrollo de una nueva especialidad jurídica, el Neuroderecho, y la conformación de nuevos neuroderechos de las personas.

La nueva tecnología supone un paso más en la invasión de la privacidad, pues ya no se trata de inferir estados emocionales a partir de datos biométricos, sino directamente de acceder a los pensamientos de la persona mediante una lectura de los patrones de estímulo eléctrico cerebral. Estamos ante la lectura de los pensamientos, cuya aplicación puede ser muy beneficiosa en medicina, como recogen **Vicente Domingo** y **Rodríguez Cachón**[64], refiriéndose al proyecto *Neuralink* impulsado por Elon Musk, que diseña implantes cerebrales para que los pacientes con parálisis cerebral puedan controlar dispositivos con la mente. Y también *Facebook*, hoy *Meta*, ha iniciado un programa de 40 millones de dólares para conseguir mediante cascos de electrodos no invasivos, convertir en texto lo que está pensando una persona, sin

63 Moreu Carbonell, Elisa, «La regulación de los neuroderechos», *Revista General de Legislación y Jurisprudencia*, 2022, n.º 1, enero-marzo, pp. 71-100, p. 72.

64 Vicente Domingo, Elena y Rodríguez Cachón, Teresa, «Derecho de la Persona, Neurodatos y Neuroderechos: *A Research Agenda*», *Revista General de Legislación y Jurisprudencia*, 2023, número 3, pp. 495-526, p. 499.

necesidad de implantes o scanner cerebral. Mientras que, desde las universidades de Singapur y Hong Kong, los investigadores Chen, Qing y Zhouy[65] dan cuenta de la posibilidad de decodificación directa de las señales cerebrales en imágenes y videos, utilizando la resonancia magnética. Mediante esta técnica, se capturan y codifican datos sobre la actividad cerebral del sujeto, y un sistema de reconocimiento traduce dichos datos a imágenes, en algunos casos muy fidedignas.

Los neuroderechos adquieren por tanto su verdadero significado frente a esta neurotecnología inteligente, que como vemos es especialmente invasiva. Recoge **Moreu Carbonell**[66] de los investigadores Marcello Ienca y Roberto Andorno cuatro neuroderechos: Derecho a la libertad cognitiva o «autodeterminación mental», una actualización del derecho a la libertad de pensamiento y de conciencia. Derecho a la privacidad mental, que protege frente a cualquier información que pueda obtenerse de nuestros cerebros por medio de neurotecnologías. Derecho a la integridad mental, contra las intrusiones en el cerebro. Y, por último, derecho a la continuidad psicológica, que garantiza la percepción de la propia identidad como seres humanos.

Vicente Domingo y **Rodríguez Cachón**[67], por su parte, recogen la enumeración de la *NeuroRights Foundation*, del neurólogo Rafael Yuste (derecho a la identidad e integridad personal y mental, al libre albedrío, a la privacidad mental, al acceso equitativo y a la protección contra los sesgos), y afirman que una eventual aceptación de esta propuesta de los neuroderechos como nueva categoría de Derechos Humanos llevaría aparejada una modificación de la Declaración Universal de los Derechos Humanos.

Sin embargo, para otros no es tan urgente crear nuevos neuroderechos, y en particular un neuroderecho a la privacidad mental, como ocurre con **Ligthart, Douglas, Bublitz** *(et al.)*[68], que opinan que no es necesario, en especial en el ámbito de la lectura mental forense, bastando con ampliar la protección derivada de la jurisprudencia del Tribunal europeo de Derecho humanos a estos casos.

---

65 Chen Zijiao, Qing Jiaxin y Zhouy Juan Helen, «CinematicMindscapes: High-quality Video Reconstruction from Brain Activity», preprint en *arXiv*:2305.11675, https://doi.org/10.48550/arXiv.2305.11675

66 Moreu Carbonell, «La regulación de los neuroderechos», cit., p. 83.

67 Vicente Domingo y Rodríguez Cachón, «Derecho de la Persona, Neurodatos y Neuroderechos: *A Research Agenda*», cit. P. 513.

68 Ligthart, Sjors, Douglas, Thomas, Bublitz, Christoph, Kooijmans, Tijs y Meynen, Gerben (2021), «Forensic Brain-Reading and Mental Privacy in European Human Rights Law: Foundations and Challenges», Neuroethics, 2021, 14, pp. 191-203, https://doi.org/10.1007/s12152-020-09438-4

3.3 Las técnicas subliminales y el *nudge*

La Ley de la IA también se ocupa, para declararlas prácticas prohibidas en su artículo 5 —como ya se ha dicho—, de los sistemas de IA que utilicen «…técnicas subliminales que trasciendan la conciencia de una persona para alterar de manera sustancial su comportamiento de un modo que provoque … perjuicios físicos o psicológicos a esa persona o a otra». O bien sistemas que aprovechen vulnerabilidades de un grupo específico de personas, ya por edad o discapacidad física o mental para alterar de manera sustancial el comportamiento de personas de dicho grupo.

En cuanto a estas técnicas, ha señalado **Ebers**[69] que ni el artículo 5 ni el Considerando 16 de la LIA definen qué es una «técnica subliminal», o en qué consiste alterar *sustancialmente* el comportamiento de una persona. Y además este art. 5 sólo cubre las prácticas de IA que tienen la intención directa de perjudicar a otras personas. En realidad, la mayoría de estas prácticas serán ilegales conforme al Derecho penal en la mayoría de los Estados, si lo que intentan es engañar a las personas (posiblemente para estafarlas), y además la Directiva sobre prácticas comerciales desleales (2005/29/CE) ya prohíbe las prácticas comerciales que distorsionan el comportamiento humano en determinadas condiciones.

Como advierte **de Miguel Asensio**[70], todas estas técnicas están prohibidas por el art. 5 de la LIA por generar riesgos inadmisibles que contravienen los valores de la Unión europea. Sin embargo, opina el autor que además de esta normativa, el perjudicado por sistemas inteligentes puede acudir a la normativa de protección del honor, intimidad y propia imagen, a la de protección de datos personales e incluso a la de protección del consumidor y a la de responsabilidad por productos defectuosos. Por ello el autor estima que la nueva LIA viene a cubrir estos espacios desprotegidos, los daños de difícil encuadre en las figuras «normales» del daño.

¿A qué técnicas de modificación de comportamientos se puede referir la LIA? La influencia inconsciente en el comportamiento de las personas fue el tema de una famosa obra del año 1957, *The Hidden Persuaders*, de Vance Packard (traducida al español como *Las formas ocultas de la propaganda*). En este libro, dedicado a la publicidad y su influencia en los consumidores, se toma como premisa la falta de racionalidad de muchos comportamientos de consumo, y la anécdota que mejor muestra este aspecto es la siguiente: «Un bazar cuyos dueños consideraban con creciente escepticismo la racionalidad de sus clientes puso a prueba un experimento. Uno de los renglones de menor salida era un artículo que valía catorce centavos. Cambió el precio

69 Ebers, «El futuro marco jurídico europeo de la inteligencia artificial», cit., p. 265.

70 De Miguel Asensio, Pedro A., «Propuesta de Reglamento sobre Inteligencia artificial», *La Ley Unión Europea*, n.º 92, mayo 2021, p. 4.

ofreciendo dos de dichos artículos por veintinueve centavos. Las ventas aumentaron rápidamente en un 30 % al ofrecerlo a precio "rebajado"»[71].

