

La inteligencia artificial y los nuevos entes cibernéticos¹

MIGUEL L. LACRUZ MANTECÓN*

Sumario: 1. ¿Es la Inteligencia Artificial verdadera inteligencia?; 2. ¿Hay que reconocer personalidad en los entes electrónicos inteligentes?; 3. La alternativa entre personas y cosas; 4. Una cierta subjetividad: el robot como “agente”; 5. Los dos órdenes de problemas de los sistemas inteligentes; 6. La atribución de los resultados del sistema inteligente; 6.1 La responsabilidad civil por daños; 6.2 Los beneficios de la actividad robótica; 7. El segundo orden de problemas: la consideración que otorgamos los humanos a los entes cibernéticos, y su efecto en nosotros; 8. Concluyendo: seamos prácticos; Bibliografía.

1. ¿Es la Inteligencia Artificial verdadera inteligencia?

En la actualidad, la Inteligencia Artificial (IA) es una tecnología de la que los seres humanos hacemos un uso intensivo. Hablan HERRERA, CORDÓN y DEL

JURISMAT, Portimão, n.º 17, 2023, pp. 167-195.

* Prof. Tit. Derecho Civil, Universidad de Zaragoza.

¹ El presente trabajo se ha realizado al amparo del Proyecto «Derecho e inteligencia artificial: nuevos horizontes jurídicos de la personalidad y la responsabilidad robóticas», IP. Margarita Castilla Barea, (PID2019-108669RB-100 / AEI / 10.13039 / 501100011033).

JESÚS² de una *Inteligencia artificial silenciosa*, que nos acompaña hoy en nuestras actividades diarias, pero por mi parte estimo que lo hace desde hace más de cuarenta años, desde que en la década de los 70 del pasado siglo se popularizaron las calculadoras electrónicas y en la de los 80 los ordenadores personales. Desde entonces, como tengo dicho en otro lugar,³ hemos subcontratado parte de nuestras tareas inteligentes (memorizar y recordar datos, cálculo mental, lectura de mapas) y las hemos trasladado a estas máquinas, para que sean nuestros nuevos esclavos electrónicos quienes realicen las tareas que a nosotros nos resultan pesadas.

Por *inteligencia* entendemos una cualidad de los seres humanos, y de algunos animales, de origen adaptativo, que se reconoce por sus resultados, y que se manifiesta a través de conductas de resolución de problemas de cualquier tipo. Para uno de los pioneros de la IA, Marvin MINSKY,⁴ la inteligencia consiste en esta capacidad de resolución de problemas, y la inteligencia artificial sería esa misma capacidad cuando se encuentra en entes no biológicos, los ordenadores o computadores. En similar sentido, BRYSON y WINFIELD⁵ dan una definición más precisa, la Inteligencia sería la capacidad de hacer lo adecuado en el momento adecuado, en un contexto en el que no hacer nada (o no cambiar de conducta) sería peor, y la IA es esta misma capacidad demostrada por la conducta de artefactos no biológicos. Resumiendo, inteligencia es la capacidad de resolver problemas inespecíficos, e IA es esa misma capacidad demostrada por máquinas. Tanto una como otra se reconocen por sus efectos: Deducimos la existencia de inteligencia a partir de las conductas inteligentes.

En cuanto a si es posible que una máquina desarrolle inteligencia, esto se logra haciendo que el sistema desarrolle comportamientos inteligentes mediante el manejo de símbolos físicos (números, letras). Como nos recuerdan los profesores WALLACH y ALLEN,⁶ fueron el informático y psicólogo Allen Newell y el economista y politólogo Herbert Simon quienes establecieron la siguiente premisa: «Un sistema de símbolos físicos tiene los medios necesarios

² HERRERA, Francisco, CORDÓN, Óscar, DEL JESÚS, M^a José, «Una visión actual de la Inteligencia artificial», en *El Derecho y la Inteligencia artificial*, F. Herrera Triguero, A. Peralta Gutiérrez, L. Torres López, coords., Universidad de Granada, Granada 2022, pág. 51.

³ LACRUZ MANTECÓN Miguel L., «Inteligencia Artificial: Daño y Prejuicio», *JURISMAT*, Portimão, 2022, n.º 15, pp. 389-414.

⁴ MINSKY, Marvin, *La sociedad de la mente*, Ediciones Galápagos, Buenos Aires, 1986, pág. 74.

⁵ BRYSON, J. J., y WINFIELD, A., «Standardizing Ethical Design for Artificial Intelligence and Autonomous Systems», *Computer*, vol. 50, n.º 5, <https://doi.org/10.1109/MC.2017.154>.

⁶ WALLACH, Wendell y ALLEN, Colin, *Moral Machines. Teaching robots right from wrong*, Oxford University Press, Nueva York, 2009, pág. 56.

y suficientes para llevar a cabo acciones inteligentes». Es decir, que la manipulación de símbolos puede generar repuestas físicas o verbales que se pueden calificar de «inteligentes» porque resultan adecuadas a los requerimientos del entorno, o resuelven problemas, que es decir lo mismo, tanto procedan dichas conductas de un ser humano como de una entidad no humana.

Imagine una pizarra magnética en la que unas manos mueven cifras, realizando, por ejemplo, una simple suma de $7 + 3 = 10$. La primera impresión que tenemos es la de que se ha producido una actividad intelectual, y que para realizarla alguien ha tenido que *pensar*. Pero si consideramos objetivamente esta actividad, únicamente se ha producido una manipulación de unas fichas que contenían unos símbolos, nada más. En cambio, si la misma operación de sumar $7 + 3$ la realiza una calculadora mecánica o una electrónica, ya no nos parece lo mismo, al no haber un ser humano detrás: Esta actividad la consideramos de carácter inteligente si la realiza un ser humano, pero no le damos la misma importancia si la realiza una calculadora. Esto es claramente un prejuicio, la conducta que se ha observado es en ambos casos la misma, una conducta de manipulación de símbolos que da respuesta a un problema matemático. La diferencia es que el gestor de dicha conducta es, en un caso, de carácter biológico, y en el otro, no biológico, pero en ambos casos es de naturaleza material.

Esta argumentación de la equivalencia entre los comportamientos inteligentes, humanos o de la máquina, es la que constituye la base del famoso «Juego de la imitación» o Test de Turing, que se ha consolidado como el argumento principal para los partidarios de la posibilidad de emulación de la inteligencia humana. Fue Alan TURING,⁷ en un famoso artículo publicado en 1950, quien propuso este «Juego de la imitación», en el que un interrogador plantea preguntas a dos personas, hombre y mujer, A y B, y tiene que averiguar, a partir de sus respuestas por escrito, cuál de ellas es la mujer y cuál el hombre. El siguiente paso es sustituir a una de las personas por un ordenador. Ahora el juego consiste en averiguar cuál de los dos interrogados es el humano y cuál la máquina: Si no podemos averiguar cuál de los dos interlocutores es el humano, entonces no podremos negar inteligencia a ninguno de los dos, y tendemos que concluir que la máquina también «piensa», que tiene inteligencia. En realidad, el test de Turing descansa sobre la base de entender que sólo cuando se alcanza la inteligencia humana se demuestra verdadera inteligencia.

⁷ TURING, Alan Mathison, «Computing Machinery and Intelligence», *Mind*, 49 (1950), pp. 433-460. Sin embargo, la idea fundamental de la Inteligencia artificial ya la expuso en anteriores estudios, singularmente en *Intelligent Machinery: A Heretical Theory*, informe para el National Physical Laboratory de 1948.

Frente a esta idea, muchos autores, del ámbito de la Filosofía o del Derecho, niegan la equivalencia de la IA con la humana. Los argumentos que se mueven para sustentar esta posición son de dos tipos.

En primer lugar está el argumento de la complejidad, que parte de la excelente opinión que los seres humanos tenemos sobre nosotros mismos, y que lleva a que la actividad intelectual humana se considere de una altura o calidad especial, que sólo entes dotados de un atributo que llamamos «espiritualidad» pueden alcanzar. El ámbito en el que la exigencia de espiritualidad aparece como un requerimiento absoluto, por la complejidad de la tarea, es el de la creación literaria, artística y científica, campo en el que se dice que la Inteligencia artificial no puede competir con la inspiración del artista o del poeta. De hecho, nuestra Ley de Propiedad Intelectual, como muchas otras, exige que el autor de una obra sea una persona física. Se comparan las grandes obras artísticas con las cumbres de montañas a las que ha llegado el espíritu humano, pero que son inaccesibles para la inteligencia artificial.

Sin embargo, Max TEGMARK⁸ recoge precisamente esta analogía de las cumbres de las montañas, y nos advierte que se trata de un ejemplo que ya fue expuesto por Hans Moravec: Se considera que los distintos logros de la humanidad son como distintas alturas en un valle montañoso, representando las más bajas las tareas mecánicas, y las cimas más altas los logros más refinados, entre ellos los de carácter científico y artístico. Pues bien, el desarrollo de la IA se puede describir como una inundación que crece en dicho valle, y va sumergiendo primero los puntos más bajos: primero consigue realizar operaciones matemáticas, o recuperar datos de una base de datos. Luego alcanza otras tareas más complejas, como jugar al ajedrez o al go, o traducir de un idioma a otro el lenguaje escrito y el hablado, hitos que ya han sido alcanzados por la IA. La creación artística, que tradicionalmente se consideraba como actividad “espiritual”, reservada a la mente humana y fruto más depurado de la inteligencia y sobre todo de la sensibilidad humanas, es una altura que ya está siendo sumergida por el avance de las aguas, pues ya existen sistemas inteligentes que crean música, obra gráfica o incluso escriben novelas. En definitiva, señala TEGMARK, la irrupción de la IA en este campo de la creación artística significa que cada vez quedan menos reductos reservados en exclusiva a la inteligencia humana. Si hay una ley de hierro en materia de IA, es que cuando se le pone una frontera del tipo: “Un ordenador nunca podrá derrotar a un gran maestro del ajedrez”, o “un ordenador nunca podrá crear música”, dicho límite es superado por la IA en no mucho tiempo.

