Inteligencia artificial, aprendizaje automático, algoritmos y protección de datos en el marco del RGPD y más allá

PID_00248576

Maja Brkan

Universitat Oberta de Catalunya

Índice

Introducción

«Me encanta», declaró Hui, el múltiple campeón europeo de Go cuando vio que AlphaGo de Google hacía un movimiento poco convencional durante una partida de este juego. Finalmente, AlphaGo venció en este antiguo juego, mucho más complejo que el ajedrez, y demostró que las máquinas pueden superar a los humanos… una vez más.

Referencia bibliográfica

Véase:
Cade Metz (2016). «The Sadness and Beauty of Watching Google's AI Play Go» <https://www.wired.com/2016/03/sadness-beauty-watching-googles-ai-play-go/>. Acceso: 21 de noviembre de 2016.

Por ejemplo, la inteligencia artificial (IA) también superó a un humano jugando al ajedrez (IA Deep Blue) y a Jeopardy (IA Watson).

Con el aumento de las máquinas inteligentes, la irrupción del *big data* y los algoritmos más o menos complejos, los agentes autónomos dotados de inteligencia artificial se están volviendo cada vez más potentes y permiten que numerosas decisiones se tomen de forma automática. El rápido crecimiento del uso de algoritmos, que procesan ingentes cantidades de datos en sectores como el financiero, la banca y los seguros, los servicios médicos, la administración pública y los mercados de valores, ofrece posibilidades infinitas de invención de nuevos algoritmos de aprendizaje automático o la configuración de otros existentes, como por ejemplo los árboles de decisión o las redes neuronales. El uso de estos algoritmos acelera la velocidad y, discutiblemente, también la precisión de la toma de decisiones. A su vez, la toma de decisiones basada en algorítmos puede conducir a decisiones sesgadas, en especial cuando lo que hay en juego son datos tan sensibles como la raza, el origen étnico o la orientación sexual.

Se espera que la inteligencia artificial (IA) sea la gran impulsora de la «cuarta revolución industrial» que se prevé que cambie el funcionamiento de nuestra sociedad y las relaciones entre los seres humanos, que altere el mercado de trabajo y las demandas de empleo, así como los sectores que seguirán el camino de la digitalización. Son numerosos los casos de despliegue de la IA en el mundo moderno actual, como por ejemplo las funciones de generación de voz en teléfonos inteligentes, como Siri, los asistentes personales como Alexa, los vehículos autónomos, las pautas de reconocimiento facial y de voz, los perfiles automatizados que permiten a las empresas enviar publicidad orientada a sus clientes, los robots para cuidados (por ejemplo, el robot terapéutico Paro, <http://www.parorobots.com/>, consultado el 11 de julio de 2017) y los médicos y, por último, los medios cada vez hablan más de la informática cuántica (Adams, 2017). La IA está actualmente en la fase de aprendizaje automático (o aprendizaje profundo) en la que las máquinas son capaces de predecir y de tomar decisiones autónomas.

Referencia bibliográfica

R. L. Adams. (2017). «10 Powerful Examples Of Artificial Intelligence In Use Today». *Forbes* (10 de enero). < https://www.forbes.com/sites/robertadams/2017/01/10/10-powerful-examples-of-artificial-intelligence-in-use-today//#2aba3ab6420d>. Acceso: 10 de julio de 2017.

El propósito de este módulo es doble. Su objetivo principal es analizar los preceptoss del Reglamento general de protección de datos (RGPD) relativos a las decisiones automatizadas en la era del *big data* y explorar cómo garantizar la transparencia de dichas decisiones, en particular la de aquellas tomadas con la ayuda de algoritmos. La segunda finalidad es abordar los otros aspectos legales del campo emergente de la IA, es decir, aquellos que no están regulados por el RPGD. Esta doble orientación del módulo se refleja en su propia estructura.

RPGD

Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 abril de 2016, relativo a la protección de las personas físicas respecto 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 n.º 119, 4 de mayo de 2016, pág.1.

La primera parte del módulo analiza las normas del RPGD y de la Directiva relativa a la protección de datos en asuntos penales con referencia a la toma de decisiones individual y automatizada. Se analizan las correspondientes disposiciones de la legislación europea que regulan la protección de datos, así como las consecuencias para el sujeto afectado por él.

Directiva sobre protección de datos en asuntos penales

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 por parte de las autoridades competentes para los fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, y a la libre circulación de dichos datos y por la que se deroga la Decisión Marco del Consejo 2008/977/JAI, n.º 119, 4 de mayo de 2016, pág. 89.

Además, este módulo aborda la necesidad de elaborar salvaguardas en la toma de decisiones automatizada, como por ejemplo facilitar al interesado una explicación sobre una decisión automatizada, garantizar la transparencia algorítmica y determinar la responsabilidad en la adopción automatizada de decisiones. Se analizan los obstáculos a la transparencia algorítmica y se presta especial atención a los de carácter técnico, a los debidos a la propiedad intelectual y a los relativos a los secretos de Estado y a cualquier otra información confidencial.

La segunda parte de este módulo trata brevemente la cuestión de la responsabilidad de los agentes autónomos dotados de IA, utilizando como ejemplo la responsabilidad de los coches sin conductor, y se aborda la posibilidad de una responsabilidad comjunta con la personalidad jurídica de dichos agentes.

1. Adopción de decisiones automatizadas

La toma de decisiones de forma automatizada se podría definir como tomar una decisión sin intervención humana. Según el RGPD, «la toma de decisiones individual automatizada» es «una decisión basada únicamente en un tratamiento automatizado»[1]. Por supuesto, el humano puede alimentar el sistema con datos –a pesar de que incluso este procedimiento puede ser automático– e interpretar la decisión una vez ha sido tomada. Si la decisión automatizada no tiene ningún efecto vinculante para los interesados y no les priva de sus derechos legítimos, dicha toma de decisión tiene un impacto leve. Sin embargo, cuando la decisión es vinculante para los individuos y afecta a sus derechos, como por ejemplo otorgar un crédito a un cliente, devolver impuestos o acceder a un puesto de trabajo, la ley debe ofrecer suficientes salvaguardas para proteger al individuo (véase más información sobre la eficacia y la imparcialidad de la toma de decisiones automatizada en Zarsky, 2016).

(1)Artículo 22(1) del Reglamento general de protección de datos.

Referencia bibliográfica

Tal Zarsky (2016) «The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making». *Science, Technology, & Human Values* (n.º 41, págs. 118-132).

Parece que la toma de decisiones automatizada abarca muchos tipos de decisiones, desde mostrar resultados de búsqueda, elaboración de perfiles (Hildebrandt, 2008), negociación de alta frecuencia (Pasquale, 2015; Loveless *et al*., 2013), decisiones de un banco para otorgar un crédito, decisiones de la administración (por ejemplo, qué empresa inspeccionar con fines fiscales, Perry, 2017) y, hasta cierto punto, incluso decisiones judiciales (Banco-Capon, 2015; Sartor, 1998; Christin, 2015). La noción de toma de decisiones automatizada no es un concepto unitario que comprenda solo un tipo específico de decisiones. Por el contrario, es más bien amplio, con múltiples caras y susceptible de ser dividido en diversas subcategorías. Antes de analizar las disposiciones del RGPD y de la Directiva sobre protección de datos en asuntos penales (Directiva 2016/680), es importante distinguir entre la toma de decisiones automatizada en el ámbito procedimental y en el sustantivo, entre decisiones automatizadas algorítmicas y no algorítmicas, y entre decisiones basadas en normas frente a las basadas en leyes:

Referencias bibliográficas

En relación a los perfiles, véase:

Mireille Hildebrandt, Serge Gutwirth (eds.) (2008). *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Springer.

En relación a la negociación de alta frecuencia, véase:

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

Jacob Loveless*et al*. (2013). «Online Algorithms in High-frequency Trading.The challenges faced by competing HFT algorithms». *ACM Queue* (vol. 11, núm. 8).

En relación a las decisiones administrativas, véase:

Melissa Perry (2017). «iDecide: Administrative Decision-Making in the Digital World». *Australian Law Journal*.

En relación a decisiones judiciales, véase:

Trevor Bench-Capon, **Thomas F. Gordon** (2015). «Tools for Rapid Prototyping of Legal Cased-Based Reasoning». *ULCS* (15-005). United Kingdom: University of Liverpool.

Giovanni Sartor, **Luther Branting** (eds.) (1998). *Judicial Applications of Artificial Intelligence*. Springer.

Angéle Christin, **Alex Rosenblat**, **Danah Boyd** (2015). «Courts and Predictive Algorithms». *Data & Civil Rights: A New Era of Policing and Justice* <http://www.law.nyu.edu/sites/default/files/upload_documents/Angele%20Christin.pdf>. Acceso: 16 de enero de 2017.