Es decir, que las técnicas de marketing y publicidad investigan los motivos que realmente deciden las conductas del consumidor, y nos dicen que la compra del producto no se basa en la lógica y racionalidad. Manipulando las imágenes asociadas al producto, se intenta influir sobre las conductas de compra del consumidor: el hecho básico es que el ser humano es manipulable.

Más allá del marketing, una de las técnicas que se usa en la actualidad para influir en los comportamientos sociales es el *nudging*. Señalan **Costas Pérez** y **Tucat**[72] que los *nudges* o «empujones» se popularizan a partir de la publicación del libro *Nudge: Improving Decisions About Health, Wealth, and Happiness,* de Cass Thaler y Richard Sunstein, en 2009. Los *nudges* (palabra que significa pequeños «empujones» o «codazos»; también se ha traducido como «acicates») son intervenciones que buscan modificar la toma de decisiones individuales, intentando cambiar el comportamiento de las personas en una dirección concreta, pero sin prohibir ninguna opción, ni alterar en gran medida los incentivos económicos. Es decir, que no son órdenes coercitivas para los ciudadanos, sino que inducen a un comportamiento que se considera adecuado o deseable, como señalaban Thaler y Sunstein.

Los *nudges* actúan precisamente partiendo de la posibilidad de influencia en el comportamiento humano de toma de decisiones, pues como hemos visto muchas decisiones se toman siguiendo criterios no estrictamente racionales sino emocionales. Esta técnica pasa del mundo del marketing al de la Administración pública, siendo utilizado para fomentar las conductas que resultan más beneficiosas (aunque no se concreta para quién lo son, si para la Administración o los ciudadanos)[73]. Como señalan los indicados autores,

---

71 Packard, Vance, *Las formas ocultas de la propaganda*, Editorial Sudamericana, Buenos Aires, Decimooctava Edición, Junio de 1992, p. 15.

72 Costas Pérez, Elena, y Tucat, Pablo, (2021). «Nudges: diseño y evaluación», *Gestión y Análisis de Políticas Públicas*, (25), 8-22. DOI: https://doi.org/10.24965/gapp.i25.10868, Pág. 9

73 Cuentan Antonio Cabrales Goitia y Pedro Rey Biel que según el modelo de Bemelmans-Videc las políticas públicas se clasifican como *palos, zanahorias o sermones*: «Los palos son herramientas reguladoras que buscan forzar el comportamiento de los ciudadanos. Las zanahorias proveen de incentivos para seducir a los individuos, mientras que los sermones, buscan persuadirles. Clasificar a los *nudges* en uno de estos tres grupos no es tarea sencilla. Los *nudges* claramente no son palos, puesto que no restringen las opciones disponibles para el individuo ni le castigan si no se comporta como el regulador pretende. Tampoco son zanahorias, puesto que buscan guiar al ciudadano de una forma inconsciente ... Por último, los *nudges* tampoco son sermones, pues no pretenden persuadir a los ciudadanos de forma abierta mediante la provisión de información… Los *nudges* se encuentran en un término medio entre la zanahoria y el sermón» («Mas allá de los *nudges*: Políticas públicas efectivas basadas en la evidencia de las ciencias del comportamiento», *Gestión y Análisis de Políticas Públicas*, (25), Pág. 40, DOI: https://doi.org/10.24965/gapp.i25.10864).

los *nudges* en las políticas públicas utilizan «los conocimientos de las ciencias económicas, la neurociencia y la psicología, para incentivar o desincentivar actuaciones concretas de los distintos agentes, y alcanzar así objetivos establecidos por los dirigentes públicos»[74].

Por su carácter no normativo, los *nudges* constituyen un importante instrumento de políticas públicas en la acción de gobierno que escapa a un control político. El ejemplo más citado es el de incluir en los documentos médicos la opción por defecto de ser donante de órganos, y estableciendo por tanto que el rechazo ha de ser expreso. Con esta modificación, el porcentaje de donantes es significativamente mayor que si la opción de ser donante se tiene que elegir expresamente. Como señala **Ortiz de Zárate-Alcarazo**[75], el *nudging* milita en el campo de lo políticamente correcto: «los *nudges* son un tipo de intervenciones que, aprovechando las limitaciones cognitivas humanas, nos ponen en el "buen camino" o la "buena dirección", sin consecuencias negativas para aquella persona que elige no seguir el camino recomendado. Por el carácter orientador de los *nudges* hacia ciertos objetivos o metas y, simultáneamente, su respeto por la autonomía y libertad individual de las personas, al *nudging* también se le denomina paternalismo liberal (libertarian paternalism) …Por tanto, los *nudges* son mecanismos coercitivos alineados con los valores de las democracias liberales (Strabheim, 2020)».

Pues bien, la unión de estas técnicas de *nudging* con la IA produce interesantes resultados. Como advierte **Ortiz de Zárate-Alcarazo**[76], la aplicación de la IA puede mejorar la implementación de *nudges* en el sector público, e incluso puede llegar a transformar este tipo de intervenciones dando lugar a lo que podríamos denominar «políticas conductuales inteligentes y, más concretamente, *intelligent nudging*. Los *nudges* inteligentes tendrían niveles de eficacia y eficiencia más elevados que los de sus hermanos no inteligentes gracias al uso de tecnologías de reconocimiento facial, detección de objetos, procesamiento del lenguaje natural, predicción, análisis en tiempo real, etc., que permitirían mejorar todos los pasos del proceso».

En este sentido, el investigador de la Universidad de Seikei, Yukari **Yamazaki**[77] nos anuncia la llegada del *Hypernudge*, como resultado de la unión del *nudge* y la IA, es decir, el *nudge* creado por una IA entrenada mediante *machine*

74 Costas Pérez y Tucat, (2021). «Nudges: diseño y evaluación», cit., p. 9

75 Ortiz de Zárate-Alcarazo, Lucía (2023). «Las Políticas conductuales inteligentes: Oportunidades y riesgos ético-políticos de la Inteligencia Artificial para el nudging», *Revista Española de Ciencia Política*, 62, 67-93. Doi: https://doi.org/10.21308/recp.62.03. P. 69.

76 Ortiz de Zárate-Alcarazo, «Políticas conductuales inteligentes: Oportunidades y riesgos ético-políticos de la Inteligencia Artificial para el nudging», cit., p. 82.

77 Yamazaki, Yukari, «An Empirical Study for The Acceptance of Original Nudges and Hypernudges», *Societal Challenges in the Smart Society* / coord. por Mario Arias Oliva, Jorge Pelegrín Borondo, Kiyoshi Murata, Ana María Lara Palma, 2020, ISBN 978-84-09-20273-7, págs. 323-336. En Dialnet: https://dialnet.unirioja.es/servlet/articulo?codigo=7867256

learning, y advierte que las condiciones de autonomía, dignidad y transparencia, que deben regir el empleo de *nudges*, desaparecen en este caso. En base a un estudio estadístico, concluye que hay que estar alerta en la utilización de los *hypernudges* por su menor aceptación que los *nudges* originales. En suma, el *hypernudge* añade a los problemas intrínsecos del *nudge* los problemas específicos derivados de la IA, especialmente la derivada de redes neuronales, como la opacidad algorítmica, los problemas de privacidad, los sesgos, la falta de rendición de cuentas, y sobre todo el ataque a la privacidad e intimidad de las personas, derivando en un paternalismo agobiante.