⁸ TEGMARK, Max, *Vida 3.0. Que significa ser humano en la era de la inteligencia artificial*, Taurus, Barcelona, 2018, pág. 104.

El segundo tipo de argumentos en contra de la admisión de una igualdad entre la inteligencia humana y la artificial está en la inexistencia en la máquina inteligente de una característica que hemos ligado siempre a la especie y a la mente humanas, y por lo tanto a su inteligencia: el tener una conciencia, un *yo*, o *yo mismo*, eso que llamamos espíritu o «alma», que todos y cada uno estamos segurísimos de tener, sin que ninguno podamos concretar en qué consiste. Y esta falta de conciencia, de sentido del “yo” determina, en consecuencia, la falta de finalidad, de voluntad o de auténtica comprensión del sentido que tienen las actividades de un sistema inteligente. Este argumento parece tener especial peso, pues sobre su base los humanos aplicamos la siguiente lógica: Si la máquina carece de conciencia, si no tiene alma, sus actuaciones, aunque parezcan presuntamente inteligentes, no generan una experiencia consciente, con sentido y significado para el ente que la realiza. Luego las máquinas, como no saben lo que hacen, no tienen auténtica inteligencia, cualidad que sigue reservada a los seres humanos.

Esto es lo que quiere demostrar el famoso experimento de la habitación china, ideado por el filósofo John SEARLE:⁹ «Imaginen que llevan a cabo los pasos señalados por un programa para responder preguntas en una lengua que no comprenden. Yo no entiendo chino, así que me imagino encerrado en una habitación con un montón de cajas de símbolos chinos (la base de datos), recibo pequeños manojos de símbolos chinos (preguntas en chino) y miro el libro de reglas (el programa) [...] Ejecuto ciertas operaciones en los símbolos de acuerdo con esas reglas (es decir, llevo a cabo los pasos del programa) y devuelvo pequeños manojos de símbolos (respuestas a las preguntas) a los que están fuera de la habitación». El operador humano en este caso se comporta exactamente como lo haría un ordenador: parece que entiende lo que hace, pero en realidad no sabe una palabra de chino. De acuerdo con esto, el ordenador, lo mismo que sucede en la habitación china, realiza una mera manipulación de símbolos sin el sentido que los humanos ligamos al significado que para nosotros tienen dichos símbolos. Es decir, que no «sabe lo que hace», su comportamiento verbal para él no tiene un significado, aunque para los de fuera de la habitación parezca tenerlo.

Frente a este modo de ver las cosas, para otros autores, la conciencia o el “yo” no tienen tanta importancia. Autores cognitivistas como Dennet y Hofstadter, y pioneros en inteligencia artificial como McCarthy y Minsky consideraban que la conciencia era un resultado de una actividad cerebral suficientemente compleja, por lo que la reproducción del funcionamiento del cerebro humano por un sistema inteligente podría alcanzar un estado de conciencia. Es lo que

⁹ SEARLE, John, *El misterio de la conciencia*, Paidós, Barcelona, 2000, págs. 24 y ss.

afirma DENNET,¹⁰ que partiendo de una concepción simple de la conciencia como conocimiento del «yo», estima que sí es posible programar la autoconciencia en una computadora (o replicar una conciencia en la máquina). La autoconciencia en su noción más elemental consistiría simplemente en distinguir entre el sujeto y el resto del mundo, y esto se puede reproducir en un ordenador, que dispone de mecanismos de autovigilancia y autocontrol.

Esta consideración de la conciencia o del “yo” como un mero estado de autorreferencia es también adoptada por pensadores como HARARI,¹¹ quien, partiendo de una concepción humeana o cognitivista, cuestiona la existencia de un auténtico «yo» como un rector imparcial y consciente de nuestra propia vida: «...el yo también es un relato imaginario, al igual que las naciones, los dioses y el dinero. Cada uno de nosotros tiene un sofisticado sistema que se deshace de la mayoría de nuestras experiencias, conserva solo unas pocas muestras bien escogidas, las mezcla con fragmentos de películas que hemos visto, novelas que hemos leído, discursos que hemos oído, y ensoñaciones propias, y con todo este revoltijo teje un relato en apariencia coherente sobre quién soy yo, de dónde vengo y adónde voy». Estima por tanto que la idea de que los humanos siempre tendrán, gracias a la conciencia, una capacidad única, fuera del alcance de los algoritmos, es una ilusión.

En resumidas cuentas, muchos autores consideran que la conciencia humana no tiene la importancia que le damos, y o bien reconocen la aptitud de la IA para emular la conciencia humana, o bien estiman que esta cualidad no le es imprescindible a la máquina para llevar a cabo una actividad inteligente. Esto lo podemos ver con un ejemplo: el comando «Ordenar», marcado como AZ y ubicado en la pestaña “Inicio” del programa Microsoft Word. La tarea de ordenar alfabéticamente las obras por autores es una tarea que, sin duda, requiere inteligencia. Y aunque mi pobre ordenador no tiene conciencia, ni esa ordenación es para él una experiencia consciente (o por lo menos, no me lo ha dado a entender), lleva a cabo dicha ordenación alfabética a la perfección.

2. ¿Hay que reconocer personalidad en los entes electrónicos inteligentes?

Dejando la cuestión de la conciencia aparte, se considera que en un plazo no muy largo la actividad de los sistemas inteligentes podrá ser indistinguible de la de los seres humanos, siendo previsible una IA que sea igual que la humana,

¹⁰ DENNETT, Daniel C., «Can machines think?», In M. G. Shafto (ed.), *How We Know*. Harper & Row (1984), pag. 143.

¹¹ HARARI, Yuval Noah, *Homo Deus. Breve historia del porvenir*, Debate, Barcelona, 2016, págs. 296 y 311.

o que por lo menos finja muy bien serlo. En base a ello, están quienes argumentan que la concesión de personalidad y derechos a los robots es un imperativo moral, derivado de la existencia en ellos de esta inteligencia, inteligencia que, superado el Test de Turing, sería semejante, y para algunos igual, que la humana.

La primera opción, por tanto, es crear un *status* jurídico específico para estos seres inteligentes, dotándoles de una personalidad jurídica especial, como base para una atribución de derechos. Fue la famosa *Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica*,¹² la que introdujo el tema en el punto 59.f: *f) crear a largo plazo una personalidad jurídica específica para los robots, de forma que como mínimo los robots autónomos más complejos puedan ser considerados personas electrónicas responsables de reparar los daños que puedan causar...* Estas afirmaciones fueron tomadas como basamento para la construcción de nuevas “personas robóticas”, olvidando que la Resolución tan solo presentaba propuestas para resolver el problema de la responsabilidad por el daño causado por sistemas inteligentes.

Pero esta idea es abandonada en textos europeos posteriores, que la rechazan explícitamente. Así, el inmediatamente posterior *Dictamen del Comité Económico y Social Europeo sobre la «Inteligencia artificial: las consecuencias de la inteligencia artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad»*, de 31 de mayo y 1 de junio de 2017, (2017/C 288/01), aunque menciona esta posible personalidad electrónica o *e-personality*, advierte categóricamente que *El CESE se opone a cualquier tipo de estatuto jurídico para los robots o sistemas de IA por el riesgo moral inaceptable que ello conlleva. La legislación en materia de responsabilidad tiene un efecto correctivo y preventivo que podría desaparecer en cuanto el riesgo de responsabilidad civil dejase de recaer sobre el autor por haberse transferido al robot (o sistema de IA)* –parágrafo 1:12-. Insiste en esta línea el Informe *Responsabilidad para la IA y otras tecnologías digitales emergentes (Liability for artificial intelligence and other emerging digital technologies)*, de 27 de noviembre 2019, del Grupo de Expertos sobre Responsabilidad y Nuevas Tecnologías, que señala expresamente: «No es necesario dar a los dispositivos o sistemas autónomos una personalidad jurídica». Con mayor radicalidad, la *Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial*, deja sentado que la gestión de la responsabilidad por daños cibernéticos no exige la

¹² Diario Oficial de la Unión Europea, C 252 de 18.7.2018, p. 239/257, en: <https://eur-lex.europa.eu/legal-content/ES/TXT/?qid=1552383660116&uri=CELEX:52017IP0051>.

conformación de una personalidad jurídica en las máquinas. Y, por último, la *Propuesta de Reglamento por el que se establecen Normas Armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial)*, de 21 de abril de 2021, nos presenta un Proyecto de Reglamento europeo sobre la IA, regulando en 85 artículos los diversos aspectos de esta tecnología, y en él la idea de una personalidad electrónica está ya olvidada.