- **Sustantivo/procedimental**. La division entre procedimental y sustantivo no se refiere a tomar decisiones de una naturaleza u otra, sino que significa que las decisiones automatizadas se tendrán que adoptar de manera que se garantice la imparcialidad y la precisión procedimental y sustantiva. El requisito de imparcialidad procedimentl implica que todas las decisiones relativas a los mismos hechos o comparables deben tomarse con arreglo al mismo procedimiento automatizado[2]. La imparcialidad procedimental está estrechamente vinculada con la sustantiva, puesto que llevará a que los mismos casos tengan los mismos resultados. No obstante, las decisiones también deben ser sustantivamente justas, es decir, no deben ser discriminatorias en ningún sentido, especialmente aquellas tomadas en base a algoritmos (véase Goodman, 2016, sobre discriminación algorítmica).

[(2)]Cada decisión en el procedimiento administrativo estará bajo la ley administrativa procesal, lo que significa que las leyes procesales tendrán que ser enmendadas para ofrecer nuevas normas procesales para las tomas de decisones automatizadas.

Referencia bibliográfica

Bryce Goodman (2016). «Discrimination, Data Sanitisation and Auditing in the European Union's General Data Protection Regulation». *European Data Protection Law Review* (n.º 493).

- **Algorítmica/no-algorítmica**. La toma de decisión algorítmica es un procedimiento automatizado hecho con la ayuda de algoritmos. Las publicaciones existentes no proporcionan una definición común de la noción de algoritmo. Sin embargo, debemos especificar que en este proceso tratamos con algoritmos informáticos que se pueden definir como «un conjunto de pasos para cumplir una tarea que está descrita de forma suficientemente precisa para que un ordenador pueda realizarla» (Coormen, 2013). En la actualidad, muchas decisiones automatizadas, si no la mayoría, se toman con el apoyo de algoritmos. Con el uso creciente del *big data* y la adopción de decisiones cada vez más complejas, la intervención algorítmica se ha vuelto casi indispensable.

Referencia bibliográfica

Thomas H. Coormen (2013). *Algorithms Unlocked*. MIT Press.

- **Decisiones automatizadas basadas en reglas / basadas en leyes.** De hecho, ambos tipos de decisiones se toman en base a normas, pero la fuente de la norma difiere en cada uno de ellos. Para las decisiones basadas en reglas, estas son esencialmente el resultado de una decisión comercial, como por ejemplo la realización de perfiles para una publicidad orientada (por ejemplo, una empresa envía un anuncio sobre vacaciones en Bali a todas las personas que busquen vacaciones en Asia). Las decisiones basadas en leyes se fundamentan en una norma jurídica que es vinculante para todo el mundo. Un ejemplo de norma proclive a ser una decisión automatizada es una que prescriba que a toda persona que exceda el límite de velocidad se le impondrá una multa. A no ser que la norma basada en una ley sea

muy clara y precisa, las decisiones automatizadas basadas en leyes tienen que enfrentarse al reto de las nociones y los textos jurídicos abiertos que requieren una interpretación. La toma de decisiones automatizada presupone que las normas que se tienen que aplicar no están abiertas a interpretación y que no dejan a quien ha de decidir mucha o ninguna discreción para actuar.

1.1. La posición del RGPD sobre la adopción de decisiones individuales automatizadas

Esta sección analiza en profundidad la disposición del RGPD relativa a la toma de decisiones automatizada y explica las circunstancias en las que dicho procedimiento es posible según el RGPD y bajo qué condiciones. No es una sorpresa que el RGPD, en su artículo 22, regule la toma de decisiones individuales automatizadas, incluida la realización de perfiles. Según el primer párrafo de esta disposición:

> «Todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o que le afecte significativamente de modo similar.»

Esta disposición sigue el legado de la Directiva de protección de datos[3], en concreto su artículo 15, según el cual el interesado también tiene derecho a no ser objeto de una decisión que tenga efectos jurídicos o que le afecte significativamente y que esté basada exclusivamente en un tratamiento de datos automatizado. Aunque el texto de la disposición no sufrió grandes cambios con la adopción del RGPD, en la práctica su importancia aumentó al crecer el uso de la adopción de decisiones automatizadas. La Directiva de protección de datos también contenía algunos ejemplos de decisiones individuales automatizadas, particularmente, decisiones «para evaluar determinados aspectos personales relativos a la persona, como su comportamiento en el trabajo, capacidad crediticia, fiabilidad, conducta, etc.». Estos ejemplos demuestran que la disposición de la Directiva de protección de datos parecía centrarse, principalmente, en casos de perfiles basados en tratamiento automatizado de datos, sin incluir otros tipos de decisiones automatizadas que también conllevan tratamiento de datos personales.

[3] Directiva 95/46/CE del Parlamento Europeo y del Consejo, de 24 de octubre de 1995, relativa a la protección de los individuos respecto al procesamiento de datos personales y a la libre circulación de dichos datos, DOUE n.º 281, 23 de noviembre de 1995, pág.31.

En este contexto, es interesante observar cómo el artículo 22 del RGPD se dasarrolló mediante un procedimiento legislativo que llevó a la adopción del mismo, puesto que muestra la evolución hecha desde centrarse específicamente en la elaboración de perfiles hasta una formulación más general que utiliza una noción más amplia de la toma de decisión individual automatizada. A

[4] Artículo 20 de la propuesta de RGPD, COM(2012) 11 final.

[5] Artículo 20(4) *Ibid.*

diferencia del RGPD final, en la propuesta inicial de la Comisión este artículo, titulado «Medidas basadas en la elaboración de perfiles»[4], regulaba este procedimiento basándose en el tratamiento automatizado y no en la toma de decisiones generalmente automatizada, como hace la redacción final del RGPD. Es más, la disposición inicial contenía un párrafo separado sobre la obligación de informar al interesado de la existencia de un tratamiento automatizado y de «los efectos previstos de dicho tratamiento en el interesado»[5]. Así pues, en contraste con la disposición actual, el alcance de aplicación de aquella disposición parece estar más limitado, puesto que solo se aplicaba a la elaboración de perfiles. Además, la obligación de informar al interesado sobre dicho tratamiento se trasladó a los artículos 13 y 14 con arreglo a la obligación general de información que se debe facilitar al interesado. Aunque en la primera lectura que se hizo en el Parlamento Europeo la disposición seguía centrándose en la elaboración de perfiles y a ella añadía el derecho a la intervención humana, se eliminó el párrafo relativo a informar al interesado de los efectos previstos de la elaboración de perfiles[6]. En su primera lectura ante el Consejo, la disposición adoptó la forma del actual artículo 22 del RGPD, sin restringir su alcance meramente a la elaboración de perfiles, sino incluyéndola en una categoría más general de toma de decisiones individual automatizada[7]. No obstante, según se menciona en la propuesta del RGPD[8], esta disposición todavía tiene en consideración la Recomendación sobre elaboración de perfiles emitada por el Consejo de Europa[9].

(6)Artículo 20 (elaboración de perfiles), resolución legislativa del Parlamento Europeo, de 12 de marzo de 2014, relativa a la propuesta de reglamento del Parlamento Europeo y del Consejo sobre la protección de los individuos respecto al tratamiento de datos personales y a la libre circulación de tales datos (Reglamento general de protección de datos) (COM(2012)0011 - C7-0025/2012 - 2012/0011(COD)) (procedimiento legislativo ordinario: primera lectura).

(7)Posición del Consejo en primera lectura con vistas a la adopción de un reglamento del Parlamento Europeo y del Consejo relativo a la protección de las personas físicas en relación con el tratamiento de datos personales y con la libre circulación de tales datos, y que derogue la Directiva 95/46/CE (Reglamento general de protección de datos) –Adoptado por el Consejo el 8 de abril de 2016.

(8)Propuesta de reglamento del Parlamento Europeo y del Consejo sobre la protección de individuos respecto al tratamiento de datos personales y a la libre circulación de dichos datos (Reglamento general de protección de datos) COM(2012) 11 final, pág. 9.

(9)Recomendación CM/Rec(2010)13 del Consejo de Ministros de los Estados miembros sobre la protección de los individuos respecto al tratamiento automático de datos personales en el contexto de la elaboración de perfiles (adoptada por el Consejo de Ministros del 23 de noviembre de 2010 en la 1099.ª reunión de los Delegados de Ministros).