Desde la Filosofía del Derecho, señala **de Asís Roig**[78] que los *nudges* se introducen en el ámbito de las aplicaciones tecnológicas, para actuar como consejeros morales en sistemas inteligentes. Cita así los modelos computacionales *Truth-Teller* y *Sirocco,* y el programa *MeEthEx*, que es una especie de asesor ético en el campo de la medicina para ayudar a resolver dilemas éticos, y que se apoya en los principios de ética biomédica. Otros modelos se configuran como asistentes-guía en la toma de decisiones, como el modelo computacional diseñado por Robbins y Wallace, «… herramienta para ayuda en la toma de decisiones basado en el modelo creencia-deseo-intención y en la resolución colaborativa de problemas, simulando diferentes papeles (asesor, facilitador de grupo, entrenador de interacción y pronosticador). O, también la propuesta de F. Lara, …de un asistente virtual para fortalecer nuestra moralidad potenciando la autonomía personal. Se trata de un asistente basado en el método dialéctico socrático si bien proyectado hacia el aprendizaje moral. El asistente, que denomina *SocrAI*, lo que pretende es formar al usuario no tanto en principios éticos sustantivos sino en pautas generales sobre cómo razonar mejor». Coincide el autor en la peligrosidad de los *hypernudges*, por la posibilidad de manipulación que conllevan. Naturalmente, el problema de estas técnicas potenciadoras de la eficacia mediante la IA, es que también potencian los problemas del *nudging*, como son la disminución de la libertad individual, de la capacidad de decisión y elección de los ciudadanos, la manipulación y el determinismo. Por ello, **Ortiz de Zárate-Alcarazo**[79] matiza su consideración acerca de estas técnicas diciendo: «En este sentido, la posible perversión de las políticas conductuales inteligentes y su principal riesgo sería el de generar escenarios en los que la ciudadanía tuviera importantes dificultades para salirse del camino establecido y quedasen fácilmente atrapados en senderos de dependencia».

Más allá del *nudge* veremos ahora una técnica de vigilancia y represión (y no modificación) de los comportamientos de los ciudadanos de consecuencias aterradoras: la calificación social.

---

78 De Asís Roig, «Ética, Tecnología y Derechos», cit. p. 33.

79 Ortiz De Zárate-Alcarazo, «Políticas conductuales inteligentes: Oportunidades y riesgos ético-políticos de la Inteligencia Artificial para el nudging», cit., p. 85.

3.4. Los sistemas de calificación o crédito social

A esta técnica, utilizada tanto para influir en el comportamiento de las personas como para controlarlo, alude el art. 5 LIA al prohibir los sistemas de IA utilizados por autoridades públicas o en su representación, *c) …con el fin de evaluar o clasificar la fiabilidad de personas físicas durante un período determinado de tiempo atendiendo a su conducta social o a características personales o de su personalidad conocidas o predichas…*[80]. Estamos ante un sistema por el que cada persona, convenientemente identificada por técnicas biométricas, de identificación y rastreo en red o por teléfono móvil, recibe una puntuación o créditos, positivos o negativos según su comportamiento sea considerado aceptable o inaceptable, cuya suma determina su calificación o crédito social. El origen del invento está en algo que, al principio —como tantas otras cosas— pareció una buena idea, en concreto en las calificaciones crediticias y las listas de morosos, información comercial que existe en cualquier país (así en España las listas ASNEF, CIRBE o RAI) combinadas para que actuasen como *nudge* para los comportamientos sociales adecuados y la promoción de la confianza. Pero luego esto se complica cuando se amplía la calificación con los ingredientes de las plataformas sociales y bases de datos, las tecnologías de reconocimiento biométrico y las de seguimiento de la actividad de los teléfonos móviles, y de la mera calificación crediticia a efectos de valorar la solvencia del sujeto se pasa a valorar al sujeto en sí, a partir de todo su comportamiento.

Apunta **Cotino Hueso**[81] que estos preocupantes sistemas biométricos de categorización, si son privados, ni están prohibidos, ni en general son de alto riesgo, sino solo sometidos al artículo 52 de la Ley de IA: Artículo 52. Obligaciones de transparencia para determinados sistemas de IA … *2. Los usuarios de un sistema de reconocimiento de emociones o de un sistema de categorización biométrica informarán del funcionamiento del sistema a las personas físicas expuestas a él. Esta obligación no se aplicará a los sistemas de IA utilizados para la categorización biométrica autorizados por la ley para fines de detección, prevención e investigación de infracciones penales*. Opina el autor que sería más adecuado que o bien se regulen como sistemas de alto riesgo o en algunos casos directamente se prohíban.

---

80 Sigue el precepto: *…de forma que la clasificación social resultante provoque una o varias de las situaciones siguientes: i) un trato perjudicial o desfavorable hacia determinadas personas físicas o colectivos enteros en contextos sociales que no guarden relación con los contextos donde se generaron o recabaron los datos originalmente; ii) un trato perjudicial o desfavorable hacia determinadas personas físicas o colectivos enteros que es injustificado o desproporcionado con respecto a su comportamiento social o la gravedad de este.*

81 Cotino Hueso, «Sistemas de inteligencia artificial con reconocimiento facial y datos biométricos…», cit., p. 73.

Para **Ebers**[82], hay que aclarar cuándo la calificación social tiene lugar *en representación* de las autoridades públicas, pues estas prácticas se realizan mayoritariamente por el sector privado, extendiendo sus efectos a la administración u otras autoridades públicas. Al restringir la prohibición de la calificación social a las autoridades públicas, la LIA «ignora el uso de tales sistemas por parte de entidades privadas, incluso en ámbitos de alto riesgo que podrían afectar a los derechos fundamentales de las personas»; por ejemplo, en las calificaciones crediticias.

Como todo el mundo sabe, pero intenta olvidar, es el sistema de partido único chino el máximo exponente de la implantación de los sistemas de calificación social. Según nos cuenta Lizzy **Rettinger**[83], la idea del «crédito social» se adecúa perfectamente al carácter chino, por reproducir la idea de la confianza social predicada por Confucio.

Como señalan Roberts, Cowls, Morley y otros[84], esta tecnología viene a ser sancionada por el *Plan de Desarrollo de la Inteligencia Artificial de la Nueva Generación,* aprobado por el supremo órgano del gobierno chino, el Consejo estatal. Este plan tiene tres ámbitos de proyección: Desafío internacional, Desarrollo económico y Gobernabilidad o «Construcción» social, siendo este último el que más nos interesa. El Sistema de Crédito Social todavía no se ha implantado a nivel nacional, pero como nos dicen los citados autores, los ambiciosos objetivos del mismo «…ofrecen un convincente ejemplo de la intención del gobierno de confiar en la tecnología digital, no sólo para gobernanza social, sino también para una regulación más detallada del comportamiento»[85]. Añade **Rettinger**[86] que, en 2018, cuarenta gobiernos municipales y provinciales establecieron planes piloto de sistemas de crédito social, y que la situación actual muestra un sistema fragmentado en tres ámbitos: «sistema de listas negras» a nivel nacional, sistemas de crédito social en determinados municipios, y sistemas de crédito social a efectos financieros pilotados por instituciones financieras.