Además, la creación de una personalidad electrónica para robots, como medio de gestionar la responsabilidad civil en que pudieran incurrir estas máquinas, es una vía superada en la doctrina. En este sentido consideró Martin EBERS¹³ que la responsabilidad debe ser siempre del ser humano y no del robot personificado, al entender que los robots «...no pueden modificar su arquitectura de control, la cual garantiza que su actuar permanece dentro de guardarraíles predefinidos. Por tanto, no se debería en absoluto hablar de inteligencia y, sobre todo, de acciones autónomas. El coche que se conduce por sí mismo no decide ni si se pone en marcha, ni sobre el lugar adonde va, sino que depende siempre de las especificaciones de su usuario». Y en un trabajo posterior, precisa EBERS¹⁴ que la creación de una personalidad jurídica electrónica tampoco resulta económicamente eficiente, ya que el mismo propósito se puede lograr más fácilmente introduciendo una responsabilidad objetiva y/o exigiendo un seguro de responsabilidad. El también alemán Gerhard WAGNER¹⁵ señala que la personificación de los sistemas inteligentes sólo resulta útil en el campo de la responsabilidad por daños, y que esta utilidad debe superar un análisis coste/beneficio: Sólo es rentable imponerla si la diferencia entre la riqueza generada por el sistema y los costes de producción (que incluyan la indemnización de daños) resulta positiva. Y esto cargando los costes a quien realmente ha producido los daños, evitando la externalización de los mismos, es decir, que los paguen terceros distintos del auténtico responsable. En este análisis, considera que soluciones como imputar la responsabilidad al fabricante del sistema y aproximar la cuestión a los daños causados por productos defectuosos tiene más ventajas que la personificación del robot.

¹³ EBERS, Martin, «La utilización de agentes electrónicos inteligentes en el tráfico jurídico: ¿Necesitamos reglas especiales en el Derecho de la responsabilidad civil?», *Indret*, julio 2016, pág. 8, en www.indret.com.

¹⁴ EBERS, Martin, «Regulating AI and robotics», *Algorithms and Law*, Edited by Martin Ebers and Susana Navas, Cambridge University Press, 2020, p. 61.

¹⁵ WAGNER, Gerhard, «Robot, Inc.: Personhood for Autonomous Systems?», *Fordham Law Review*, Vol. 88, No. 2, 2019, p. 600.

Por lo demás, que la utilización de la personalidad jurídica era una solución de conveniencia se aprecia también en GARCÍA MEXÍA,¹⁶ que señala que «la instrumentalidad de la personalidad electrónica sería la de perfeccionar la responsabilidad civil (e incluso criminal) respecto de los sistemas inteligentes». Pero además añade que esta personalidad electrónica no constituye un instrumento adecuado para la resolución de cuestiones de responsabilidad, y que el atributo de la personalidad es mejor reservarlo al ser humano. Como recientemente advierte el brasileño AVILA NEGRI,¹⁷ el hecho de que teóricamente sea posible atribuir personalidad a sistemas inteligentes no significa que ésta sea una buena idea. Los juristas, ante las nuevas realidades, reaccionamos aplicándoles por sustitución conceptos ya conocidos, en una actividad que califica este autor de «utilización de metáforas jurídicas», siendo en este sentido metafórico en el que se aplica a los entes robóticos el término «personalidad».

3. La alternativa entre personas y cosas

La doctrina francesa trae a colación en esta discusión la dicotomía personas-cosas, como una alternativa al reiterado empleo del concepto de persona para categorizar a los entes cibernéticos. Como señala LARREGUE,¹⁸ existe en la doctrina francesa actual un rechazo al otorgamiento a los robots de personalidad y subjetividad jurídicas, porque se hace de la distinción entre personas y cosas una *summa divisio* inalterable. En este sentido, apunta Xavier LABÉE,¹⁹ existe una posición muy crítica sobre una posible personalidad o incluso subjetividad robótica, sustentada por autores como Malaurie: «Nuestro Derecho todavía hace distinción entre personas y cosas. Esta distinción es fundamental, y constituye según Monsieur Malaurie, el reto crucial del tercer milenio». Así opinan también juristas como LOISEAU²⁰ (que destaca además que todo el problema radica en gestionar la responsabilidad por daños

¹⁶ GARCÍA MEXÍA, Pablo, «Inteligencia artificial. Una mirada desde el Derecho», *Anales de la Academia Matritense del Notariado*, tomo LX, curso 2019/2020, pág. 155.

¹⁷ AVILA NEGRI, Sergio, SMC (2021) «Robot as Legal Person: Electronic Personhood in Robotics and Artificial Intelligence», *Frontiers in Robotics and AI* | www.frontiersin.org, December 2021, Volume 8, Article 789327, pág. 2. doi: 10.3389/frobt.2021.789327

¹⁸ LARREGUE, Julien, «Un tournant relativiste chez les juristes? La distinction entre les personnes et les choses n'est pas menacée par les robots humanoïdes», *ZILSEL, Sociologie, histoire, anthropologie et philosophie*, 23 février 2019. En https://zilsel.hypotheses.org/3619?utm_source=lettre.

¹⁹ LABÉE, Xavier, «L'intrusion de l'humanoïde dans le Droit de la famille», en *Lex robotica. Le Droit à l'épreuve de la robotique*, Lextenso-LGDJ, Issy-les-Moulineaux, 2018, pág. 95.

²⁰ LOISEAU, Grégoire, «La personnalité juridique des robots: une monstruosité juridique», *La semaine juridique, édition générale*, Juris Classeur, 2018, vol. 22, p. 1039-1042.

producidos por el robot), o FRISON-ROCHE,²¹ contrarios a la admisión de figuras intermedias entre las categorías clásicas de personas o cosas, que reproduce la división entre sujetos y objetos.

Suavizando esta dicotomía, LARREGUE y otros estiman que nada impide calificar a los robots y vehículos-robot como cosas especiales, otorgarles un estatuto jurídico especial, intermedio entre las personas y las cosas, como hacemos con las personas jurídicas, o actualmente con los animales, o con espacios naturales, aludiendo a los casos de los ríos Whanganui (Nueva Zelanda), Ganges y Yamuna (India). Se trataría de un estatuto específico, con capacidades bien delimitadas, y en cualquier caso con una dependencia jurídica de control respecto de una persona, física o jurídica, que tutelaría a este peculiar ente y controlaría su peculiar capacidad. En definitiva, aunque no consideremos a los sistemas inteligentes como seres dotados de una personalidad propia, tampoco debemos reducirlos a meras cosas: estaríamos ante la tradicional solución del *estatuto jurídico propio*.

En la doctrina española, BARRIO ANDRÉS²² incide en la idea de que los robots van a tener una cierta subjetividad, trayendo a colación estudios psicológicos conforme a los cuales los robots serían una nueva categoría ontológica diferenciada de los meros objetos, lo que aconseja «equiparlos» con una «personalidad electrónica independiente». Y ERCILLA²³ estima que los robots capaces de aprender del entorno a través de la inteligencia artificial, y de distinguirse de los demás de su misma versión tomando sus propias decisiones, precisarán un estatus jurídico propio, valorando la necesaria creación de un *tertium genus* de personalidad, distinta de la de la persona física y la persona jurídica. ERCILLA conecta la voluntad que estima que tiene el robot autónomo que toma decisiones propias, con la teoría de la voluntad de Savigny como base del derecho subjetivo, y con la de la autonomía de la voluntad de Kant. Con tales bases filosóficas concluye el autor que la creación de una personalidad jurídica específica para los robots es una cuestión de conveniencia para gestionar la actuación de los futuros robots autónomos.

Frente a esta construcción de una figura intermedia, estimo que estos entes siguen siendo cosas, todo lo especiales que se quiera, pero cosas, o

²¹ FRISON-ROCHE, Marie-Anne, «La disparition de la distinction de jure entre la personne et les choses: gain fabuleux, gain catastrophique», *Recueil Dalloz*, 2017, vol. 41, p. 2386-2389.

²² BARRIO ANDRÉS, «Hacia una personalidad electrónica para los robots», *Revista de derecho privado*, n° 102, Mes 2, 2018, pág. 100.

²³ ERCILLA GARCÍA, Javier, «Aproximación a una Personalidad Jurídica Específica para los robots», *Revista Aranzadi de Derecho y Nuevas Tecnologías*, 2018, Número 47 (mayo-agosto), pág. 4.

“productos”, como se sigue de los más recientes textos europeos. La citada *Propuesta de Reglamento por el que se establecen Normas Armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial)*, de 21 de abril de 2021, abandona completamente la personificación del sistema inteligente y, como precisa DE MIGUEL ASENSIO,²⁴ preve la creación de un Comité Europeo de Inteligencia Artificial, y la designación de autoridades nacionales de supervisión y vigilancia, supervisión relacionada con la vigilancia de la seguridad de productos: En definitiva, la IA aparece cosificada como un mecanismo o producto. Esto nos lo ratifica la *Propuesta de Directiva del Parlamento Europeo y del Consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial*, COM(2022) 496 final 2022/0303 (COD), de 28 de septiembre 2022, que frente a la opacidad y dificultades de prueba derivadas de los sistemas inteligentes, propugna que la IA será fiable cuando las normas garanticen «que las víctimas de daños causados por la IA obtengan una protección equivalente a la de las víctimas de daños causados por los demás productos». Es decir, que ubica la responsabilidad por daños cerca de la derivada del producto defectuoso o inseguro.²⁵

De forma paralela a esta Propuesta, se dicta también la *Propuesta de Directiva acerca de la modificación de la responsabilidad por productos defectuosos*, COM(2022) 495 final, de 28 de septiembre 2022. Señala NAVAS NAVARRO²⁶ que esta segunda Propuesta de Directiva parte de considerar al sistema de IA como producto: «El art. 4 (1) de la Propuesta considera que dentro de la categoría “producto” debe comprenderse el “software”. Y según el considerando núm. 12 tal expresión, que no define, debe comprender también a los “sistemas de IA”. Por supuesto, sin diferenciar entre sistemas de alto y bajo riesgo». Estima la autora que pese a esta ampliación de la regulación existente

²⁴ DE MIGUEL ASENSIO, Pedro A., “Propuesta de Reglamento sobre inteligencia artificial”, *La Ley Unión Europea*, nº 92, mayo 2021, pág. 8.