Con independencia de una formulación más amplia del RGPD, se puede cuestionar hasta qué punto el alcance de la aplicación de su artículo 22 cubre la adopción de decisiones que sean realmente más amplias que las decisiones basadas en la elaboración de perfiles.[10] El interesado tiene derecho a no ser objeto de una decisión basada únicamente en el *tratamiento automatizado* y la elaboración de perfiles es un tipo de tratamiento que conduce, mayoritariamente, a decisiones de esta naturaleza. Con arreglo al RGPD, la elaboración de perfiles significa tratar datos personales de forma que se utilizan para «evaluar aspectos personales relativos a la actuación natural de la persona en el trabajo, a su situación económica, salud, preferencias personales, intereses, fiabilidad, comportamiento, localización de movimientos» (artículo 4). Resulta difícil imaginar ejemplos en los que los datos personales del individuo que *no* conducen a una elaboración de perfil lleven a una decisión automatizada. Un posible ejemplo sería la aplicación automatizada de normas fiscales para determinar qué devolución de impuestos tendría un residente. Existen decisiones y predicciones automatizadas que no implican elaboración de perfiles, como

Referencia bibliográfica

Isak Mendoza, Lee A. Bygrave. (2017) «The Right not to be Subject to Automated Decisions based on Profiling». *Legal Studies Research Paper Series* (n.º 20/2017, pág. 7). University of Oslo, Faculty of Law Legal Studies. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2964855>. Acceso: 11 de julio de 2017.

por ejemplo la negociación de alta frecuencia o las predicciones de resultados de decisiones judiciales, pero no conllevan el tratamiento de datos personales y, en consecuencia, no entran en el ámbito del artículo 22 del RGPD.

(10)Isak Mendoza y Lee A. Bygrave («The Right not to be Subject to Automated Decisions based on Profiling». *Legal Studies Research Paper Series* (n.º 20/2017, pág. 7) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2964855>. Acceso: 11 de julio de 2017) señalan acertadamente que el procedimiento legislativo que conduce a la adopción del RGPD da como resultado que no otorga al interesado el derecho a objetar a toda realización de perfiles, sino tan solo a determinados tipos de decisiones que surgen de la elaboración de perfiles.

El artículo 22 del RGPD refleja, por un lado, el escepticismo europeo hacia sesgos y decisiones potencialmente falsas que se pueden adoptar a través de medios automatizados si no hay humanos que las verifiquen. Por otro lado, esta disposición, al ofrecer ciertas garantías al interesado, principalmente el derecho a la intervención humana, aborda preocupaciones en torno a la falta de capacidad de los interesados de influir en decisiones tomadas, cada vez con mayor frecuencia, por medios automatizados (Mendoza, 2017). En una primera impresión, esta posición general negativa hacia dichas decisiones automatizadas aparece como una poderosa fortaleza para proteger a los individuos con firmeza e incluso, potencialmente, para dificultar el futuro desarrollo de la IA en el ámbito de las decisiones. Sin embargo, si se hace una evaluación más extensa, se puede argumentar que esta disposición, aun conteniendo numerosas limitaciones y excepciones, se parece bastante a un queso suizo con agujeros gigantescos.

1.2. Decisiones automatizadas no permitidas por el RGPD y la Directiva relativa a la protección de datos en asuntos penales

Con el fin de que el interesado tenga derecho a no ser objeto de decisiones automatizadas, la propia decisión debe cumplir ciertos requisitos establecidos en el artículo 22(1). Sin embargo, antes de profundizar en estas condiciones es esencial analizar la naturaleza de este «derecho» del interesado.

El «derecho» a no ser objeto de toma de decisiones automatizadas se puede entender bien como un derecho que el interesado puede ejercer de forma activa, bien como un derecho «pasivo» que deben observar los responsables de la decisión automatizada sin que exista una petición expresa por parte del interesado. Si este «derecho» que emana del artículo 22(1) del RGPD se configura siguiendo la primera forma, el ejercicio del derecho dependería de la libre voluntad y elección del interesado. El no ejercicio de este derecho conduciría, en una interpretación apropiada, a que las decisiones automatizadas que tuvieran las características descritas en el artículo 22(1) se adoptarían con arreglo a la ley. Ello podría dar lugar, por ejemplo, a que una decisión totalmente automatizada tuviera consecuencias jurídicas para un interesado, sin ofrecerle las salvaguardas necesarias establecidas en el párrafo 3 de dicha disposición. Por otro lado, la decisión del interesado de ejercer este derecho tendría conse-

cuencias jurídicas poco claras. Por ejemplo, ¿el ejercicio de este derecho podría traducirse como un derecho a objetar, no permitiendo así la decisión automatizada? ¿Podría entenderse como una petición de intervención humana?

En consecuencia, interpretar el artículo 22(1) de forma que proporciona al interesado un derecho que tiene que ejercer de forma activa podría producir efectos perjudiciales para él y resultar contrario al propósito de esta disposición, cuyo objetivo es proteger al interesado frente a una posibilidad general de ser objeto de toma de decisiones automatizadas. Una interpretación sistemática del artículo 22 implica que el RGPD solo autoriza las decisiones automatizadas que cumplan los requisitos del párrafo 2 y que permitan las salvaguardas del párrafo 3 de esta disposición. Por lo tanto, como afirman correctamente Mendoza y Bygrave, es más apropiado interpretar el «derecho» del interesado como una prohibición de la toma de decisiones totalmente automatizadas que los responsables deben cumplir (Mendoza, 9). Esta interpretación del artículo 22(1) alinea esta disposición con el artículo 11 de la Directiva relativa a la protección de datos en asuntos penales, que establece para los Estados miembros una clara obligación de prohibir decisiones automatizadas que tengan determinadas características.

Interpretar el «derecho» del interesado como una prohibición de determinados tipos de decisiones automatizadas también da una visión diferente a las condiciones del artículo 22(1). En base a esta lectura, esta disposición prohíbe una decisión que tenga las siguientes características:

- La decisión debe ser individual (la directiva impone la misma condición).
- La decisión debe basarse exclusivamente en un tratamiento automatizado (lo mismo ocurre en la directiva).
- La decisión debe tener efectos jurídicos o significativos respecto al interesado (la directiva contiene un requisito adicional de efectos legales «adversos»).

Desde esta perspectiva, se debe entender que la primera condición prohíbe las decisiones automatizadas **individuales**, es decir, aquellas que hacen referencia solo a una persona física[11], a un único interesado. Las decisiones individuales pueden ser vinculantes para un individuo (como una decisión para conceder un crédito, la petición de una tarjeta de crédito, otorgar un visado, escoger a un contribuyente para auditar) o no vinculantes (por ejemplo, una elaboración de perfil, el envío de publicidad orientada para un viajero de líneas aéreas en base a su perfil). En línea con el alcance general de la aplicación *ratione personae* del RGPD, que cubre la protección de las personas físicas (artículo 1(1)) y, por ende, regula únicamente la protección de los individuos y no de los grupos, la interpretación textual del artículo 22 del RGPD parece excluir, asimismo, las decisiones *colectivas* que afectan a varias personas físicas o a un grupo de individuos unidos en virtud de sus características comunes, su pertenencia a

[11]Según su artículo 1(1), el RGPD se aplica solo a las personas físicas, no a las jurídicas.

Referencia bibliográfica

Alessandro Mantelero (2016). «Personal Data for Decisional Purposes in the Age of Analytics: From an Individual to a Collective Dimension of Data Protection». *Computer Law & Security Review* (vol. 32, n.º 2, págs. 238-255).

un grupo o por ser residentes de una zona concreta (véase Mantelero, 2016, para los aspectos de la protección de datos colectivos en la era de la analítica del *big data*). Se puede aplicar el mismo razonamiento en relación al artículo 11 de la Directiva relativa a la protección de datos en asuntos penales, el cual también prohíbe las decisiones exclusivamente individuales. Un ejemplo de decisión colectiva en asuntos penales es una decisión automatizada de la policía para aumentar el control policial en una zona geográfica determinada, que afectaría a todos los sujetos residentes en dicha zona. Una decisión colectiva en asuntos no penales sería, por ejemplo, una política de precios dinámica para un producto en función del nivel de ingresos de las personas.

En su actual redactado, ni el artículo 11 de la Directiva 2016/680 ni el artículo 22 del RGPD, leídos junto con sus respectivos artículos 1, cubrirían *ratione personae* (y, por lo tanto, tampoco prohibirían) dicha decisión colectiva. Tanto el RGPD como la directiva parecen seguir la lógica según la cual la razón subyacente para la protección de datos de grupos de individuos difiere de la razón para la protección de datos de un sujeto individual. A este respecto, a veces se defiende que una decisión colectiva no está necesariamente ligada a datos personales de una persona física individual, sino que es fácil que se base en datos anonimizados, lo cual haría que la legislación de la UE relativa a la protección de datos fuera inaplicable[12]. Aun así, la anonimización de los datos no es suficiente si el interesado sigue siendo identificable (sobre identificabilidad, véase Gedefa, 2016). Con el uso y la importancia crecientes del *big data*, se facilita de forma significativa la re-identificación de un individuo perteneciente a un determinado grupo. La clasificación de los individuos en una categoría específica (hombre/mujer, renta baja/alta) permite llevar a cabo decisiones colectivas relativas a dicho grupo. Excluir las decisiones automatizadas colectivas del alcance de aplicación del RGPD no solo daría lugar a un desequilibrio enorme en el tratamiento de las decisiones automatizadas, sino que también abriría la puerta a sortear la prohibición de decisiones automatizadas individuales mediante la adopción de decisiones colectivas siempre que fuera posible[13]. Por consiguiente, en base a la exigencia de un elevado nivel de protección del sujeto, se sostiene que las decisiones automatizadas colectivas deben quedar cubiertas por el ámbito de aplicación del artículo 22 del RGPD y el del artículo 11 de la Directiva 2016/680. Una manera posible de incluir dichas decisiones bajo el ámbito de aplicación de estos dos instrumentos legales es considerar la decisión relativa a un grupo como, en realidad, un conjunto de decisiones individuales. Una interpretación en base al significado (teleolológica) del artículo 22 del RGPD y del artículo 11 de la Directiva 2016/680, junto con la necesidad de garantizar al individuo un nivel elevado de salvaguarda de su derecho fundamental a la protección de datos, podría llevar al Tribunal de Justicia de la UE a adoptar esta posición interpretativa.