En definitiva, se trata del control de los comportamientos sociales, cuya implantación es un proyecto perfectamente establecido, como se deduce del documento «Esquema para el establecimiento de un sistema de crédito social» del Consejo estatal para la implantación del sistema, publicado en

82 Ebers, «El futuro marco jurídico europeo de la inteligencia artificial», cit., p. 267.

83 Rettinger, Lizzy, «The Human Rights Implications of China's Social Credit System», *Journal of High Technology Law*, vol. XXI, n.º 1, (2021), pp. 1-33, p. 3.

84 Roberts, Huw, Cowls, Josh, Morley, Jessica *et al.*, «The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation». *AI & Society* 36, 59-77 (2021). https://doi.org/10.1007/s00146-020-00992-2

85 Roberts, Cowls, Morley *et al.*, «The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation», cit. p. 66.

86 Rettinger, «The Human Rights Implications of China's Social Credit System», cit., p. 6.

2014. Este documento subrayó que el Sistema de Crédito Social no solo tenía como objetivo regular las finanzas y acciones corporativas de empresas y ciudadanos, sino el comportamiento social de los individuos, persiguiendo conductas como evasión fiscal, alarmas sobre la seguridad alimentaria y deshonestidad académica, mediante el sistema de «listas negras». Pero a ello hay que añadir otros datos, como que la ciudad de Fuzhou enriquece el currículo social de sus ciudadanos con una cifra que expresa su empleabilidad, según datos de desempeño y constancia en el trabajo; a lo que hay que sumar el desarrollo de ciudades inteligentes, con tecnologías de vigilancia basadas en el reconocimiento facial y seguimiento de teléfonos móviles para rastrear a quienes el gobierno presenta como potenciales disidentes o terroristas, sobre todo de la etnia uigur.

Señala **Cotino Hueso**[87] que, en China, el uso de las tecnologías biométricas combina las tecnologías de identificación, categorización y reconocimiento de emociones para su famoso sistema de crédito social, pero también para el control policial de la ciudadanía, de la lealtad al partido o para el seguimiento de la atención atención de los escolares y evaluación y control mental en el ámbito educativo. Naturalmente, la base para el funcionamiento del sistema de crédito social está en la perfecta identificación tanto de personas como empresas, para lo que China cuenta con un sistema de seguimiento en Internet, un análisis de grandes datos y 626 millones de cámaras de reconocimiento facial[88]. Una vez identificado el sujeto, nos dice **Rettinger**, se publica en las plataformas de crédito social, la nacional o las municipales, como la de Pekín. La reputación de los sujetos deriva de su inclusión en listas negras (mala reputación) o rojas (buena).

Las consecuencias pueden ser muy peligrosas para las personas, pues como señala el periodista Serrano Martínez[89], «los ciudadanos pueden entrar en una lista negra con bastante facilidad, lo que tiene graves consecuencias en la vida real. Acciones tan cotidianas como saltarse un semáforo, fumar en lugares prohibidos, tener deudas impagadas o cometer fraude, además de su correspondiente sanción administrativa, conlleva ciertas restricciones: como la prohibición de viajar en avión o en trenes de alta velocidad, y la compra de artículos de lujo. En algunas ciudades, se publicita la información de las

87 Cotino Hueso, «Sistemas de inteligencia artificial con reconocimiento facial y datos biométricos…», cit., p. 71.

88 Rettinger, «The Human Rights Implications of China's Social Credit System», cit., p. 11.

89 Serrano Martínez, Alejandro, «Crédito social chino: el sistema de puntos que ya se exporta a otras sociedades», *El Economista,* 17/06/2023. Añade el autor que también Rusia quiere completar una red de reconocimiento biométrico combinando sus propios algoritmos de IA con su enorme sistema de vigilancia pública: Así, en enero de 2020, Moscú implementó un nuevo sistema de reconocimiento facial en tiempo real en toda la ciudad, con más de 160.000 cámaras: https://www.eleconomista.es/economia/noticias/12325879/06/23/credito-social-chino-el-sistema-de-puntos-que-ya-se-exporta-a-otras-sociedades.html

personas morosas en pantallas LED de centros comerciales, camiones o paradas del autobús, desvelando datos personales y suponiendo un escarnio social para la persona afectada y su familia. Por si fuera poco, si el crédito social es negativo, consecuencias de estar en una lista negra pueden consistir en que no se puede viajar en avión, o en determinados trenes, no se pueden contratar algunos destinos turísticos, o alojarse en ciertos hoteles, o no se puede inscribir a los hijos en los mejores colegios». Desde esta perspectiva, está claro que las palabras de **Cotino Hueso** sobre la devastadora incidencia de esta técnica sobre los derechos fundamentales no son nada exageradas.

Para **Turley**[90], la realidad es que el gobierno chino está intentado abiertamente crear una sociedad-pecera, en la que ni siquiera sea necesario el control policial. Si se implanta una tecnología de reconocimiento facial completa, las personas serán reacias a asistir a protestas o manifestaciones si el gobierno puede determinar su identidad, y tampoco tendrán contactos con personas o empresas que sean consideradas problemáticas por el gobierno, especialmente teniendo en cuenta las consecuencias de su sistema de «calificación social» o «puntuación ciudadana». Además, gran parte de los esfuerzos del reconocimiento biométrico en China han estado dirigidos a identificar minorías, especialmente los uigures y otras etnias vistas como una amenaza para el régimen comunista.

Zygmunt **Bauman**[91] advertía que las dos antiutopías más célebres en la segunda mitad del siglo pasado, *Un mundo feliz*, de Huxley, y *1984*, de Orwell, compartían el dato de que ambas reflejaban mundos en los que la población estaba estrechamente controlada, y la libertad individual no sólo era inexistente, sino que ofendía gravemente a todos los pobladores de ambas distopías. A algo parecido se está llegando, pues la aceptación del sistema de crédito social entre la población china parece buena. Como nos dicen **Roberts, Cowls y Morley**[92], según una encuesta realizada en China por especialistas occidentales, se detectaron altos niveles de aprobación dentro de la población, aunque ello más por falta de conocimiento de las repercusiones del sistema que por un apoyo explícito. **Rettinger**[93] precisa que, en un estudio efectuado en 2019, hasta un 80 % de los encuestados aprobaba el sistema de crédito social, y solamente un 1 % lo desaprobaba. Decía **Bauman**[94] que el sentimiento de libertad «… implica alcanzar un equilibrio entre los deseos,

90 Turley, «Anonymity, Obscurity, And Technology: Reconsidering Privacy …», cit., pág. 2185.

91 Bauman, Zygmunt, *Modernidad líquida*, Fondo de Cultura Económica, Buenos Aires, 2004, p. 58.

92 Roberts, Cowls, Morley *et al.*, «The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation», cit. p. 67.

93 Rettinger, «The Human Rights Implications of China's Social Credit System», cit., p. 12.

94 Bauman, *Modernidad líquida*, cit., p. 22.

la imaginación y la capacidad de actuar: nos sentimos libres siempre y cuando nuestra imaginación no exceda nuestros verdaderos deseos y ni una ni los otros sobrepasen nuestra capacidad de actuar. Por lo tanto, el equilibrio puede alcanzarse y conservarse inalterable de dos maneras diferentes: agostando, recortando el deseo y/o la imaginación, o ampliando la capacidad de acción». El sistema de crédito social controla a la población haciendo que recorte sus deseos a la necesidad de tener una puntuación alta.

3.5. La falsificación de la realidad

Hemos visto hasta ahora las amenazas que para los derechos fundamentales de las personas suponen los sistemas de IA basados en la utilización de datos biométricos, persuasión subliminal y calificación social, todos los cuales aparecen definidos o mencionados en la LIA y en otros textos europeos.