²⁵ En este sentido, el texto de la Propuesta configura un sistema de responsabilidad subjetiva o por culpa en el que las reclamaciones de daños se ven favorecidas por un deber de exhibición de pruebas por parte del responsable del sistema de alto riesgo, deber de exhibir o conservar las pruebas que si se incumple permite presumir el incumplimiento por parte del demandado de la diligencia pertinente, y por tanto su culpa. A ello se añade la presunción del nexo causal entre la culpa del demandado y los daños, siempre y cuando se cumplan todas las condiciones siguientes: a) la demostración o presunción de culpa del demandado; b) la probabilidad de la intervención en la causación del resultado, y, finalmente, «c) que el demandante haya demostrado que la información de salida producida por el sistema de IA o la no producción de una información de salida por parte del sistema de IA causó los daños».

²⁶ NAVAS NAVARRO, Susana. «Seguimos necesitando normas de responsabilidad civil en caso de daños ocasionados por sistemas de inteligencia artificial de alto riesgo». *Revista CESCO de derecho de consumo*, octubre 2022, págs. 1-5.

para gestionar cuestiones de responsabilidad por daños cibernéticos, quedarían fuera de la misma las actuaciones dañosas de sistemas autónomos de *machine learning*, para los que sería precisa una nueva regulación *ad hoc*.

Esta es un poco la idea en la que insisten recientes elaboraciones prelegislativas en el ámbito europeo y en el americano: la inteligencia, por sí sola, no debe implicar la personalidad. Para evitar la contaminación entre estos conceptos, el reciente Informe europeo sobre «Inteligencia artificial en la era digital», de 5 de abril 2022,²⁷ revisa el término IA y nos dice que «...es un término general que designa una amplia gama de tecnologías, técnicas y enfoques, antiguos y nuevos, mejor entendidas como ‘sistemas de inteligencia artificial’, que se refiere a cualesquiera sistemas ...guiados por un conjunto dado de objetivos definidos por el ser humano, con diversos grados de autonomía en sus acciones, y que responden con predicciones, recomendaciones o toma de decisiones basadas en los datos disponibles». Añadiendo que, si bien algunas de estas tecnologías ya son de uso general, «otras están todavía en desarrollo o son solo meras especulaciones, que pueden llegar o no a existir en el futuro».

Y en EEUU, la reciente *Algorithmic Accountability Act* de 2022²⁸ deja de hablar de IA, y para referirse a los sistemas inteligentes emplea la expresión *sistema de decisión automatizada*, que se refiere, según esta misma Ley, a «cualquier sistema, software o proceso que utiliza la informática (incluido el derivado del aprendizaje automático, las estadísticas u otras técnicas de procesamiento de datos o inteligencia artificial y excluyendo la infraestructura informática pasiva), cuyo resultado sirve de base para una decisión o juicio».

²⁷ *Report on artificial intelligence in a digital age*, (2020/2266(INI)), Special Committee on Artificial Intelligence in a Digital Age, Rapporteur: Axel Voss. Pág. 12/65: ... *the term AI is an umbrella term covering a wide range of old and new technologies, techniques and approaches better understood as ‘artificial intelligence systems’, which refers to any machine-based systems that often have little more in common than being guided by a given set of human-defined objectives, with varying degrees of autonomy in their actions, and engaging in predictions, recommendations or decision-making based on available data; notes that while some of these technologies are already in widespread use, others are still under development or are even just speculative concepts that may or may not exist in the future.*

²⁸ Introduced in House (02/03/2022), 117th CONGRESS, 2d Session, H. R. 6580.

4. Una cierta subjetividad: el robot como “agente”

Como dice FLORIDI,²⁹ lo que realmente importa es el reconocimiento de la creciente presencia de tecnologías cada vez más inteligentes en nuestras vidas, con enormes efectos en cuanto a cómo nos concebimos a nosotros mismos, al mundo y a nuestras interacciones con el mundo: «El asunto no consiste en que nuestras máquinas sean conscientes, o inteligentes, o capaces de saber algo como nosotros. No lo son... El asunto es que cada vez son más capaces de hacer frente a más y más tareas, mejor que nosotros, incluida la predicción de nuestros comportamientos. No somos los únicos agentes inteligentes en el mundo, de ninguna manera».

Aunque las máquinas inteligentes no tengan personalidad jurídica, aunque sean cosas y no personas, esto no significa que no estemos ante entes que actúan en la vida social, y por tanto en la jurídica. Vamos a ver, la última vez que compró un libro o un cacharro cualquiera por Amazon ¿De verdad pensó que estaba interactuando con un vendedor de carne y hueso? Sabe usted muy bien que no es así, todo el proceso de contratación tuvo lugar entre un ser humano, usted, y un sistema inteligente de venta, que además y una vez terminada la transacción se ocupa de monitorizar la entrega del objeto y sus incidencias. Ante esta realidad cotidiana ¿Podemos, como juristas, ignorar la incidencia en la vida jurídica de estos nuevos entes? ¿Cómo podemos regular su existencia y su interacción con los seres humanos?

Como señala AVILA NEGRI,³⁰ siguiendo a Teubner, el individualismo metodológico racionalista hizo que el número de actores en el mundo jurídico fuera drásticamente reducido por el discurso filosófico de la modernidad, conforme al cual el único actor plausible es el individuo humano. A partir de aquí, la pretendida personalidad jurídica electrónica, o para robots, no es sino parte de un debate más amplio: el del reconocimiento de nuevas subjetividades y, en consecuencia, nuevos actores legales. Esto siempre que el debate no imponga la aceptación de nuevos sujetos de derechos, porque entonces estaríamos reproduciendo lo dicho sobre la personalidad, y aceptando nuevas personas.

Entre los especialistas en Inteligencia Artificial, la consideración subjetiva de los sistemas de IA se da por supuesta como derivada de su naturaleza

²⁹ FLORIDI, Luciano, «Singularitarians, Atheists, and Why the Problem with Artificial Intelligence is H.A.L. (Humanity at Large), not HAL», *American Psychological Association Newsletter*, Spring 2015, vo-lumen 14, number 2., p. 10.

³⁰ AVILA NEGRI, Sergio, SMC (2021) «Robot as Legal Person: Electronic Personhood in Robotics and Artificial Intelligence», *Frontiers in Robotics and AI* | www.frontiersin.org, December 2021, Volume 8, Article 789327, pág. 7. doi: 10.3389/frobt.2021.789327

inteligente, así RUSSELL y NORVIG,³¹ que nos dicen que los sistemas inteligentes son «agentes», en el sentido de que «hacen cosas», y en particular operan de modo autónomo, perciben el medio, persisten en su actuación durante un tiempo prolongado, se adaptan a los cambios y crean y persiguen objetivos: «Un agente racional es uno que actúa de manera que alcanza el mejor resultado, o, cuando éste es incierto, el mejor resultado esperable». WALLACH y ALLEN,³² por su parte, dan un paso más y consideran a los sistemas inteligentes como «agentes morales» (AMA, Agente Moral Artificial o *Artificial Moral Agent*), aunque esto más en el sentido de agentes que toman decisiones propias, es decir, un significado más semejante al de autonomía decisional que permite actuaciones calificables de correctas o erróneas; en este sentido se expresa el diccionario Merriam-Webster en su primera acepción, letra e: *capable of right and wrong action //a moral agent*.

SIAU y WANG,³³ por su parte, exponen las características más consolidadas en la doctrina para calificar a un sistema inteligente como «agente moral», que serían las de *autonomía*, en cuanto no han de estar controladas por otro agente; *intencionalidad*, en el sentido de que sus acciones sean deliberadas y estén debidamente calculadas; y finalmente, *responsabilidad*, que indica que las máquinas cumplen un rol social que implica la asunción de responsabilidades. Mientras que FLORIDI y SANDERS³⁴ estiman que una definición de agente de un adecuado nivel de abstracción debe acomodarse a los siguientes tres criterios: interactividad, autonomía y adaptabilidad: «Interactividad significa que el agente y su entorno (pueden) actuar entre sí... Autonomía significa que el agente puede cambiar de estado... Adaptabilidad significa que las interacciones del agente (pueden) cambiar las propias reglas por las que cambia de estado». Esta última propiedad asegura que un agente puede aprender de su experiencia.

Ahora bien, la idea de "moralidad" del agente indica que su actividad puede ser objeto de valoración moral, pero no que el sujeto tenga una moral al modo humano. Como señala COECKELBERGH, las máquinas pueden ser «agentes», pero no propiamente «agentes morales», pues les falta la condición del control que para la responsabilidad moral exigía Aristóteles en su *Ética a Nicómaco*: el

³¹ RUSSELL, Stuart, y NORVIG, Peter, *Artificial Intelligence. A Modern Approach*, Pearson Education Limited, Harlow, 2016, pág. 2.

³² WALLACH y ALLEN, *Moral Machines. Teaching robots right from wrong*, cit., pág. 16.

³³ SIAU Keng y WANG Weiyu, "Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI", *Journal of Database Management (JDM)*, vol. 31, Issue 2, abril-junio 2020, p. 81.