En segundo lugar, el RGPD y la Directiva 2016/680 no permiten que una decisión se base única y exclusivamente en un proceso automatizado. Que una decisión sea totalmente automatizada o no depende, en primer lugar, de si es

(12)En virtud del considerando 26 del RGPD y del considerando 21 de la Directiva relativa a la protección de datos en asuntos penales, los «principios de protección de datos […] no serían de aplicación para la información anónima».

(13)Puesto que un grupo de individuos múltiples no constituye necesariamente un grupo de individuos con las mismas características o similares, quizá esto no siempre sea posible

Referencia bibliográfica

Worku Gedefa Urgessa (2016). «The Protective Capacity of the Criterion of "Identifiability" under EU Data Protection Law». *European Data Protection Law Review* (vol. 2, n.º 4, págs. 521–531).

posible la intervención humana en el proceso de toma de decisión. Por ejemplo, si el precio de un producto vendido *online* se determina en base a los datos de ingresos del individuo y se muestra automáticamente en la página web, sin que un ser humano intervenga en el proceso de definir el precio, con toda seguridad dicha decisión se basa únicamente en un tratamiento automatizado. Sin embargo, si el proceso permite la intervención humana, se debe verificar si la mera posibilidad de que un humano tenga poder para cambiar una decisión automáticamente hace que dicha decisión no se base exclusivamente en un proceso automatizado. En otras palabras, si un humano se limita a poner un sello en una decisión automatizada sin verificar si esta es correcta, ¿se puede asumir que dicha decisión no se tomó de forma totalmente automatizada?

La respuesta a esta pregunta debería ser negativa. Dicha interpretación formalista, que implica al humano tan solo como una parte necesaria del proceso pero, en última instancia, dejando el poder de decisión a la máquina, no garantizaría un nivel suficientemente elevado de protección de los datos del interesado. Para que la decisión no se base únicamente en un proceso automatizado, el juicio humano debe llegar a la verificación de la decisión generada por la máquina y el humano debe valorar (Mendoza y Bygrave, 10, señalan que el sujeto tiene que «evaluar activamente el resultado del proceso antes de su formalización como decisión») el contenido de la decisión y participar en la misma de manera simplemente formal, como otra etapa del procedimiento. Es decir, para eludir la prohibición del artículo 22 del RGPD o la del artículo 11 de la Directiva 2016/680, la persona debe utilizar la máquina solo como una ayuda para la adopción de la decisión: la decisión final debe ser adoptada por el humano.

En tercer lugar, el RGPD y la Directiva relativa a la protección de datos en asuntos penales impiden únicamente la toma de decisiones, incluída la elaboración de perfiles, que tenga efectos jurídicos (en el caso de la directiva, efectos jurídicos «adversos») para el individuo o que le afecte significativamente. Aunque ninguno de estos dos instrumentos define la noción de efectos jurídicos, se puede asumir que una decisión que tenga efectos jurídicos es vinculante y tiene un impacto sobre la situación legal o los intereses jurídicos del sujeto. Por ejemplo, la decisión de una autoridad fiscal sobre la devolución de impuestos de un individuo concreto, calculada en base a sus ingresos, es una decisión que tiene efectos legales para dicho individuo dentro de la interpretación del RGPD. La decisión que tome un policía para interrogar a un sujeto o para incautarle el dispositivo móvil, tomada en base a sus datos personales, es una decisión que tiene efectos jurídicos adversos para la persona bajo la perspectiva de la mencionada directiva. Aunque parece relativamente claro determinar qué decisión tendrá efectos jurídicos en un individuo, es menos evidente qué tipo de toma de decisión o de elaboración de perfil le «afectaría significativamente». El RGPD proporciona ejemplos de rechazo de una aplicación de créditos *online* o del uso de decisiones automatizadas para prácticas de contratación electrónica. Estos son casos en los que el interesado actúa como solicitante de una tarjeta de crédito, de un contrato de seguro con

una determinada bonificación o de un puesto de trabajo. Sin embargo, parece que no está tan claro el hecho de establecer un efecto significativo para un individuo cuando se trata de la elaboración de perfiles. Por ejemplo, ¿cuándo «afecta significativamente» a un individuo que Google y Facebook le envíen publicidad? Dados los diferentes impactos potenciales que dicha publicidad orientada puede tener en una persona, resulta casi imposible responder claramente de forma afirmativa o negativa. A modo de ejemplo, si un individuo ignora dicha publicidad orientada y no la sigue, es bastante difícil argumentar que la publicidad le «afecta significativamente». En sentido contrario, si una persona, sistemáticamente, adapta sus decisiones de compra teniendo en cuenta dicha publicidad orientada, el efecto significativo se podría establecer muy fácilmente. Por supuesto, todo ello plantea la cuestión de si se requiere que haya un vínculo causal entre la elaboración del perfil y la acción del sujeto para que se cumpla el criterio del efecto significativo. Por un lado, el requisito de que existiera un vínculo causal aseguraría que solo los casos de publicidad orientada tendrían un efecto significativo en el individuo. Por otro lado, la exigencia de dicho vínculo causal complicaría extremadamente el análisis del efecto significativo. Una manera alternativa de probar la existencia de un efecto significativo en dichos casos sería tomar como modelo a un consumidor medio en lugar del consumidor real a quien iba dirigida la publicidad.

1.3. Decisiones automatizadas autorizadas por el RGPD y la Directiva relativa a la protección de datos en asuntos penales

El RGPD, en su artículo 22(2), y la Directiva 2016/680, en su artículo 11, autorizan expresamente determinados tipos de decisiones automatizadas. Con arreglo al RGPD, la prohibición establecida en el párrafo 1 de dicha disposición «no se aplicará si la decisión:

a) es necesaria para la celebración o la ejecución de un contrato entre el interesado y un responsable del tratamiento;

b) está autorizada por el Derecho de la Unión o de los Estados miembros que se aplique al responsable del tratamiento [...], o

c) se basa en el consentimiento explícito del interesado.» La directiva solo autoriza las decisiones «autorizadas por el Derecho de la Unión o de los Estados miembros que se aplique al responsable del tratamiento.»

La primera posibilidad de decisiones automatizadas permitidas por el RGPD engloba a aquellas que son necesarias para perfeccionar o llevar a cabo un contrato entre el individuo y el responsable del tratamiento. Si el significado de esta disposición debe ceñirse a una interpretación estricta del texto, es cuestionable que alguna vez se abra la puerta a las decisiones automatizadas. Por ejemplo, se puede argumentar que la formalización del seguro o del contrato de préstamo requiere una evaluación de riesgos pero, ¿necesariamente este riesgo debe ser evaluado mediante procedimientos automatizados? Los precios de los vuelos con frecuencia se determinan a través de una «política de precios dinámica» que tiene en cuenta el perfil del comprador potencial. ¿Es realmente necesaria esta determinación automatizada del precio para concluir la realización de este contrato de compra-venta? En consecuencia, se supone que el requisito de «necesidad» deberá entenderse más como «posibilitación» para la firma de un contrato. Si la evaluación automatizada de un riesgo crediticio permite la firma de un contrato en virtud del cual el sujeto recibe una tarjeta de crédito, dicha evaluación autorizó la firma de ese contrato. En ocasiones, estos contratos se denominan «contratos algorítmicos» (para más información sobre este tema, véase Scholz, 2017) y son cada vez más frecuentes en el comercio electrónico. Amazon es el ejemplo más utilizado para ello.

Referencia bibliográfica

Lauren Henry Scholz (2017). «Algorithmic Contracts». *Stanford Technology Law Review* <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=274770>. Acceso 10 de noviembre de 2016

En segundo lugar, las decisiones automatizadas y la elaboración de perfiles están permitidas si las autoriza la legislación de la Unión o del Estado miembro que ofrezca suficientes garantías para proteger los derechos, las libertades y los intereses legítimos del interesado. Un ejemplo de legislación de la Unión que permite potencialmente las decisiones automatizadas con suficientes salvaguardas es la Directiva de PNR[15]. Aunque esta directiva, en principio, no permite una decisión automatizada «que produzca un efecto jurídico adverso en una persona o que le afecte significativamente» (artículo 7(6)), ofrece la posibilidad de una correspondencia o identificación automatizada de las personas que deberá ser examinada en profundidad por las autoridades competentes con vistas a una posible implicación en terrorismo, siempre que dicha correspondencia se revise de forma individual con medios no automatizados[14].