Queda tratar ahora una nueva aplicación de la IA claramente lesiva para los derechos humanos, invasiva de la intimidad, pero también atentatoria contra otros derechos, como el derecho a una información veraz, a la libre formación de la opinión, a la libertad de expresión, lesiones que derivan de la falsificación de la realidad mediante sistemas de IA, y la creación de versiones adulteradas de la misma, generalmente con fines fraudulentos.

Esta amenaza deriva no de la capacidad de la IA de controlar gran cantidad de identificadores o de calcular a partir de los datos biométricos los estados de ánimo, tendencias o pensamientos de las personas inspeccionadas. Esta nueva amenaza no aparece mencionada en textos internacionales ni en estudios gubernamentales, porque proviene de una capacidad de la IA que no pensamos que pueda lesionar ningún derecho: Se trata de la capacidad y fuerza creativa de los sistemas inteligentes, de la que en principio sólo derivan ventajas y ayuda a los seres humanos.

La IA puede ser creativa en un nivel medio-alto, esto es algo que en la actualidad viene a ser admitido, si acaso se discute si puede llegar a un nivel de genialidad, pero el nivel medio se alcanza en la actualidad sin problemas, como explico en otro trabajo[95]. Pues bien, esta fuerza creativa se puede dirigir a la creación de nuevas obras intelectuales y artísticas, o al descubrimiento de nuevos avances en ciencia, y de hecho así se está haciendo. Pero también se puede dedicar a la creación de réplicas de la realidad, procediendo el sistema como algunos pintores hiperrealistas, que aspiran a una imitación tan fiel de la realidad que no sea posible distinguir sus obras de las fotografías de los objetos pintados en el lienzo.

95 Lacruz Mantecón, Miguel L., *Inteligencia Artificial y Derecho de Autor*, Editorial Reus, Madrid, 2021.

Con una capacidad creativa mejorada, un sistema entrenado en utilizar datos de la realidad, también biométricos, puede combinarlos para emitir versiones nuevas (y fingidas) de la realidad, ya sustituyendo a las personas por otros sujetos, ya creando sujetos enteramente nuevos. De esta manera se puede fingir la existencia de una realidad inventada, en videos o en fotografías, que puede ser admitida como auténtica por los sujetos destinatarios y alcanzar el fin que se persigue, generalmente de engaño, o para modificar su opinión, o simplemente —se dice— para entretenerse o reírse. Estos sistemas, nos dice **Navas Navarro**[96], se ubican dentro de lo que la LIA (que es una Propuesta de Reglamento, no lo olvidemos), tras las enmiendas introducidas en el Parlamento, versión de 14 de junio de 2023, llama «modelos fundacionales generativos», *generative foundational models*, que define en su art. 3.1.c como «… modelos de IA que son entrenados con una amplia cantidad de datos a escala, designados para obtener una generalidad de resultados y que pueden adaptarse a un amplio abanico de tareas diversas». Se caracterizan porque pueden ser utilizados para muy distintas tareas, siendo capaces de generar nuevos resultados a partir de los datos de entrenamiento iniciales sin necesidad de recibir nuevos *inputs*, pues son capaces de recolectar sus propios datos de la realidad o de Internet.

Los ejemplos más claros son los últimos sistemas inteligentes polivalentes com *ChatGpt, Dall-E, Stable Diffusion*, etc., y su uso es sencillo, así el sistema de producción gráfica *Midjourney* convierte las palabras e instrucciones escritas de sus usuarios en imágenes, más o menos realistas, a demanda del cliente. Son muy comunes para los teléfonos móviles las aplicaciones y sitios web para la creación de fotos y videos falsos, que generalmente repiten el esquema de cambiar la cara del utilizador por la de un famoso o a la inversa: *Neocortext, Inc., Reface, FacePlay, FaceJoy, Deep Nostalgia, Doublicat, Familiar, Reflect, Wombo.ia*… Pero también sitios web en la que se puede crear videos de contenido sexual cambiando las caras por las de conocidos o menores, como ocurrió recientemente en nuestro país en Almendralejo, caso en el que unos menores crearon videos porno con la cara de unas niñas compañeras de colegio. Aquí el ataque que se produce entra de lleno en el Código penal. Las falsificaciones llegan también a la voz, con aplicaciones como *Lingojam, Resemble AI, Freemium, ReadSpeaker, Modulate.ai*, que pueden ser utilizadas para timos telefónicos en los que la voz sintetizada reproduce perfectamente a la del sujeto imitado.

Estas creaciones o imitaciones de la realidad, o mejor, de una realidad paralela, puede incidir en la lesión de algunos de los derechos fundamentales, pero en mi opinión el daño es más importante, porque al falsificar la realidad se produce una pérdida de los referentes reales y del marco auténtico de la vida humana. Corremos un serio peligro de perder la certeza acerca de

96 Navas Navarro, Susana (2023), *Chatgpt y modelos fundacionales. Aspectos jurídicos de presente y de futuro*, Reus, Madrid, 2023, p. 18.

nuestros registros de la realidad, con el resultado de producirse una desinformación general y una inverificabilidad de hechos y pruebas. En medios periodísticos e informativos ya no son novedad las intoxicaciones y falsas noticias o *fake news* creadas con IA y expandidas con propósitos malintencionados. En el campo del arte se anuncian producciones de nuevas obras de pintores clásicos, fallecidos hace siglos, cuya obra se crea hoy con su estilo y su arte. En cinematografía se «resucita» a actores fallecidos, y se les hace actuar junto a los vivos, o se rejuvenece a los actores ya ancianos, para hacer efecto de regresión en el tiempo. En la educación, corremos el riesgo de lo que yo llamo «deslocalizar» nuestra inteligencia, al trasladar la producción de resultados inteligentes de nuestro propio cerebro al de la máquina. Ya no es posible distinguir la labor intelectual humana de la de la máquina, produciéndose un general descenso del nivel de inteligencia, que es precisamente lo que nos indica la llamada «regresión del efecto Flynn» de la que hablo en otros trabajos[97].

Sin embargo, la peligrosidad de estos «modelos fundacionales» de propósito general no suele apreciarse en su justa medida, es más, en el texto del Parlamento europeo sobre LIA de 14 de junio 2023 todo lo más somete a estos sistemas inteligentes a los requisitos y obligaciones de los sistemas de alto riesgo, como nos señala **Navas Navarro**[98]. Ello posiblemente deriva de su carácter general y polivalencia, que hace que las utilizaciones más habituales, como instrumentos de trabajo o de entretenimiento desvíen la atención de sus utilizaciones lesivas.

IV. Epílogo: *Doomers* contra *boomers*

La situación actual nos muestra una utilización cada vez mayor de estos sistemas biométricos en los ámbitos de vigilancia y perfilado de personas, sin entrar en el sistema de crédito social chino, vigente en ámbitos por ahora limitados. Y desde luego, una invasión masiva de los sistemas generalistas o modelos fundacionales, capitaneados por el famoso *ChatGpt*. Como decía Luciano **Floridi**[99], la revolución digital inaugura un nuevo capítulo de la historia de la humanidad, añadiendo a continuación «Las futuras generaciones no sabrán lo que era una realidad analógica, predigital y off-line. Somos la última generación en experimentarla». Lo que significa, en definitiva, dos cosas: que somos dinosaurios, y que ya no hay vuelta atrás, que no podemos «desinventar» la IA.