³⁴ FLORIDI Luciano, SANDERS, J.W., "On the Morality of Artificial Agents", *Minds and Machines* 14(3): 349-379, August 2004, DOI: 10.1023/B:MIND.0000035461.63578.9d.

que la acción tenga origen en el agente, cuando aquí lo tiene en el humano que ordena la agencia al robot.³⁵

Nathalie NEVEJANS³⁶ nos explica que la voluntad de atribuir personalidad jurídica a un robot parte de una visión limitada del concepto de robot, que se entiende como robot humanoide. Si los constructores dan forma humana a los robots es para así provocar en el utilizador un sentimiento de empatía, que dará paso a la aceptación, necesaria para los robots-cuidadores, sobre todo. Pues bien, es este fenómeno de la empatía artificial el que nos hace creer que el robot está dotado de comportamiento humano o de emociones: «En el actual estadio de desarrollo de la robótica, querer atribuir personalidad jurídica a un robot humanoide equivale a ser víctima de esta empatía artificial».

Esta última consideración nos lleva a un ámbito distinto, el del robot humanoide, es decir, el mismo sistema de IA pero incorporado (nunca mejor dicho) a una máquina que imita a un ser humano. Este dispositivo, según nos advierten los autores, es capaz de despertar en nosotros una pulsión que nos hace tratarlos como seres vivos. Señala David GUNKEL³⁷ que existe un componente humano que apoya la tendencia a extender la personalidad a los robots, y es nuestra propensión al antropomorfismo: «Los científicos sociales, desde mediados de la década de 1940 (Heider y Simmel 1944; Reeves y Nass 2003) han demostrado que los seres humanos concederán estatus y posición social a los objetos. Esta tendencia es a menudo denigrada y descartada como antropomorfismo. Pero el antropomorfismo no es un error; es una característica de la sociabilidad humana». Y la pulsión para conceder esta personalización al objeto proviene no de exigencias del objeto inteligente, sino del grupo social en el que interactúa. Pone el autor el ejemplo de los robots desactivadores de explosivos, que acaban siendo tratados por los especialistas humanos como uno más del grupo. Más adelante nos señala que los robots tienen presencia social, y que no podemos evitar responder a estos artefactos de una manera similar a cómo se responde a un ser humano, porque como predijo Norbert Wiener en los años 50, la sociedad actual no se define simplemente por las transacciones de persona a persona sino que también incluye interacciones entre humanos y máquinas.

³⁵ COECKELBERGH, Mark, *Ética de la Inteligencia artificial*, Cátedra, Madrid, 2021, pág. 97. Otra condición es que el agente «sepa lo que hace», que sea consciente de ello, consciencia que es discutible que exista en el robot: “Esta es una condición epistémica: necesitas ser consciente de lo que estás haciendo y conocer las consecuencias que puede tener”.

³⁶ NEVEJANS, Nathalie, *Traité de Droit et d'Éthique de la Robotique civile*, LEH Éditions, Burdeos, 2017, pág. 79.

³⁷ GUNKEL, David J. y WALES, J. Joseph, «Debate: what is personhood in the age of AI?», *AI & SOCIETY* (2021) 36, p. 475, <https://doi.org/10.1007/s00146-020-01129-1>

En definitiva, la consideración del robot en la vida jurídica se puede hacer desde dos puntos de vista: el de la actividad del robot en la sociedad humana y el de la sociedad humana hacia el robot como ente que actúa. Esto nos lleva a dos problemas distintos.

5. Los dos órdenes de problemas de los sistemas inteligentes

En mi opinión, tanto la atribución de personalidad como el reconocimiento de una propia subjetividad no son sino artificios para resolver dos órdenes de problemas completamente distintos.

En primer lugar, tenemos el problema de la utilización del sistema inteligente como herramienta para ampliar las capacidades de los seres humanos y para gestionar intereses humanos, con sus consiguientes repercusiones jurídicas. En este sentido cabe destacar una modalidad del término «agente» que nos aporta bastante significado: aparece con claridad en las acepciones 3ª y 4ª del término (ambas como sustantivo) en el Diccionario de la Real Academia Española de la Lengua: *3. m. y f. Persona que obra con poder de otra. 4. m. y f. Persona que tiene a su cargo una agencia para gestionar asuntos ajenos o prestar determinados servicios.* En suma, el agente, en su significación más jurídica, normalmente actúa para otro o en beneficio de otro. Las repercusiones jurídicas de su uso van a consistir básicamente en las obligaciones que presiden su utilización, los efectos de la actuación del sistema y los resultados producidos, y las obligaciones derivadas de los daños causados por los robots en sus actuaciones –en todas, programadas o no– es decir, el problema de la responsabilidad civil por el daño cibernético.

En segundo lugar, está el problema del encaje del robot, especialmente el robot humanoide, en la sociedad humana, es decir, el robot como sujeto de interacciones por parte de los humanos, y su consideración como ente merecedor de una especial atención moral: esto es lo que los autores denominan la *valencia social* del robot, que va a ser objeto de atención por parte de la ética pero posiblemente también tenga consecuencias jurídicas.

6. La atribución de los resultados del sistema inteligente

6.1 La responsabilidad civil por daños

El primer orden de problemas se centra en la gestión de los resultados (buenos o malos) de la actividad robótica, ámbito en el que la personificación fue ante

todo un recurso para intentar solucionar los problemas en materia de responsabilidad por daños, careciendo de base filosófica alguna. Así, como opina PALMERINI,³⁸ mediante la personalidad electrónica sólo se trataba de resolver el problema de la responsabilidad robótica, siendo éste un mecanismo poco adecuado: Muchos autores estiman que bastan los recursos jurídicos habituales para resolver los eventos de daño y responsabilidad civil, sin necesidad de acudir a esta novedad. Algunos optan por un sistema de responsabilidad objetiva unido a la contratación de un seguro obligatorio, tomando como modelo la responsabilidad en la circulación de vehículos a motor. Así NAVAS NAVARRO³⁹ que estima que para la responsabilidad por daños a tercero no sería necesaria una norma específica que regulara la responsabilidad robótica, sino que el legislador debería simplemente regular "...una responsabilidad civil por la tenencia de bienes potencialmente peligrosos, de modo general, mediante la modificación del Código civil". EBERS, como ya hemos visto, considera que la solución de la personalidad electrónica es ineficaz y que basta para resolver la mayor parte de los problemas con establecer una responsabilidad objetiva y un contrato de seguro de responsabilidad obligatorio.

Es también la idea que se sigue en la Resolución de 20 de octubre de 2020 sobre *Régimen de responsabilidad civil en materia de inteligencia artificial*,⁴⁰ que viene acompañada de un Proyecto de Reglamento en 15 artículos. Impone una responsabilidad objetiva para los sistemas de IA de alto riesgo, determinando la acción para reclamarla (arts. 4 a 7). Los sistemas de IA que no sean de alto riesgo se van a ver sujetos a un sistema de responsabilidad subjetiva que entiendo que no será la ordinaria, pues se acerca al sistema de responsabilidad por riesgo, o a la responsabilidad por el producto defectuoso.

Por su parte el Informe europeo sobre *Inteligencia artificial en la era digital*, de 5 de abril 2022⁴¹, señala que debido a las características propias de los sistemas de IA (complejidad, conectividad, opacidad, vulnerabilidad, capacidad de ser modificado mediante actualizaciones, capacidad de autoaprendizaje y autonomía potencial), así como la variedad de actores

³⁸ PALMERINI Erica, «Robótica y derecho: sugerencias, confluencias, evoluciones en el marco de una investigación europea», *Revista de Derecho Privado*, n.º 32, enero-junio de 2017, pág. 77.

³⁹ NAVAS NAVARRO, Susana, "Smart robots y otras máquinas inteligentes en nuestra vida cotidiana", *Revista CESCO de Derecho de Consumo*, N° 20/2016, <http://www.revista.uclm.es/index.php/cesco>, pág. 91.

⁴⁰ P9_TA-PROV(2020)0276. Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial (2020/2014(INL)).

⁴¹ *Report on artificial intelligence in a digital age*, (2020/2266(INI)), cit., pág. 37/65.

involucrados en su desarrollo, implementación y uso, «... considera, por lo tanto, que aunque no hay necesidad de una revisión completa del buen funcionamiento de los regímenes de responsabilidad, son necesarios ajustes específicos y coordinados de las normas europeas y nacionales sobre regímenes de responsabilidad para evitar una situación en la que las personas que sufren daños o cuyos bienes resultan dañados queden sin compensación; especifica que mientras que los sistemas de IA de alto riesgo deben estar sujetos a leyes de responsabilidad objetiva, combinadas con un seguro obligatorio, cualquier otra actividad, dispositivo o proceso impulsado por sistemas de IA que pueda causar daño debe permanecer sujeto a la responsabilidad basada en la culpa; cree que la persona afectada no obstante, debe beneficiarse de una presunción de culpa por parte del operador, a menos que este último pueda probar que ha cumplido con su deber de cuidado».

Sin embargo TEUBNER,⁴² en lo que se refiere a responsabilidad extracontractual, rechaza la idea de aplicar una responsabilidad objetiva o por riesgo derivado del uso de un instrumento peligroso. Opina el autor que lo apropiado sería aplicar una responsabilidad por conducta indebida o ilegal de los agentes de software o «asistentes digitales» que toman las decisiones autónomas como responsabilidad vicaria o indirecta. En esta línea, estima CARRASCO⁴³ que el esquema en el que se insertaría esta construcción es el de la responsabilidad por hecho de otro, como la recogida en nuestro sistema en el art. 1903 Cc. En realidad, la indemnización del daño quedaría garantizada, lo mismo que en la responsabilidad objetiva, porque se incluye en el esquema la necesidad de un seguro obligatorio que extienda la reparación al mayor número posible de casos, acallando las discusiones teóricas. Pero esto da como resultado que no haya diferencia entre la responsabilidad objetiva y la vicaria por empleo del «asistente digital». Aunque también valora como un tratamiento adecuado dar a los agentes de software el trato de los animales, aplicando la responsabilidad del art. 1905 Cc., en lugar de considerarlos empleados conforme al art. 1903 Cc.