(14)Artículo 6(5) de la Directiva PNR. Compárese también el párrafo 2 de esta disposición y Mendoza y Bygrave, pág. 6.

(15)Directiva (UE) 2016/681 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativa a la utilización de datos del registro de nombres de los pasajeros (PNR) para la prevención, detección, investigación y enjuiciamiento de los delitos de terrorismo y de la delincuencia grave (DOUE n.º 119, 4 de mayo de 2016, pág.132).

Un ejemplo de legislación de un Estado miembro que regula la toma de decisiones automatizada es la ley alemana que implementa el RGPD[16], la cual autoriza expresamente las decisiones automatizadas en el sector de los seguros. Por un lado, se autoriza una decisión si esta se toma en el marco de la realización de un contrato de seguro y si la petición de la persona en cuestión se ha aprobado. Tal y como se clarifica en las explicaciones de esta norma alemana, esta disposición permite dicha decisión automatizada en una relación extracontractual (¡y no contractual!) entre la compañía de seguros de la per-

(16)Véase § 37 (Automatisierte Entscheidungen im Einzelfall einschließlich Profiling) de Gesetz zur Anpassung des Datenschutzrechts an die Verordnung (UE) 2016/679 und zur Umsetzung der Richtlinie (UE) 2016/680 (Datenschutz-Anpassungs- und -Umsetzungsgesetz UE – DSAnpUG-UE).

sona que causó el daño y la persona que lo sufrió, bajo la condición de que esta última gane la demanda[17]. Por otro lado, la ley alemana también permite la decisión automatizada en servicios de seguros sanitarios privados cuando la decisión se base en normas vinculantes relativas a la remuneración por tratamiento médico (*ibid*.). Además, la ley administrativa alemana permite, asimismo, la adopción automatizada de actos administrativos en el marco de los procedimientos administrativos totalmente automatizados[18].

(17)Véase *ibid*., pág.106: explicaciones a § 37 (Automatisierte Entscheidungen im Einzelfall einschließlich Profiling).

(18)Véase § 35a Verwaltungsverfahrensgesetz (VwVfG) y explicación a § 37 de DSAnpUG-UE *ut supra*.

En tercer lugar, el RGPD permite de igual modo la toma de decisión automatizada si esta se basa en el consentimiento explícito del interesado. En algunos casos, notablemente en relación con decisiones basadas en la elaboración de perfiles, donde el interesado tiene que dar su consentimiento *online*, puede resultar problemático determinar si el consentimiento obtenido *online* era verdaderamente explícito o no. Con frecuencia, los perfiles se elaboran sin que el interesado siquiera lo sepa.[19] Y si el interesado no dio su consentimiento explícito para la elaboración de perfiles, tampoco consintió una decisión tomada en base a dicha elaboración. Por ejemplo, el consentimiento explícito para *cookies* no significaría, necesariamente, el consentimiento para una decisión automatizada basada en el perfil que resulte de ellas. Aunque el RGPD permite en sí la elaboración de perfiles siempre que se respeten los requisitos en él establecidos[20], las decisiones basadas en elaboraciones de perfiles deberían ser conformes a determinadas salvaguardas.

(19)Grupo de Trabajo del Artículo 29, «Advice Paper on Essential Elements of a Definition and a Provision on Profiling within the EU General Data Protection Regulation», adoptado el 13 de mayo de 2013, <http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2013/20130513_advice-paper-on-profiling_en.pdf> (acceso: 11 de noviembre de 2016).

(20)Véase el considerando 72 del RGPD en virtud del cual «la elaboración de perfiles está sujeta a las normas del presente Reglamento que rigen el tratamiento de datos personales, como los fundamentos jurídicos del tratamiento o los principios de la protección de datos [...]".».

Salvaguardas adicionales en diferentes países

La mayoría de los países que específicamente permiten la elaboración de perfiles exigen unas salvaguardas adicionales a este respecto. Italia puede utilizarse como ejemplo de estos países, pero el interesado debe recibir una notificación con anterioridad al tratamiento de los datos para los perfiles. Véanse las directrices sobre elaboración de perfiles en línea emitidas por el Garante per la protezione dei dati personali (resumen disponible en <http://blogs.dlapiper.com/iptitaly/?p=56970>. Acceso: 26 de mayo de 2017). Además, algunos Estados, como los Países Bajos, incluso permiten una elaboración de perfiles étnica, lo cual puede ser problemático desde la perspectiva de la protección de datos y de la no discriminación. Para más información sobre este tema, véase Simone Vromen, «Ethnic Profiling in the Netherlands and England and Wales: Compliance with International and European Standards». Public Interest Litigation Project (PILP-NJCM). Utrecht University <https://pilpnjcm.nl/wp-content/uploads/2015/06/Research-project-B-FINAL-version.pdf>. Acceso: 16 de mayo de 2017.

1.4. Salvaguardas en la toma de decisiones automatizada: de revelar la lógica tras la decisión a la rendición de cuentas

1.4.1. Salvaguardas en el artículo 22 del RGPD

En todos aquellos casos en los que se permita una decisión automatizada, el interesado debe contar con las salvaguardas apropiadas para evitar una decisión perjudicial o discriminatoria que no respete sus derechos e intereses. Siempre que se autorice una decisión automatizada con arreglo a la legislación de la

Unión o de un Estado miembro, dicha normativa también tiene que ofrecer «medidas adecuadas para salvaguardar los derechos, las libertades y los intereses legítimos del interesado» (artículo 22(2)(b)). En otros dos ejemplos, firma de un contrato y consentimiento explícito, el RGPD también requiere dichas salvaguardas, pero clarifica qué medidas mínimas deben ofrecerse para que el interesado tenga derecho, por lo menos, a:

- Obtener intervención humana por parte del controlador
- Expresar su punto de vista.
- Impugnar la decisión.

El interesado siempre tiene derecho a obtener una intervención humana, es decir, a que la decisión totalmente automatizada se convierta en no automatizada gracias a la intervención de una persona. Por ejemplo, en un contrato de seguro, la evaluación de riesgo se realiza por medios automatizados, pero la pesona valora los resultados y toma la decisión final. En algunas ocasiones, puede resultar difícil ejercer este derecho en la práctica. En el caso de un individuo que firma un contrato *online* con precios dinámicos, ¿cómo puede solicitar intervención humana si la página web no ofrece tal posibilidad? Es más, el individuo también tiene derecho a expresar su punto de vista, aunque el RGPD no especifica qué consecuencia legal se derivaría si se expresara dicha opinión. Y, por último, el individuo tiene derecho a impugnar la decisión. En la práctica, esto significa que el procedimiento se convierte en contradictorio y, por consiguiente, es cuestionable quién debería decidir sobre dicha objeción del individuo. Si, por ejemplo, un individuo diera su consentimiento explícito a una valoración automatizada de su solvencia y después se opusiera a dicha decisión, ¿debería esta objeción ser tratada por el empleado de banca que gestiona su cuenta, por otro empleado de la organización o por un organismo independiente?

1.4.2. ¿La existencia del derecho a la explicación?

En el caso de decisiones automatizadas que implican datos personales del interesado, el RGPD obliga al responsable del tratamiento a ofrecer al individuo «información significativa sobre la lógica aplicada» en dicha toma de decisión, con independencia de que los datos se recojan a partir del interesado[21] o no (deber de notificación del responsable del tratamiento[22]). Además, en el marco del derecho al acceso, el RGPD prevé un derecho similar del individuo a recibir no solo información sobre la existencia de una toma de decisión automatizada, sino también «información significativa sobre la lógica aplicada, así como la importancia de las consecuencias prevists de dicho procesamiento para el interesado»[23]. Estas disposiciones encajan bien dentro del marco más amplio que constituye la búsqueda del RGPD de un nivel más elevado de transparencia, lo cual requiere que el procesamiento de los datos persona-

(21) Artículo 13(2)(f) RGPD

(22) Artículo 14(2)(g) RGPD

(23) Artículo 15(1)(h) RGPD.

(24) Considerando 39 RGPD.

(25) Considerando 58 RGPD.