97 Lacruz Mantecón, Miguel L., «La deslocalización de la Inteligencia», *Diario La Ley*, N.º 62, Sección Ciberderecho, 20 de Mayo de 2022, Wolters Kluwer.

98 Navas Navarro (2023), *Chatgpt y modelos fundacionales*…, cit., p.22.

99 Floridi, *The Ethics of Artificial intelligence*, cit., p. xi.

La IA ha venido para quedarse, y las consecuencias de esta invasión empiezan a hacerse sentir en ámbitos como el de la traducción profesional, el diseño gráfico, la música y las listas de reproducción *on line*, los bancos de imágenes, o la publicidad. En todos éstos, se puede decir que se comienza a ver las orejas al lobo. ¿Qué es lo que está pasando? Que por fin está llegando la IA «de verdad», la IA fuerte o de nivel similar al humano. Pero también que el mundo va a ser muy distinto, y que puede parecerse a una distopía de pesadilla.

Sin embargo, ya se aprecia una cierta reacción de cautela por parte de los especialistas y los responsables gubernamentales, así la UNESCO, en un texto de 2023 titulado *Kit de herramientas global sobre IA y el Estado de derecho para el poder judicial*[100], señala directamente la incidencia de la IA en los derechos fundamentales: «…el despliegue de IA podría utilizarse para limitar la libertad de expresión de las personas o su capacidad para participar en actividades políticas o para identificar a los disidentes políticos. La IA también podría dañar los derechos humanos en situaciones en las que se utilizan datos de capacitación de baja calidad, diseño de sistemas o interacciones complejas entre el sistema de IA y su entorno». El resultado de esta utilización de la IA puede ser una potenciación del discurso de odio o la incitación a la violencia, así como la desinformación sobre asuntos políticos y públicos, especialmente en tiempos de elecciones, con la consiguiente manipulación de los resultados.

En la UE tenemos la *Recomendación (UE) 2023/2425 de la Comisión, de 20 de octubre de 2023*, «sobre la coordinación de las respuestas a incidentes, especialmente los derivados de la difusión de contenido ilícito, en anticipación de la plena entrada en aplicación del Reglamento (UE) 2022/2065». Se emite para luchar contra la difusión en línea de contenidos ilícitos y nocivos, o de desinformación e información errónea, en relación con las crisis internacionales provocadas por la agresión de Rusia a Ucrania y el ataque terrorista de Hamás a Israel. Y el reciente proyecto de ley francés *Proposition de loi n°1630 visant à encadrer l'intelligence artificielle par le droit d'auteur,* de 12 de septiembre de este año 2023, nos pone en guardia en su exposición de motivos sobre las fotografías, textos y vídeos creados digitalmente, que, aparte de problemas de Derecho de autor, nos dice que plantea cuestiones de ética, entorpece nuestro libre albedrío e incluso causa problemas para la supervivencia de nuestra creatividad humana.

El 30 de octubre el presidente Biden dicta una *Orden ejecutiva sobre el desarrollo y uso seguro y confiable de la inteligencia artificial*[101]*,* que tiene

100 Publicado por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura © UNESCO 2023, acceso abierto bajo la licencia Attribution-ShareAlike 3.0 IGO (CC-by-sa 3.0 IGO), https://unesdoc.unesco.org/ark:/48223/pf0000387331_spa

101 https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

como objeto promover un uso responsable de la IA. Señala que la IA tiene un gran potencial para hacer que nuestro mundo sea más próspero, productivo, innovador y seguro, pero también puede «…exacerbar daños sociales como el fraude, la discriminación, los prejuicios y la desinformación, así como sustituir y quitar poder a los trabajadores, reprimir la competencia y plantear riesgos para la seguridad nacional. Aprovechar la IA para siempre y aprovechar sus innumerables beneficios requiere mitigar sus riesgos sustanciales. Este esfuerzo exige un esfuerzo de toda la sociedad que incluya al gobierno, el sector privado, la academia y la sociedad civil». Para evitar estos males se imponen nuevos estándares oficiales de seguridad, requiriendo que las empresas que desarrollen sistemas que represente un riesgo grave para la seguridad nacional, la economía o la salud y seguridad públicas notifiquen al gobierno federal cuando entrenen al sistema, y compartan con éste los resultados de todas las pruebas de seguridad. Asimismo, se involucra a las agencias federales de supervisión de la investigación en la preservación de la intimidad de los ciudadanos y vigilancia de los datos que se manejan para el entrenamiento de los sistemas.

Luego los días 1 y 2 de noviembre de 2023 se reúnen en Inglaterra, en Bletchley Park precisamente[102], representantes de los países más avanzados y de instituciones académicas o de investigación (la Universidad de Oxford, el Instituto Alan Turing) en una «Cumbre sobre seguridad de la IA» *(AI Safety Summit)*. En la Declaración[103] que se suscribe, se alerta sobre los importantes riesgos que conlleva la IA, ante los que se hace necesario abordar la protección de los derechos humanos, así como «…la transparencia y la explicabilidad, la equidad, la rendición de cuentas, la regulación, la seguridad, la supervisión humana adecuada, la ética, la evitación de sesgos, la privacidad y la protección de datos. También observamos el potencial de riesgos imprevistos derivados de la capacidad de manipular contenido o generar contenido engañoso». Se consideran especialmente peligrosos los sistemas inteligentes de propósito general, o modelos fundacionales, que podrían ser objeto de usos indebidos, ya intencionalmente, ya por pérdida del control del sistema: «Estos problemas se deben en parte a que esas capacidades no se comprenden completamente y, por lo tanto, son difíciles de predecir».

Como vemos, empiezan las reacciones a la intromisión de sistemas inteligentes en la vida de los ciudadanos. Otro síntoma de esta respuesta está en la dimisión del director ejecutivo de la compañía *OpenAI*, creadora de *ChatGpt*, y su fulgurante reposición en el puesto pocos días después, suceso del

102 Bletchley Park, a medio camino entre Oxford y Cambridge, fue el lugar donde se instaló en la II Guerra mundial el centro de codificación y cifrado en el que Alan Turing descifró el código de la máquina *Enigma*.

103 https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023

que han dado cuenta todos los periódicos. Nos cuenta así Pérez Giménez[104], recogiendo los hechos de *The Economist*, que el 17 de noviembre, la junta directiva de *OpenAi* formada por seis miembros expulsó a Sam Altman, creador de *ChatGpt*, y con él, renunciaba también el presidente de *OpenAi*, Greg Brockman, yéndose a trabajar a *Microsoft*. Pero tres días después «…casi la totalidad de los 770 empleados de la startup firmaban una carta en la que exigían el retorno de Altman o todos se irían con él a trabajar a *Microsoft*. Al día siguiente, 21 de noviembre, Altman y Brockman regresaban a *OpenAi*».