6.2 Los beneficios de la actividad robótica

La actividad de los sistemas inteligentes no siempre va a ser dañina, al contrario, normalmente producirá beneficios para su utilizador. Es el caso ya

⁴² TEUBNER, Gunther, «Digitale Rechtssubjekte? Zum privatrechtlichen Status autonomer Softwareagenten – Digital personhood? The status of autonomous software agents in private law», *Ancilla Iuris*, 2018, 106, pág. 137.

⁴³ CARRASCO PERERA, Ángel, “A propósito de un trabajo de Gunter Teubner sobre la personificación civil de los agentes de inteligencia artificial avanzada”, *Centro de Estudios de Consumo, Publicaciones Jurídicas*, en <http://centrodeestudiosdeconsumo.com>, Fecha de publicación: 11 de enero de 2019.

citado de los sistemas comerciales de venta por Internet, o el de los programas de gestión financiera, compraventa de acciones y activos, o los robots cuidadores en residencias o prestadores de servicios de cualquier tipo, dentro de poco, servicios de transporte. En todos estos casos, la actuación del sistema inteligente es meramente vicaria de su utilizador, es una mera herramienta. En mi monografía *Robots y personas* estudio su parificación con el esclavo romano, que se consideraba un instrumento –inteligente y humano, eso sí– de su amo. Esta figura del esclavo en el Derecho romano presenta mucha utilidad para el tratamiento jurídico de los entes o «agentes» robóticos: El esclavo, aun siendo considerado como un ser humano, carecía de personalidad, y su conducta podía producir efectos jurídicos, pero no para él, sino para su dueño. En palabras de ÁLVAREZ SUÁREZ,⁴⁴ las actuaciones jurídicas que realiza el esclavo «...producen sus efectos a favor de su señor; éste es el que adquiere y no el esclavo, sin que sea necesario que tales actos se realicen con conocimiento o por mandato de éste». Cuando adquiere algún bien lo adquiere para el dueño, no para sí, y no puede contraer obligaciones como propias, ni tampoco para su dueño pues éste es un extraño al vínculo obligatorio; todo lo más puede generar obligaciones naturales, cuyo cumplimiento el dueño asumirá, pero sólo voluntariamente.

La consideración y el estatus jurídico del esclavo en Roma es un modelo que puede ayudarnos a la hora de construir un estatus especial para los robots autónomos. Como dice ERCILLA,⁴⁵ en el robot encontramos, lo mismo que en los esclavos en el Derecho romano, «...que las dos características citadas se encuentran íntimamente entrelazadas, a saber, el elemento físico y el elemento intelectual y autónomo... el elemento físico, como mecanismo, estaría inexorablemente unido al elemento algorítmico, que determina la inteligencia y la autonomía. Del mismo modo que en el esclavo, el carácter jurídico de cosa estaba íntimamente vinculado a la consideración de ser humano».

La consideración del esclavo como instrumento inteligente de su dueño, le acerca a los modernos sistemas dotados de inteligencia artificial. Abunda en esta idea Gunther TEUBNER:⁴⁶ «Los agentes de software son esclavos digitales, si bien esclavos con capacidades sobrehumanas». Perfila este autor un estatus en el que el «agente electrónico» o «asistente digital» que intervenga en la negociación de contratos tenga al menos la capacidad jurídica necesaria para

⁴⁴ ÁLVAREZ SUÁREZ, Ursicino, *Instituciones de Derecho romano – III. Personas jurídicas y colectivas en el Derecho romano*, EDERSA, MADRID, 1977, pág. 31.

⁴⁵ ERCILLA GARCÍA, «Aproximación a una Personalidad Jurídica Específica para los robots», *Revista Aranzadi de Derecho y Nuevas Tecnologías...*, cit., pág. 6.

⁴⁶ TEUBNER, Gunther, «Digitale Rechtssubjekte? Zum privatrechtlichen Status autonomer Softwareagenten – Digital personhood? The status of autonomous software agents in private law», *Ancilla Iuris*, 2018, 106, pág. 114.

trasladar los efectos de la declaración negocial a su utilizador, principal o *dominus*, a título de representante, alegando que es la solución propuesta por la *Section 14* de la *Uniform Electronic Transactions Act* estadounidense, que declara que los contratos celebrados entre algoritmos son válidos y vinculantes: *A contract may be formed by the interaction of electronic agents of the parties, even if no individual was aware of or reviewed the electronic agents' actions or the resulting terms and agreements.*

La idea de representación aparece también en la doctrina alemana, que considera que la declaración contractual automatizada es la del utilizador del agente, pues éste carece de voluntad propia y se limita a anunciar la de su operador en base a su programación. Señala TEUBNER que el Tribunal Federal estima que la declaración de un sistema computerizado es la de la persona o compañía que lo utiliza como medio de comunicación. Estima este autor que, aplicando por analogía el esquema del contrato de agencia hay que entender que la declaración es propia del agente electrónico, pero que actúa como representante de su principal, para lo cual basta con otorgar al sistema la capacidad parcial y restringida de actuar como representante. El inconveniente de la falta de una genuina voluntad contractual en el agente no humano (en realidad de una voluntad propia) se remediaría acudiendo al principio de la confianza y buena fe entre las partes contractuales.

Sin embargo, seguimos teniendo el obstáculo de la falta de subjetividad y personalidad del sistema inteligente, que hace que no pueda ser considerado representante de nadie: El mecanismo por el que se traslada el efecto de la declaración al *dominus* o principal no puede ser el de la representación, porque para ser representante se exige personalidad y capacidad, y ya hemos visto sobradamente que el sistema inteligente carece de lo primero y por tanto de cualquier capacidad.

Mezclando ideas ya expuestas, la doctrina más ajustada apoya el construir una entidad representativa electrónica que remedie la falta de subjetividad, que además se basaría en la idea de agencia, configurándose como una suerte de dependiente o empleado del principal, ostentando una representación funcional.

La alternativa sería la ficción de entender que el agente electrónico es un mero instrumento y que la declaración es del utilizador, como si se tratase de un teléfono que se limita a reproducir la voz del que lo utiliza. Sin embargo, esta solución tampoco es viable, simplemente porque esto no es así. Hoy por hoy los sistemas inteligentes llevan a cabo decisiones con absoluta autonomía respecto de sus dueños: Es posible que la decisión esté programada, o que se mantenga dentro de un rango decisonal cerrado, o que el sistema goce de mayor o menor libertad según circunstancias de mercado o campañas

promocionales, por ejemplo, para la fijación de precios. Pero dentro de estos parámetros la decisión es del sistema, de la máquina, y no de su dueño o de su programador. Se dirá que lo mismo ocurre con cualquiera de los dependientes de una gran empresa, que deciden sin que el Consejo de administración tenga idea de sus decisiones, pero esta situación no es parificable porque el dependiente tiene personalidad, y actúa en representación de su principal.

Para que la decisión de un sistema inteligente tenga plena eficacia jurídica es necesario que se reconozca primero que la actuación del sistema inteligente es autónoma y propia de la máquina. Otra cosa es que la falta de personalidad de la misma haga necesario encontrar un ser dotado de personalidad, natural o jurídica, que asuma los resultados de la actuación del robot. Es exactamente esto lo que ocurre con los sistemas inteligentes de creación artística, musical, literaria o pictórica, en los que el autor de la obra es, en mi opinión, la propia máquina, pero la ley impide esta autoría robótica o cibernética porque sólo reconoce como posible autor a la persona natural.⁴⁷ Mi propuesta, en este ámbito, consiste en separar la autoría de la titularidad de derechos sobre la obra intelectual o artística, y poner estos derechos en manos del ser humano o la persona jurídica que se halle en una mejor posición para la utilización de la máquina y la reclamación de sus resultados, es decir, las obras.

Esta idea, trasladada a cualquier resultado de la actuación de un sistema inteligente, consistiría en imputar dichos resultados a la persona física o jurídica que tenga mejor derecho a la utilización del sistema, así por ejemplo la empresa de ventas asume tanto las ganancias de la plataforma inteligente de ventas como la responsabilidad por los fallos o daños que el sistema haya causado, o la sociedad gestora de la residencia de ancianos asume tanto los beneficios de la actuación de los cuidadores robóticos como la indemnización de los daños que su actuación pueda generar.

Tampoco se trata de algo realmente nuevo: si volvemos a la consideración jurídica de la actuación del esclavo en Derecho romano, cuando el esclavo que gestionaba un comercio vendía y compraba género, o el esclavo capataz de una finca tomaba decisiones para la siembra o la venta de animales, está claro que quienes actuaban eran ellos. Pero desde luego los resultados de su actuación jamás ingresaban en su patrimonio, sino en el del *dominus*. Y la falta de

⁴⁷ Sin perjuicio de que también dispense protección a las personas jurídicas que intervienen decisivamente en el proceso de creación de la obra. Véase art. 5 del TR de la Ley de Propiedad Intelectual de 1996: *Autores y otros beneficiarios*. 1. *Se considera autor a la persona natural que crea alguna obra literaria, artística o científica*. 2. *No obstante, de la protección que esta Ley concede al autor se podrán beneficiar personas jurídicas en los casos expresamente previstos en ella*.

personalidad jurídica del esclavo no era ningún obstáculo para la admisión de este efecto.