(26) Considerando 60 RGPD.

les sea transparente para las personas físicas cuyos datos se «recogen, usan, consultan o tratan de cualquier otra forma»[24]. El principio de transparencia del tratamiento de datos, consagrado en el artículo 5(1)(a) del RGPD, exige no solo que la información para el individuo sea «concisa, fácilmente accesible y fácil de entender»[25], sino que también se informe al sujeto «de la existencia de la operación de tratamiento y de sus fines»[26]. Dada la circunstancia de que la transparencia en el RGPD hace referencia al individuo y no a la sociedad en su conjunto, se puede entender como «transparencia individual» puesto que, en principio, ofrece al sujeto derecho a acceso, explicación y comprensión de las razones subyacentes a una decisión en caso de un tratamiento automatizado. El Supervisor Europeo de Protección de Datos (EDPS, por sus siglas en inglés) señala acertadamente que no les corresponde a los individuos buscar la explicación a esa lógica, sino que las organizaciones tienen que buscar proactivamente dicha transparencia.[27]

(27)European Data Protection Supervisor, «Opinion 7/2015. Meeting the challenges of big data. A call for transparency, user control, data protection by design and accountability» < https://secure.edps.europa.eu/edpsweb/webdav/site/mysite /shared/documents/ consultation/opinions/2015/15-11-19_big_data_en.pdf>. Acceso: 15 de noviembre de 2016.

No obstante, esta búsqueda de la transparencia plantea diversas preguntas: ¿qué se tiene que revelar exactamente al interesado?, ¿revelar la lógica significativa implica que el interesado tiene derecho a una explicación de la decisión automatizada? De ser así, ¿cuán detallada ha de ser la explicación? Cabe subrayar que el derecho explícito a una explicación no se menciona ni en el artículo 22 del RGPD ni en los artículos 13 y 14 relativos a las obligaciones de información, que dan derecho al interesado a obtener información significativa sobre la lógica aplicada. La única parte donde se menciona el derecho a la explicación en el RGPD es el considerando 71, en virtud del cual el tratamiento según el artículo 22:

> «En cualquier caso, dicho tratamiento debe estar sujeto a las garantías apropiadas, entre las que se deben incluir la información específica al interesado y el *derecho a obtener intervención humana*, a expresar su punto de vista, a *recibir una explicación de la decisión tomada después de tal evaluación* y a impugnar la decisión.» (Cursiva añadida.)

En la doctrina existe un intenso debate sobre si realmente se debe otorgar al interesado ese derecho a la explicación. Fueron Goodman y Flaxman quienes lo iniciaron al inferir ese derecho del requisito de dar al interesado información significativa sobre la lógica aplicada (artículos 13 y 14) (Goodman, 2017). Wachter *et al*. afirman que el RGPD solo exige una explicación *ex ante* del funcionamiento del sistema y no una explicación *ex post* de las razones subyacentes de la decisión (véase Wachter, 2017, sobre una postura respecto a la no existencia en el RGPD del derecho a la explicación de una decisión automatizada). Edwards y Veale aceptan la posibilidad del derecho a la explicación, pero señalan dificultades prácticas en su ejercicio desde la perspectiva de los algoritmos de aprendizaje automático (Edwards). Asimismo, Mendoza y Bygrave plantean argumentos a favor del derecho a la explicación (Mendoza, pág. 16).

Referencias bibliográficas

Bryce Goodman, **Seth Flaxman** (2017). *European Union Regulations on Algorithmic Decision-making and a "Right to Explanation"* <https://arxiv.org/pdf/1606.08813v3.pdf>. Acceso: 18 de julio de 2017.

Sandra Wachter, Brent Mittelstadt, Luciano Floridi (2017). *Why a Right to Explanation of Automated Decision-making Does not Exist in the General Data Protection Regulation* <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2903469>. Acceso: 18 de mayo de 2017.

Lilian Edwards, Michael Veale *(2017). Slave to the Algorithm? Why a "Right to an Explanation" is probably not the Remedy You Are Looking For* <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2972855>.

Se sostiene que las disposiciones del RGPD deben interpretarse de manera que se dote al interesado del derecho a la explicación y que el Tribunal de Justicia (TJ) debe seguir este enfoque en sus decisiones sobre este asunto. La información sobre la lógica aplicada tiene que permitir al interesado poder expresar su punto de vista y oponerse a la decisión automatizada[28]. Esta información tendría que ir más allá de la que se tiene que ofrecer al interesado en todos los casos de tratamiento de datos, como por ejemplo la identidad de su responsable o los fines con los que se procesan los datos personales[29]. Por tanto, entendemos que la información significativa sobre la lógica aplicada debería comprender, idealmente:

(28)Artículo 22(3) RGPD.

(29)Véanse los artículos 13 y 14 RGPD.

a) Información sobre los datos que sirvieron como elementos de la decisión automatizada.

b) Información sobre la lista de factores que influyeron en la decisión.

c) Información sobre la importancia relativa de los factores que influyeron en la decisión.

d) Explicación razonable sobre *porqué* se tomó una determinada decisión (información textual).

En realidad, y dados los numerosos obstáculos para los puntos b) y c) que veremos más adelante, el derecho a la explicación probablemente englobaría «solo» información textual que explicara las razones cruciales de las decisiones. Se pueden exponer varios argumentos a favor de tal derecho a la explicación.

En primer lugar, la aproximación metodológica que debería utilizarse a este respecto es interpretar varias disposiciones del RGPD de forma conjunta. En esta línea, las disposiciones del RGPD, en concreto el artículo 22, leído en el sentido del considerando 71, en combinación con los artículos 13(2)(f), 14(2)(g) y 15(1)(h) del RGPD, deberían interpretarse de forma que otorgaran al interesado el derecho a una explicación *ex post* de la decisión automatizada. Esta agrupación metodológica de las diferentes disposiciones relativas a la protección de datos para crear un derecho determinado del individuo no es inusual en la jurisprudencia del TJ. Por ejemplo, en el caso Google España, el tribunal se basó en la combinación del derecho de acceso y del derecho de oposición de la Directiva 95/46[30] para construir judicialmente el derecho al borrado (popularmente descrito como «derecho al olvido»[31]). Contrariamente a la aproxi-

(30)Directiva 95/46/CE del Parlamento Europeo y del Consejo, de 24 octubre de 1995, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos [1995] DOUE n.º 281/31.

(31)Más precisamente, el TJ se basó en el artículo 12(b) y el subpárrafo (a) del primer párrafo del artículo 14 de la Directiva 95/46; Caso C-131/12 Google España y Google ECLI:UE:C:2014:317.

mación de Wachter *et al*., quienes analizan estas disposiciones por separado, aquí se sostiene que el TJ, al interpretar las disposiciones del RGPD, debería leerlas conjuntamente si pretende construir el derecho a la explicación.

En segundo lugar, ante la ausencia de un derecho a la explicación del interesado, su derecho a impugnar la decisión adoptada por medios automatizados sería totalmente inefectivo (véase también Mendoza, pág. 16). Si el interesado quiere oponerse a dicha decisión de forma sustantiva, debe obtener información, por lo menos, sobre los datos que se utilizaron como elementos para la decisión automatizada, así como una explicación razonable de los fundamentos de dicha decisión. El derecho a impugnar hace referencia a la sustancia de la decisión y estaría vacío de contenido si el interesado se enfrentara tan solo a una decisión final sin ninguna explicación respecto a la misma.

En tercer lugar, los que se oponen al derecho a la explicación se basan en el argumento según el cual este derecho no aparece expresamente en el texto del RGPD y añaden que el considerando 71 del reglamento, que contiene ese derecho, no es jurídicamente vinculante (Wachter, pág. 4). Si bien el procedimiento legislativo muestra que el derecho a la explicación, en efecto, se omitió en el texto del artículo 22 (Wachter, pág. 9), la circunstancia de que siga apareciendo en el considerando demuestra que el legislador no quería eliminarlo del todo. Colocarlo en el considerando era una solución de compromiso que partió de un desacuerdo sobre si este derecho debía quedar ratificado en el RGPD o no (compárese Edwards, pág. 33). Esta redacción de la norma también demuestra que el legislador dejó la decisión final sobre la existencia de este derecho en manos del TJ quien, como se ha demostrado repetidamente en la jurisprudencia reciente, es bastante intencionado y activista a la hora de interpretar la legislación relativa a la protección de datos. Desestimar totalmente la posibilidad de que exista el derecho a la explicación porque los considerandos no son jurídicamente vinculantes es una postura demasiado formalista, en especial a la luz de la jurisprudencia del TJ, quien utiliza habitualmente los considerandos como una ayuda interpretativa[32]. Si se observa con mayor atención, veremos que la jurisprudencia de este tribunal revela que, efectivamente, «el preámbulo de una ley de la Unión Europea no tiene fuerza vinculante» y no se puede utilizar para derogar disposiciones existentes o «para interpretar dichas disposiciones de una forma claramente contraria a su redactado»[33]. Así pues, la jurisprudencia marca que los considerandos no se pueden utilizar para una interpretación *contra legem* de la legislación de la UE. Sin embargo, el basarse en un enfoque que interpreta varias disposiciones del RGPD, como se ha sugerido más arriba, juntamente con el considerando 71 para reforzar la interpretación que apoya la existencia del derecho a la explicación, no conduciría a una interpretación *contra legem*. Por el contrario, sería un medio para resolver la ambigüedad que emerge de una lectura normal de las disposiciones relevantes del RGPD[34]. Teniendo en cuenta la postura activista y los esfuerzos

(32)Véase, por ejemplo: Caso C-283/16 M.S.ECLI:UE:C:2017:104, párrafos 34-35; Caso C-436/16 Leventis y Vafias ECLI:UE:C:2017:497, párrafo 33; Caso C-578/16 PPU C.K., H.F., Un.S.ECLI:UE:C:2017:127, párrafo 43; Caso C-111/17 PPU OL ECLI:UE:C:2017:436, párrafo 40.