Esta secuencia es una muestra de la lucha entre los «pragmáticos o aceleracionistas» de la IA, también llamados *boomers*, frente a los «idealistas académicos o fatalistas», *doomers*. Microsoft, y ahora también *OpenAi*, son un terreno tomado por los *boomers*, y el desarrollo de los sistemas generalistas o modelos fundacionales, en particular *ChatGpt-5*, sigue adelante sustituyendo en cada vez más tareas inteligentes a los humanos. Por su parte, los *doomers* o catastrofistas se hicieron visibles desde que el *Future of life Institute* promovió la petición de una moratoria de al menos 6 meses en el desarrollo de la IA, mediante una carta firmada entre otros por Elon Musk, Stuart Russell o Yuval Noah Harari. En ella se pregunta si debemos arriesgarnos a perder el control de nuestra civilización, afirmando que «Estas decisiones no deben delegarse en líderes tecnológicos no electos. Sólo se deberían desarrollar potentes sistemas de IA una vez que estemos seguros de que sus efectos serán positivos y sus riesgos manejables».

Lo que late en esta discusión entre aceleracionistas y catastrofistas es el miedo a la IA superhumana, la «Superinteligencia», escenario planteado por Nick Bostrom[105] que plantea la posibilidad de que, mediante autoaprendizaje, un sistema inteligente supere a la inteligencia humana y reemplace al ser humano como gestor de la vida social y económica. La posibilidad de una Superinteligencia aparece en autores como Ray Kurzweil o Stuart Russell, pero, en general, la doctrina la rechaza, así el científico-jefe de Meta, Yann LeCun, o Andrew Ng, investigador de Meta, para quienes es una hipótesis sin sentido. Desde luego, la visión apocalíptica de una Superinteligencia que domine y dirija los destinos de la especie humana, es algo difícil de aceptar. Pero la idea de unos sistemas de IA general que gestionen los asuntos humanos, de acuerdo con las directrices que los humanos decidamos, eso es un escenario que parece cada día más cercano. Y este uso general de la IA va a suponer, como he expuesto, un cambio social de primer orden. La incertidumbre es completa.

104 Pérez Giménez, Alberto «Inteligencia Artificial: 5 días de guerra en Openai que cambian el mundo y "dan miedo"», *Vozpopuli*, 25/11/2023 04:4, https://www.vozpopuli.com/opinion/inteligencia-artificial-openai.html

105 Bostrom, Nick, *Superinteligencia*, TEELL, Zaragoza, 2016, pág. 19.

Somos la última generación en valorar la inteligencia humana como gestora de los asuntos humanos, porque pensar es agotador, y ya tenemos máquinas que piensan por nosotros. La sociedad que se avecina pondrá la esencialidad de lo humano en otros valores, y desde luego, no cultivará la inteligencia como don preciado y del que se obtienen ventajas evolutivas. Vamos hacia una sociedad que será gestionada por sistemas inteligentes, en la que los resultados inteligentes se obtienen de máquinas, dirigiéndose los esfuerzos humanos a otros fines: No va a ser una sociedad particularmente inteligente.

Frente a este panorama, ¿qué puede hacer la sociedad, qué el Derecho? Estimo que la respuesta debe venir de tres ámbitos. Primero, y, ante todo, una respuesta técnica: todo contenido generado por una IA debe delatar criptográficamente su origen, debe ser verificable, para evitar toda confusión entre resultados inteligentes humanos y los que vienen de una máquina. Segundo, una respuesta educativa: en todos los niveles educativos se debe dejar bien claro que la utilización de la IA desplaza a la propia inteligencia, y no produce resultados evaluables. El uso de la IA en educación pocas veces tiene una utilización meramente instrumental, suele serlo sustitutiva. Tercero, una respuesta jurídica: Regulación y medidas de seguridad. Se debe asegurar la indicada verificabilidad de las producciones sintéticas, y aumentar la protección de las personas contra la intoxicación producida por IA. La vía de los neuroderechos parece iniciada, pero se trata de un concepto vago y quizá sólo sea necesaria una interpretación correcta de los derechos ya existentes.

Y, para terminar, una reflexión: La mayoría de los problemas que nos causa la IA no provienen de su mal funcionamiento, sino de la perfección no humana de su funcionamiento. Que se suma a un error humano bastante común: tendemos a pensar que un sistema inteligente, como produce resultados parecidos a los de la inteligencia humana, tiene algo de humano. Y eso sí que es un gran error, la inteligencia artificial no es en absoluto humana, para empezar, es incansable. Este es el error que comete Mickey Mouse (en *Fantasía*, Walt Disney, 1940) cuando hechiza a la escoba para que haga su trabajo llenando el depósito de agua: piensa que, como haría él mismo, cuando el depósito estuviera lleno, la escoba, cansada, dejaría de trabajar.

Bibliografía

Bauman, Zygmunt, *Modernidad líquida*, Fondo de Cultura Económica, Buenos Aires, 2004.

Cabrales Goitia, Antonio y **Rey Biel, Pedro**, (2021). «Mas allá de los nudges: Políticas públicas efectivas basadas en la evidencia de las ciencias del comportamiento», *Gestión y Análisis de Políticas Públicas*, (25), pp. 38-45. DOI: https://doi.org/10.24965/gapp.i25.10864.

Chen Zijiao, **Qing Jiaxin** y **Zhouy**, **Juan Helen**, (2023). «Cinematic Mindscapes: High-quality Video Recons-truction from Brain Activity», preprint en *arXiv*:2305.11675, https://doi.org/10.48550/arXiv.2305.11675

Coca Payeras, **Miguel**, (2023), «Las iniciativas de la Unión europea sobre inteligencia artificial: de la persona electrónica, al difícil equilibrio entre la necesidad de impulsarla y evitar sus riesgos», *Revista de Derecho Civil,* vol. X, núm. 2, especial (junio, 2023). En http://nreg.es/ojs/index.php/RDC

Costas Pérez, **Elena**, y **Tucat**, **Pablo**, (2021). «Nudges: diseño y evaluación», *Gestión y Análisis de Políticas Públicas*, (25), 8-22. DOI: https://doi.org/10.24965/gapp.i25.10868

Cotino Hueso, **Lorenzo**, (2023). «Reconocimiento facial automatizado y Sistemas de identificación biométrica bajo la regulación superpuesta de Inteligencia artificial y protección de datos», *Derecho Público de la inteligencia artificial,* Coordinadores **Balaguer Callejón**, Francisco, **Cotino Hueso**, Lorenzo, Fundación Manuel Giménez Abad, Zaragoza, 2023.

Cotino Hueso, **Lorenzo**, (2022). «Sistemas de inteligencia artificial con reconocimiento facial y datos biométricos. Mejor regular bien que prohibir mal», *El Cronista del Estado Social y Democrático de Derecho*, N.° 100 (Septiembre-Octubre), 2022 (Ejemplar dedicado a: Inteligencia artificial y derecho).

De Asís Pulido, **Miguel**, (2022). «La justicia predictiva: tres posibles usos en la práctica jurídica», en *Inteligencia Artificial y Filosofía del Derecho*, pp. 285-308, Ediciones Laborum, Murcia, 2022.

De Asís Roig, **Rafael Francisco**, (2022). «Ética, Tecnología y Derechos», en *Inteligencia Artificial y Filosofía del Derecho*, pp. 25-40, Ediciones Laborum, Murcia, 2022.

De Miguel Asensio, **Pedro A.**, (2021). «Propuesta de Reglamento sobre Inteligencia artificial», *La Ley Unión Europea*, n.° 92, mayo 2021.

Drummond, **Víctor**, (2004). *Internet, privacidad y datos personales*, Traducción y notas de Isabel **Espín Alba**, Reus, Madrid, 2004.

Ebers, **Martin**, (2023). «El futuro marco jurídico europeo de la inteligencia artificial», *Revista General de Legislación y Jurisprudencia*, 2023, número 2, pp. 185-21.