Aplicando este modelo al actual esclavo electrónico, el problema se reduce al reconocimiento de su virtualidad como *sujeto agente* con eficacia en el tráfico jurídico, con una subjetividad no propia sino prestada por su dueño o principal, que le presta también su capacidad (o si se quiere, que actúa su capacidad mediante el instrumento robótico). Esta actuación de la máquina queda reducida en sus efectos al de la traslación de los resultados de la actuación a dicho dueño: Una subjetividad o agencia puramente instrumental, que traslada al ámbito jurídico la capacidad del dueño, y que no conlleva reconocimiento de personalidad alguna. La actuación sería propia del agente, pero realizada para el principal y al amparo de su personalidad, con el resultado de que dicho principal no podría eludir la imputación de resultados por tratarse de actos realizados por una extensión de su personalidad. La idea es clásica, en realidad: Quien crea un Golem o el que utiliza un Golem es responsable de los actos del Golem.

7. El segundo orden de problemas: la consideración que otorgamos los humanos a los entes cibernéticos, y su efecto en nosotros

Esta pregunta incide en el segundo orden de problemas que nos planteará la IA, que es el del tratamiento social del robot. Por ahora, los humanos solamente utilizamos a los simplísimos robots domésticos, como la aspiradora Roomba, o a los industriales, y en este sentido son cosas. El problema está en lo que sentimos o mejor, sentiremos acerca de los robots y sistemas inteligentes futuros, especialmente los robots que interactúen en el medio social.

Desde EE. UU. nos dice Ryan CALO⁴⁸ que la relevancia jurídica de las nuevas criaturas robóticas viene no sólo de su papel instrumental, sino de lo que se conoce como su «valencia social», o valor como sujeto en la apreciación de la gente: «...los robots pertenecerían a una categoría ontológica completamente nueva. Las personas no tienden a considerar a los robots humanoides como seres vivos, pero tampoco consideran que sean objetos. Los sujetos del experimento tienden a atribuir estados mentales a los robots y encuentran difícil observar comportamientos hacia ellos como si se tratase de simples objetos». Concluye el autor que, en definitiva, los robots son seres fronterizos entre los agentes (sujetos) y los meros objetos. Y ya vimos cómo advertía

⁴⁸ CALO, Brian, «Robotics and the Lessons of Cyberlaw», California Law Review, DOI <http://dx.doi.org/10.2139/ssrn.2402972>. June 2015, vol. 103, Issue 3, pág. 546.

David GUNKEL⁴⁹ la tendencia que tenemos los seres humanos al antropomorfismo, tendencia a conceder estatus y posición social a determinados objetos que interactúan socialmente.

Este trato a los robots como algo más que cosas, como cuasipersonas, nos recuerda poderosamente el tratamiento que se da a las mascotas, que ya no son meras cosas sino, de acuerdo con la última reforma legal, seres sensibles. Sin embargo, y con esto profundizaremos en el efecto que la IA puede causar en la sociedad, como destaca WALES,⁵⁰ hay una importante diferencia entre mascotas y robots, y es que, por mucho que esté bien amaestrado y nos obedezca siempre, un perro sigue teniendo su propia vida perruna, mientras que, a diferencia de las mascotas, la personalidad sintética de nuestras IAs estará en todo momento a nuestra disposición. De acuerdo con su diseño, su razón de ser será el servicio a nuestros intereses, y además, y como seremos sus dueños, veremos muy natural que esto suceda de tal modo. Por ello, como señala el citado autor, la propiedad de cuasipersonas va configurar una sociedad en la que se puede producir una pérdida del respeto a la dignidad de las personas, una sociedad esclavista, en suma, cruel con los robots y posiblemente también con los humanos.

Para muchos autores, en este ámbito las conductas relacionales entre humanos y robots (HRI – Human-Robot Interaction), van a primar las normas derivadas de la Ética sobre las elaboradas por el Derecho.

Así, partiendo de la Ética de la Virtud, que desde bases aristotélicas considera que las acciones inmorales perjudican tanto a la sociedad como al propio actor, CAPUCCIO⁵¹ propone lo que llama *Roboetics Virtue*, una Roboética de la Virtud, centrada en el diseño de robots para la mejora del ser humano. Su afirmación central es que la interacción bidireccional humano-robot, puede generar experiencias moralmente significativas de autoconciencia a través de las cuales el ser humano puede desarrollar sus virtudes. Para asegurar este buen fin, el diseño del robot es fundamental, puesto que también se puede producir robots que, al contrario de lo dicho, fomenten el vicio y no la virtud. Pone el

⁴⁹ GUNKEL, David J. y WALES, Jordan Joseph, «Debate: what is personhood in the age of AI?», *AI & SOCIETY* (2021) 36:473–486, p. 475, <https://doi.org/10.1007/s00146-020-01129-1>

⁵⁰ WALES, Jordan Joseph Y GUNKEL, David J., «Debate: what is personhood in the age of AI?», *AI & SOCIETY* (2021) 36:473–486, p. 481, <https://doi.org/10.1007/s00146-020-01129-1>

⁵¹ CAPPUCCIO Massimiliano L., SANDOVAL, Eduardo B., Mubin, Omar, OBAID, Mohammad, VELONAKI Mari, «Can Robots Make us Better Humans? Virtuous Robotics and the Good Life with Artificial Agents», *International Journal of Social Robotics*, <https://doi.org/10.1007/s12369-020-00700-6>, Accepted: 14 September 2020.

autor algunos ejemplos, como el robot servil, que satisface todas las exigencias de su dueño, por ejemplo, un robot cuidador de niños que no sabe decir no, y que va educando a un pequeño monstruo, o robots que satisfacen comportamientos psicopáticos de sus utilizadores. Es decir, tenemos que afinar el diseño de los sistemas inteligentes y robots de manera que su sumisión a las órdenes y deseos del ser humano no favorezca el desarrollo de la peor parte de nuestro yo.

8. Concluyendo: seamos prácticos

Advierte WAGNER,⁵² retomando el argumento de la concesión de la personalidad a las personas jurídicas, que la concesión del status jurídico de entidad dotada de personalidad no tiene lugar sobre base de teorías filosóficas, sino que se trata de una decisión pragmática: «El ordenamiento jurídico es libre para conferir la personalidad jurídica a los sistemas autónomos o negársela. No es la semejanza entre humanos y robots la que obliga al sistema legal a conferir estatus de entidad a sistemas autónomos. El problema real no es si los robots deben ser tratados como seres humanos, sino si existen razones para tratarlos legalmente como personas». Y ni siquiera dándose estas razones tiene por qué concederse una personalidad total o completa, con plenitud de derechos, bastando con reconocer al sistema inteligente cierta entidad solamente en algunos campos de la realidad, aquéllos en los que esta personalidad o subjetividad nos sea útil (a los humanos, por supuesto).

Pues bien: en los epígrafes anteriores se ha examinado la poca utilidad que tiene la personificación de los sistemas inteligentes para resolver los problemas que pueden plantear y gestionar los intereses concurrentes en su funcionamiento. Para solucionar cuestiones de responsabilidad e indemnización del daño cibernético, ya se ha visto que existen suficientes remedios sin tener que acudir a la personificación. Para resolver la cuestión de la imputación de los resultados positivos de la actividad cibernética también se ha visto que basta con estimar que el sistema inteligente es un «agente» que actúa en nombre y para otro, para su usuario, que es tanto responsable de los daños como beneficiario de los rendimientos de la actividad del robot.

En otro ámbito, se trata de resolver los problemas que pueda plantear la interacción (y no convivencia) entre *robots sociales* y seres humanos, y aquí la cuestión no es si se concede personalidad o no a la máquina, sino si nuestra tendencia humana a la antropomorfización, a atribuir valor subjetivo a

⁵² WAGNER, Gerhard, «Robot, Inc.: Personhood for Autonomous Systems?», *Fordham Law Review*, Vol. 88, No. 2, 2019, p. 599.

determinadas máquinas debe llevarnos a imponer un trato jurídico especial hacia las mismas.

La aproximación hacia un trato “humano” de los robots sociales viene de la mano de la Ética, como ya se ha dicho. Mois NAVON⁵³ nos dice que algunos pensadores proponen tratar a los Robots Sociales como “herramientas”. Pero se argumenta que este tratamiento implica que neguemos nuestra propia empatía natural, «...inculcando finalmente disposiciones viciosas en oposición a las virtuosas. Muchos pensadores aplican aquí el enfoque de Kant sobre el trato a los animales: «el que es cruel con los animales será duro también en su trato con los hombres»—afirmando que no debemos maltratar robots para que no pasemos a maltratar a los humanos».

Nos advierte NAVON de que una de las mayores contradicciones en la materia, que señala Coeckelbergh, es el «problema de la brecha», la existencia de una brecha entre lo que nuestra razón nos dice que son los robots sociales, meras máquinas, nuestros sentimientos que nos obligan a tratar al robot como algo más que una máquina. Y estos sentimientos de empatía son provocados por el comportamiento del robot, que es similar al humano : «Este fenómeno de ver a los robots sociales como humanos es conocido como antropomorfismo, pero no termina simplemente con la proyección de creencias, deseos e intenciones humanas al robot, los seres humanos damos un paso más y nos comprometemos emocionalmente con los robots sociales» (*loc. cit.*). Es aquí donde tenemos el problema, la brecha entre nuestra razón y nuestros sentimientos se decide a favor de lo que nos dicen los segundos, y pasamos a considerar al robot no como un objeto, sino como un sujeto.

Por tanto, va a ser en este ámbito de la valoración social del robot, especialmente el robot androide, donde la subjetividad y, posiblemente, personificación del robot va a alcanzar su máxima expresión. Y la finalidad que perseguirá dicha personificación no será económica, pensada para la indemnización de daños, sino puramente moral, para reforzar un especial trato social hacia los robots.