(33)Caso C-308/97 Manfredi ECLI:UE:C:1998:566, párrafo 30; Caso C-136/04 Deutsches Milch-Kontor ECLI:UE:C:2005:716, párrafo 32; Caso C-134/08 Tyson Parketthandel ECLI:UE:C:2009:229, párrafo 16; Caso C-7/11 Caronna ECLI:UE:C:2012:396, párrafo 40; Caso C-345/13 Karen Millen Modas ECLI:UE:C:2014:2013, párrafo 31.

(34)Sobre la función de los considerandos para resolver la ambigüedad de las disposiciones legislativas, véase Tadas Klimas, Jurate Vaiciukaitè (2008), «The Law of Recitals in European Community Legislation». ILSA Journal of International & Comparative Law (n.º 15, pág. 26).

del TJ para asegurar un nivel elevado de protección de datos, es posible que el Tribunal optara por una interpretación que otorgara al interesado el derecho a la explicación.

En cuarto lugar, también es importante subrayar las sutilezas del redactado de las disposiciones relevantes. Una lectura detallada del artículo 22(3) revela que las salvaguardas establecidas en esta disposición (intervención humana, expresión del propio punto de vista, impugnación) no son necesariamente las únicas posibles. En concreto, esta disposición requiere que al interesado se le garanticen, *por lo menos*, esas salvaguardas. Por lo tanto, añadir una adicional, el derecho a la explicación del considerando 71, mediante la interpretación judicial no supondría claramente una interpretación que derogase esta disposición o que fuera contraria a su redactado. Es más, también se ha afirmado que los artículos 13(2)(f) y 14(2)(g) del RGPD omiten referirse directamente al artículo 22(3) del RGPD relativo a las salvaguardas y solo se refieren al artículo 22(1) y (4) (Wachter, pág. 13). Se puede afirmar lo mismo del artículo 15(1)(h). Sin embargo, el redactado de los primeros exige que se dé al interesado información significativa sobre la lógica aplicada en la decisión automatizada, *por lo menos*, en el caso del artículo 22(1) y (4). Parece que se dejó un redactado explícitamente abierto para no imposibilitar del todo una interpretación judicial que otorgue al interesado el derecho a una explicación.

En quinto lugar, aunque no se reconozca un derecho general a la explicación, este debería existir, al menos, cuando la decisión automatizada se basa en datos sensibles. En principio, las decisiones automatizadas no deberían apoyarse en categorías especiales de los datos personales (artículo 22(4)), excepto si el individuo da su consentimiento explícito al tratamiento con fines específicos o si dicho tratamiento es necesario para salvaguardar un interés público importante (artículo 9(2)(a) y (g)). El artículo 22(4) es bastante impreciso en lo relativo a las salvaguardas y declara que se deben aplicar «medidas adecuadas para salvaguardar los derechos y las libertades del interesado». En cambio, todas las disposiciones del RGPD que exigen que el interesado esté familiarizado con la lógica aplicada (artículos 13(2)(f), 14(2)(g) y 15(1)(h)), requieren expresamente que esto se garantice en caso de que la decisión automatizada se base en datos sensibles. Lo que debe tenerse en cuenta, asimismo, no es solo la interpretación literal de estas disposiciones, sino también el objetivo de un nivel elevado de protección cuando se trata de datos sensibles. Si la inclusión de este tipo de datos conduce a una decisión sesgada, el interesado tendría que poder entender las razones que subyacen a dicha decisión. Sus derechos no están suficientemente garantizados si solo recibe información general sobre el funcionamiento del sistema.

Para cumplir con el requisito del RGPD según el cual la lógica de la decisión tiene que explicarse al interesado, no basta con asegurar tan solo lo que Kroll *et al*. denominan «regularidad procedimental». Este principio solo garantiza que las decisiones se basen en la misma política de decisiones, la cual se definió antes de que los elementos de entrada y los resultados se pudieran reproducir (Kroll *et al*., 2017). Por lo tanto, solo se dirige a la regularidad procedimental agregada de todos los casos, salvaguardando que todos los casos se decidan con arreglo a las mismas normas. Sin embargo, el concepto de regularidad procedimental no responde la pregunta de porqué el algoritmo llegó a una determinada decisión tomando para ello un conjunto concreto de datos. La transparencia que requiere el RGPD tiene una naturaleza distinta según la cual el interesado debe entender las razones subyacentes de la decisión.

Referencias bibliográficas

Joshua A. Kroll *et al*. *(2017). Accountable Algorithms*. *University of Pennsylvania Law Review* (n.º 165, 1, pág. 18) [en línea] <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2765268>.

Anupam Datta, Sen Shayak, Yair Zick (2016). *Algorithmic Transparency via Quantitative Input Influence* [en línea] <https://www.andrew.cmu.edu/user/danupam/datta-sen-zick-oakland16.pdf>. Acceso: 10 de diciembre de 2016.

La transparencia individual relativa a las decisiones automatizadas *no algorítmicas* no planteará problemas concretos en relación con la explicación de la lógica tras la decisión. Por ejemplo, si una cámara que detecta la velocidad de un conductor comunica a las autoridades públicas que se excedió el límite de velocidad, automáticamente se genera la multa. La lógica subyacente a esta decisión, así como la norma en la cual se basa, se pueden explicar fácilmente al interesado: la multa se emite si se excede el límite de velocidad (teniendo en cuenta el factor de corrección aplicable).

A diferencia de lo anterior, la toma de decisiones automatizadas basadas en *algoritmos* se enfrentan a numerosas complicaciones cuando se tienen que explicar las razones subyacentes. A medida que avanza la tecnología y que el uso de algoritmos para tomar decisiones crece exponencialmente, tanto la normativa jurídica como el sector académico reclaman decisiones algorítmicas más transparentes, con frecuencia descritas con la expresión de moda «transparencia algorítmica». La búsqueda básica de los que proponen la transparencia algorítmica es revelar la lógica que hay tras el algoritmo que adopta una decisión determinada. Aunque algunos expertos consideran que es casi imposible explicar un algoritmo porque ni siquiera sus desarrolladores pueden señalar claramente las razones por las que se adoptó una decisión en concreto[35], otros se sitúan en un enfoque más optimista (Kroll, 2017) e incluso proponen soluciones técnicas (Datta, 2016) que resultarían en una mayor transparencia algorítmica. Se considera que este concepto debe cubrir diferentes grados de transparencia, desde revelar un código fuente hasta la explicación de su funcionamiento. Para los fines de este trabajo, creemos que la transparencia algorítmica, jurídicamente hablando, debe englobar la transparencia del proceso de toma de decisión algorítmica hasta el punto que sea necesario para garantizar el respeto de los derechos según el RGPD, especialmente proporcionar al sujeto afectado información relevante de la lógica aplicada. En términos técnicos, el grado de revelación del funcionamiento del algoritmo puede diferir para distintas decisiones.

(35)Véase por ejemplo <https://www.wired.com/2016/07/artificial-intelligence-setting-internet-huge-clash-europe/> y <https://www.wired.com/2016/07/artificial-intelligence-setting-internet-huge-clash-europe/>

Referencias bibliográficas

Joshua A. Kroll *et al*. *(2017). Accountable Algorithms*. *University of Pennsylvania Law Review* (n.º 165, 1, págs. 15-16) [en línea] <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2765268>.

Anupam Datta, Shayak Sen, Yair Zick (2016). *Algorithmic Transparency via Quantitative Input Influence* [en línea] <https://www.andrew.cmu.edu/user/danupam/datta-sen-zick-oakland16.pdf>. Acceso: 10 de diciembre de 2016.

Además, la transparencia algorítmica es una herramienta necesaria para la prevención de la discriminación en las decisiones algorítmicas, que en ocasiones son discriminatorias, incluso cuando el algoritmo no está programado para discriminar. Una decisión puede ser discriminatoria, por ejemplo, porque los datos en los que se basa son en sí mismos discriminatorios. Esto se puede dar, en concreto, cuando la decisión implica datos sensibles tales como la raza o el género. A este respecto, Zarsky señala que «un conjunto de datos desviado y sesgado puede dar lugar a resultados del proceso algorítmico que discriminen a los grupos protegidos» (Zarsky, 2016). La decisión también puede ser sesgada si el algoritmo es entrenado con datos sesgados (Goodman, 2016). Sin embargo, otros autores afirman que, para evitar la discriminación algorítmica, es necesario utilizar datos personales sensibles en el proceso de construcción de los modelos de toma de decisión (Žliobaité, 2016). En cualquier caso, cuando el algoritmo adopta la decisión, los datos personales sensibles, como la raza, no deberían considerarse como «variables de entrada» relevantes para la toma de decisión (*ibid*.). Además de la discriminación causada por conjuntos de datos de entrada sesgados, la decisión algorítmica también puede ser discriminatoria debido a la programación sesgada del algoritmo, que produce decisiones discriminatorias. En todo caso, la transparencia algorítmica ayudaría a comprender las razones que subyacen en decisiones sesgadas y, en consecuencia, sería el primer paso para prevenir dicha discriminación algorítmica.