Fernando Pablo, **Marcos M.**, (2023). «Capítulo I: Construyendo la dignidad digital de la persona en el entorno digital. De los datos de tráfico, a

la plaza y mercado de los servicios de comunicaciones electrónicas», en *Desafíos éticos, jurídicos y tecnológicos del avance digital* / coord. por Alicia **Rodríguez Sánchez**, Pilar **Talavera Cordero**; José Luis **Domínguez Álvarez** (dir.), Daniel Terrón Santos (dir.) Iustel, Madrid, 2023, págs. 21-40.

Floridi, **Luciano**, *The Ethics of Artificial intelligence*, Oxford University Press, Oxford, 2023.

García Inda, **Andrés**, **González Ordovás**, **María José** y **Muñoz Soro**, **José Félix**, «IA y Filosofía del Derecho: derechos, normas y sesgos», en Proyecto IASAC Unizar, Ciencias sociales y jurídicas, http://unidigitaliasac.unizar.es/ficha/la-ai-vista-desde-la-filosofia-del-derecho

Garriga Domínguez, **Ana**, (2022) «Inteligencia artificial y el fenómeno de la desinformación: el papel del RGPD1 y las garantías recogidas en la propuesta de la ley de servicios digitales», en *Inteligencia Artificial y Filosofía del Derecho*, pp. 451-473. Ediciones Laborum, Murcia, 2022.

González Tapia, **M.ª Isabel**, (2022). «Protección penal de los neuroderechos: el uso directo de las neurotecnologías sobre el ser humano», *Inteligencia Artificial y Filosofía del Derecho*, pp. 313-335 Ediciones Laborum, Murcia, 2022.

Hu, **Margaret**, (2022). «Biometrics and an AI Bill of Rights», 60 *Duquesne Law Review* pp. 283-301 (2022). En *William & Mary Law School Scholarship Repository*, Faculty Publications, Summer 2022. https://scholarship.law.wm.edu/facpubs/2078

Juliá-Pijoan, **Miquel**, (2023). «Una aproximación al perfilaje criminal desde la investigación neurocientífica», *FODERTICS 11.0 Derecho, entornos virtuales y tecnologías emergentes,* pp. 441-453, Comares, Granada, 2023.

Lacruz Mantecón, **Miguel L.**, *Inteligencia Artificial y Derecho de Autor*, Editorial Reus, Madrid, 2021.

— «La deslocalización de la Inteligencia», *Diario La Ley*, N.º 62, Sección Ciberderecho, 20 de Mayo de 2022, Wolters Kluwer.

Ligthart, **Sjors**, **Douglas**, **Thomas**, **Bublitz**, **Christoph**, **Kooijmans**, **Tijs** y **Meynen**, **Gerben** (2021), «Forensic Brain-Reading and Mental Privacy in European Human Rights Law: Foundations and Challenges», *Neuroethics*, 2021, 14, pp. 191-203, https://doi.org/10.1007/s12152-020-09438-4

Lucasiewicz, **Rafal**, (2022). «Facial recognition. Matching in gamete donation using AI», *Tratado de Inteligencia artificial y Derecho en el nuevo milenio*, pp. 385-401. Ediciones Olejnik, Santiago de Chile, 2022.

MATEFI, Roxana, y DARIUS, Cupu, (2022). «Artificial intelligence and its impact on personality rights», en *Tratado de Inteligencia artificial y Derecho en el nuevo milenio*, pp. 70-89, Olejnik, Santiago de Chile, 2022.

MEGÍAS QUIRÓS, José Justo, (2022). «Derechos humanos e Inteligencia artificial», *Revista DIKAIOSYNE*, N.° 37, pp. 140-163, número especial sobre DDHH, en coedición con el Observatorio de Derechos Humanos de la Universidad de Los Andes. Mérida - Venezuela. Enero, 2022.

MOREU CARBONELL, Elisa, (2022). «La regulación de los neuroderechos», *Revista General de Legislación y Jurisprudencia*, 2022, n.° 1, enero-marzo, pp. 71-100.

NAVAS NAVARRO, Susana (2023), *Chatgpt y modelos fundacionales. Aspectos jurídicos de presente y de futuro*, Reus, Madrid, 2023.

ORTIZ DE ZÁRATE-ALCARAZO, Lucía. (2023). «Las Políticas conductuales inteligentes: Oportunidades y riesgos ético-políticos de la Inteligencia Artificial para el nudging», *Revista Española de Ciencia Política*, n.° 62, julio 2023, pp. 67-93. Doi: https://doi.org/10.21308/recp.62.03.

PACKARD, Vance, (1992). *Las formas ocultas de la propaganda*, Editorial Sudamericana, Buenos Aires, Decimoctava Edición, Junio de 1992.

POLLICINO, Oreste y DE GREGORIO, Giovanni, (2021). «Constitutional Law in the Algorithmic Society», *Constitutional Challenges in the Algorithmic Society*, Cambridge University Press, 2021, pp. 3-24. DOI: https://doi.org/10.1017/9781108914857.002, p. 4.

POLLICINO, Oreste y PAOLUCCI, Federica, (2022). «Digital constitutionalism to the test of the smart identity», *Journal of E-Learning and Knowledge Society,* Vol. 18, No. 3 (2022), pp. 8-21. DOI: https://doi.org/10.20368/1971-8829/113581

RETTINGER, Lizzy, «The Human Rights Implications of China's Social Credit System», *Journal of High Technology Law*, vol. XXI, n.° 1, (2021), pp. 1-33.

ROBERTS, Huw, COWLS, Josh, MORLEY, Jessica, TADDEO, Mariarosaria, WANG, Vincent, FLORIDI, Luciano, (2021). «The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation». *AI & Soc.* 36, 59-77 (2021). https://doi.org/10.1007/s00146-020-00992-2

SALVI, Nicolás y NIGRI, Santiago, (2022). «Minority Report: the Road to a Deterministic Theory for the Philosophy of Criminal Law», June 28th, 2022, *Opinión Jurídica*, 21(46), Special Edition 2022, pp. 1-18, https://doi.org/10.22395/ojum.v21n46a2

Solar Cayón, José Ignacio, (2022). «Inteligencia artificial y justicia digital», en *Inteligencia Artificial y Filosofía del Derecho*, pp. 381-425, Ediciones Laborum, Murcia, 2022.

Turley, Jonathan, (2020). «Anonymity, Obscurity and Technology: Reconsidering Privacy in the Age of Biometrics», *Boston University Law Review*, Vol. 100: 2020, pp. 2179-2261. En https://www.bu.edu/bulawreview/files/2021/01/TURLEY.pdf

Vicente Domingo, Elena y **Rodríguez Cachón, Teresa**, (2023). «Derecho de la Persona, Neurodatos y Neuroderechos: *A Research Agenda*», *Revista General de Legislación y Jurisprudencia*, 2023, número 3, pp. 495-526. P. 499.

Yamazaki, Yukari, (2020). «An Empirical Study for The Acceptance of Original Nudges and Hypernudges», *Societal Challenges in the Smart Society* / coord. por Mario **Arias Oliva**, Jorge **Pelegrín Borondo**, Kiyoshi **Murata**, Ana María **Lara Palma**, 2020, ISBN 978-84-09-20273-7, págs. 323-336. En *Dialnet*: https://dialnet.unirioja.es/servlet/articulo?codigo=7867256