Esta empatía artificial de la que hablaba Nevejans, y nuestra implicación emocional con las máquinas, se verá reforzada si se logra que los comportamientos del robot incorporen respuestas morales, lo que para Antonio

⁵³ NAVON, Mois (2021) «The Virtuous Servant Owner—A Paradigm Whose Time has Come (Again) ». *Frontiers in Robotics and AI*, September 2021, Volume 8, Article 715849, pag. 1. doi: 10.3389/frobt.2021.

CHELLA⁵⁴ y otros investigadores de la Universidad de Palermo se puede lograr mediante la implementación de una autoconciencia en los robots. En concreto sus investigaciones se han extendido en la programación de esta funcionalidad en robots para incrementar la confianza dentro de las unidades militares compuestas por humanos y sistemas inteligentes: «En este escenario, investigamos la capacidad de autoconciencia como un componente de las interacciones de confianza. En particular, implementamos capacidades de autoconciencia al permitir que el robot genere un modelo propio de sus acciones y habilidades».

Pero este es otro problema en el que incidiré en posteriores trabajos.

Bibliografía

- ÁLVAREZ SUÁREZ, Ursicino, *Instituciones de Derecho romano – III. Personas jurídicas y colectivas en el Derecho romano*, EDERSA, MADRID, 1977.
- AVILA NEGRI, Sergio, SMC (2021) «Robot as Legal Person: Electronic Personhood in Robotics and Artificial Intelligence», *Frontiers in Robotics and AI* | www.frontiersin.org, December 2021, Volume 8, Article 789327. doi: 10.3389/frobt.2021.789327
- BARRIO ANDRÉS, «Hacia una personalidad electrónica para los robots», *Revista de derecho privado*, nº 102, Mes 2, 2018.
- BRYSON, J. J., y WINFIELD, A., «Standardizing Ethical Design for Artificial Intelligence and Autonomous Systems», *Computer*, vol. 50, nº 5, <https://doi.org/10.1109/MC.2017.154>.
- CALO, Brian, «Robotics and the Lessons of Cyberlaw», *California Law Review*, DOI <http://dx.doi.org/10.2139/ssrn.2402972>. June 2015, vol. 103, Issue 3.
- CAPPUCCIO Massimiliano L., SANDOVAL, Eduardo B., Mubin, Omar, OBAID, Mohammad, VELONAKI Mari, «Can Robots Make us Better Humans? Virtuous Robotics and the Good Life with Artificial Agents», *International Journal of Social Robotics*, <https://doi.org/10.1007/s12369-020-00700-6>, Accepted: 14 September 2020.

⁵⁴ CHELLA, Antonio, «Self-Consciousness and Theory of Mind for a Robot Developing Trust Relationships», Air Force Research Laboratory - Air Force Office of Scientific Research - European Office of Aerospace Research and Development. En <https://livelink.ebs.afrl.af.mil/livelink/lisapi.dll>, 2 abril 2019.

- CARRASCO PERERA, Ángel, «A propósito de un trabajo de Gunter Teubner sobre la personificación civil de los agentes de inteligencia artificial avanzada», *Centro de Estudios de Consumo, Publicaciones Jurídicas*, en <http://centrodeestudiosdeconsumo.com>, Fecha de publicación: 11 de enero de 2019.
- COECKELBERGH, Mark, *Ética de la Inteligencia artificial*, Cátedra, Madrid, 2021.
- CHELLA, Antonio, «Self-Consciousness and Theory of Mind for a Robot Developing Trust Relationships», Air Force Research Laboratory - Air Force Office of Scientific Research - European Office of Aerospace Research and Development. En <https://livelink.ebs.afrl.af.mil/livelink/llisapi.dll>, 2 abril 2019.
- DE MIGUEL ASENSIO, Pedro A., «Propuesta de Reglamento sobre inteligencia artificial», *La Ley Unión Europea*, nº 92, mayo 2021.
- DENNETT, Daniel C., «Can machines think?», In M. G. Shafto (ed.), *How We Know*. Harper & Row (1984).
- EBERS, Martin, «Regulating AI and robotics», *Algorithms and Law*, Edited by Martin Ebers and Susana Navas, Cambridge University Press, 2020.
- EBERS, Martin, «La utilización de agentes electrónicos inteligentes en el tráfico jurídico: ¿Necesitamos reglas especiales en el Derecho de la responsabilidad civil?», *Indret*, julio 2016, en www.indret.com.
- ERCILLA GARCÍA, Javier, «Aproximación a una Personalidad Jurídica Específica para los robots», *Revista Aranzadi de Derecho y Nuevas Tecnologías*, 2018, Número 47 (mayo-agosto).
- FLORIDI Luciano, SANDERS, J.W., «On the Morality of Artificial Agents», *Minds and Machines* 14(3): 349-379, August 2004, DOI: 10.1023/B:MIND.0000035461.63578.9d.
- FLORIDI, Luciano, «Singularitarians, AItheists, and Why the Problem with Artificial Intelligence is H.A.L. (Humanity at Large), not HAL», *American Psychological Association Newsletter*, Spring 2015, vol. 14, number 2.
- FRISON-ROCHE, Marie-Anne, «La disparition de la distinction de jure entre la personne et les choses: gain fabuleux, gain catastrophique», *Recueil Dalloz*, 2017, vol. 41.
- GARCÍA MEXÍA, Pablo, «Inteligencia artificial. Una mirada desde el Derecho», *Anales de la Academia Matritense del Notariado*, tomo LX, curso 2019/2020.
- GUNKEL, David J. y WALES, Jordan Joseph, «Debate: what is personhood in the age of AI?», *AI & SOCIETY* (2021) 36:473-486, p. 475, <https://doi.org/10.1007/s00146-020-01129-1>
- HERRERA, Francisco, CORDÓN, Óscar, DEL JESÚS, M^a José, «Una visión actual de la Inteligencia artificial», en *El Derecho y la Inteligencia artificial*, F. Herrera Triguero, A. Peralta Gutiérrez, L. Torres López, coords., Universidad de Granada, Granada 2022.

- HARARI, Yuval Noah, *Homo Deus. Breve historia del porvenir*, Debate, Barcelona, 2016.
- LABÉE, Xavier, «L'intrusion de l'humanoïde dans le Droit de la famille», en *Lex robotica. Le Droit à l'épreuve de la robotique*, Lextenso-LGDJ, Issy-les-Molineaux, 2018.
- LACRUZ MANTECÓN Miguel L., «Inteligencia Artificial: Daño y Prejuicio», *JURISMAT*, Portimão, 2022, n.º 15, pp. 389-414.
- LARREGUE, Julien, «Un tournant relativiste chez les juristes? La distinction entre les personnes et les choses n'est pas menacée par les robots humanoïdes», *ZILSEL, Sociologie, histoire, anthropologie et philosophie*, 23 février 2019. https://zilsel.hypotheses.org/3619?utm_source=lettre..
- LOISEAU, Grégoire, «La personnalité juridique des robots: une monstruosité juridique», *La semaine juridique, édition générale*, Iuris Classeur, 2018, vol. 22.
- MINSKY, Marvin, *La sociedad de la mente*, Ediciones Galápagos, Buenos Aires, 1986.
- NAVAS NAVARRO, Susana, «Smart robots y otras máquinas inteligentes en nuestra vida cotidiana», *Revista CESCO de Derecho de Consumo*, N° 20/2016, <http://www.revista.uclm.es/index.php/cesco>.
- NAVAS NAVARRO, Susana. «Seguimos necesitando normas de responsabilidad civil en caso de daños ocasionados por sistemas de inteligencia artificial de alto riesgo». *Revista CESCO de derecho de consumo*, octubre 2022.
- NAVON, Mois (2021) «The Virtuous Servant Owner—A Paradigm Whose Time has Come (Again) ». *Frontiers in Robotics and AI*, September 2021, Volume 8, Article 715849. doi: 10.3389/frobt.2021.
- NEVEJANS, Nathalie, *Traité de Droit et d'Éthique de la Robotique civile*, LEH Éditions, Burdeos, 2017. SIAU Keng y WANG Weiyu, "Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI", *Journal of Database Management (JDM)*, vol. 31, Issue 2, abril-junio 2020.
- PALMERINI Erica, «Robótica y derecho: sugerencias, confluencias, evoluciones en el marco de una investigación europea», *Revista de Derecho Privado*, n.º 32, enero-junio de 2017.
- RUSSELL, Stuart, y NORVIG, Peter, *Artificial Intelligence. A Modern Approach*, Pearson Education Limited, Harlow, 2016.
- SEARLE, John, *El misterio de la conciencia*, Paidós, Barcelona, 2000.
- TEGMARK, Max, *Vida 3.0. Que significa ser humano en la era de la inteligencia artificial*, Taurus, Barcelona, 2018.
- TEUBNER, Gunther, «Digitale Rechtssubjekte? Zum privatrechtlichen Status autonomer Softwareagenten – Digital personhood? The status of autonomous software agents in private law», *Ancilla Iuris*, 2018, 106.
- TURING, Alan Mathison, «Computing Machinery and Intelligence», *Mind*, 49(1950).

- WAGNER, Gerhard, «Robot, Inc.: Personhood for Autonomous Systems?», *Fordham Law Review*, Vol. 88, No. 2, 2019.
- WALES, Jordan Joseph y GUNKEL, David J., «Debate: what is personhood in the age of AI?», *AI & SOCIETY* (2021) 36:473–486, p. 481, <https://doi.org/10.1007/s00146-020-01129-1>
- WALLACH, Wendell y ALLEN, Colin, *Moral Machines. Teaching robots right from wrong*, Oxford University Press, Nueva York, 2009.