Referencias bibliográficas

Tal Zarsky (2016). «The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making». *Science, Technology & Human Values* (n.º 41, pág. 126).

Bryce Goodman (2016). «Discrimination, Data Sanitisation and Auditing in the European Union's General Data Protection Regulation». *European Data Protection Law Review* (vol. 2, n.º 4, pág. 498).

Indré Žliobaité, Bart Custers (2016). «Using Sensitive Personal Data May Be Necessary for Avoiding Discrimination in Data-driven Decision Models». *Artif Intell Law* (n.º 24, págs. 183–201).

Finalmente, además de proporcionar al interesado las razones de la decisión que le atañe y aparte de evitar la discriminación, la transparencia algorítmica también tiene un papel importante en determinar la distribución de la responsabilidad en la toma de decisiones automatizada (en ocasiones también denominada «responsabilidad algorítmica»). Lógicamente hablando, la transparencia algorítmica es una predisposición para la responsabilidad algorítmica. ¿Quién sería responsable si un algoritmo se equivoca o adopta una decisión discriminatoria, los desarrolladores, el usuario o incluso el propio agente autónomo? ¿Quién es responsable si un motor de búsqueda utiliza un algoritmo que favorece a un partido político en concreto en lugar de ser políticamente neutro, o si muestra resultados relativos a determinadas empresas por encima de otras y busca desfavorecer a la competencia? Aunque la doctrina no trata específicamente el problema de la responsabilidad *algorítmica*, se puede buscar inspiración en la literatura que estudia la de los *robots* cuando actúan como agentes autónomos. Las posibilidades de atribución de responsabilidades van desde la responsabilidad objetiva del programador y la responsabilidad del usuario a la responsabilidad del propio agente autónomo (Palmerini, 2016). Un análisis exhaustivo de las cuestiones relativas a la responsabilidad exceden el objetivo de este módulo. Sin embargo sostenemos que por un lado, la responsabilidad debería ser compartida entre el programador y el usuario del algoritmo, siempre que el algoritmo tenga características de autoaprendizaje. Por el otro lado, tanto la responsabilidad objetiva como la plena responsabilidad del agente autónomo deberían ser rechazadas. Existe un argumento de peso contra la responsabilidad objetiva por las acciones de un algoritmo que se basa en que es bastante improbable poder considerar un algoritmo como

Referencias bibliográficas

Erica Palmerini, Andrea Bertolini (2016). *Liability and Risk Management in Robotics*. En: Reiner Schulze, Dirk Staudenmayer (eds.) (2016). *Digital Revolution: Challenges for Contract Law in Practice* (págs. 225–260). Nomos.

Sobre la responsabilidad de los agentes autónomos, véase también:

Jape Hage (2016). *Should Autonomous Agents be Liable for What They Do?* [en línea] <http://www.jaaphage.nl/pdf/LiableAutonomousAgents.pdf>. Acceso: 18 de mayo de 2017.

Erica Palmerini, Andrea Bertolini (2016). *Liability and Risk Management in Robotics*. En: Reiner Schulze, Dirk Staudenmayer (eds.). *Digital Revolution: Challenges for Contract Law in Practice* (pág. 240). Nomos.

un producto peligroso (véase, por analogía respecto a los robots, Palmerini, 2016). Es más, afirmamos que la responsabilidad exclusiva de los agentes autónomos podría llevar a evitar toda responsabilidad y, por lo tanto, no debería ser defendida. Esto es especialmente cierto respecto a una potencial responsabilidad penal que es imposible atribuir a un algoritmo. En lo que respecta a la responsabilidad civil, es cierto que es factible crear un sistema de seguro para cubrir los daños derivados de la responsabilidad del agente autónomo. No obstante, en derecho civil, la responsabilidad requeriría, en principio, culpa o negligencia por parte del agente autónomo, lo cual sería imposible de establecer. Por consiguiente, la responsabilidad de los propios agentes autónomos tal vez sería una herramienta para eximir a los programadores o a los usuarios de la responsabilidad que se les debería atribuir.

1.5. Obstáculos a la transparencia algorítmica

Se pone de relieve que existen varios obstáculos para proporcionar al interesado una explicación razonable de la lógica subyacente a las decisiones algorítmicas (Burell, 2016), entre los que cabe destacar los siguientes:

- Obstáculos técnicos.
- Obstáculos de propiedad intelectual.
- Secretos de Estado y otra información confidencial de autoridades estatales.

Referencia bibliográfica

Burell distingue entre tres tipos de opacidad de algoritmos: la que surge de las características de aprendizaje automático, el secreto corporativo o de Estado y el analfabetismo técnico. Véase:

Jenna Burrell (2016). «How the Machine "thinks": Understanding Opacity in Machine Learning Algorithms». *Big Data & Society* (n.º 1, pág. 12).

1.5.1. Obstáculos técnicos

Referencia bibliográfica

Stuart J. Russell, Peter Norvig (eds.) (2010). *Artificial Intelligence. A Modern Approach*. Pearson (3.ª edición, pág. 638).

La cantidad de obstáculos técnicos que dificultan la explicación de decisiones autónomas basadas en algoritmos depende de la complejidad de estos. Numerosos autores afirman que es casi imposible explicar la lógica que hay tras un algoritmo que adopta una decisión. Las razones de una decisión basada en un simple árbol de decisiones tal vez podrían explicarse porque estas estructuras son una forma de razonamiento o de apoyo a decisiones que utilizan un gráfico similar a un árbol para llegar a una conclusión y fueron la principal herramienta para adoptar decisiones hasta los años 1980 aunque, en la actualidad, todavía se utilizan (Stuart, 2010). Un árbol de decisiones llega a una conclusión «al realizar una secuencia de pruebas» (*ibid.*, pág. 698). En un árbol de decisiones sencillo, cada «nódulo de decisión» representa un punto de inflexión, una pregunta para la cual hay diferentes respuestas posibles. Al final de estos puntos de inflexión hay «nódulos de hojas» que representan posibles respuestas distintas a la pregunta contenida en el «nódulo raíz» (para una explicación matemática de los árboles de decisión, véase *ibid.*, pág. 698 y sigs.). Este proceso hace que los árboles adopten decisiones lineales que son fáciles de analizar *ex post* y que permiten la transparencia algorítmica. Por ejemplo, un árbol de decisión sencillo podría proporcionar ayuda en una decisión con

determinación, en función de si la persona que solicita un empleo cumple los requisitos formales. Los nódulos de decisión podrían contener preguntas como ¿la persona ha completado sus estudios a un nivel requerido?, ¿tiene esta persona la experiencia profesional necesaria?, ¿tiene recomendación de su empleador anterior? En función de si las respuestas a estas preguntas son afirmativas o negativas, los nódulos de hojas podrán facilitar respuestas afirmativas o negativas a la pregunta inicial de si se cumplen los requisitos para el puesto de trabajo.

Aun así, si el algoritmo utilizado para tomar la decisión es una red neuronal, proclive a un aprendizaje muy rápido (Masnick, 2016), resultará casi imposible explicar las decisiones que haya detrás de dicha decisión (Metz, 2016; Goodman, 2016). Esencialmente, las redes neuronales funcionan de manera diferente a los algoritmos sencillos, puesto que están construidas siguiendo el modelo del cerebro humano, donde los diferentes «nódulos» se conectan entre sí en una red (Russell, 2010). Gracias a esta habilidad de interconexión y al hecho de que reproduce el cerebro humano, la red neuronal puede aprender mientras procesa datos. Según Russell y Norvig, «las redes neuronales siguen siendo una de las formas más populares y efectivas de sistema de aprendizaje» (*ibid*., pág. 728). Debido a su capacidad de aprendizaje, es casi imposible garantizar su transparencia e identificar los factores que influyen en una decisión tomada por una red neuronal. Sin embargo, Datta *et al*. desarrollaron un sistema denominado quantitative input influence (QII, 'influencia de entrada cuantitativa') que podría explicar las decisiones tomadas autónomamente. La idea tras el QII es que se podría medir el grado de influencia de los datos de entrada sobre los de salida (*ibid*.). Al parecer, para alcanzar la transparencia de un algoritmo se tendría que desarrollar otro para clarificar qué factores se tuvieron en cuenta y qué peso tenían[36].

[36]Compárese: documento de sesión para la 38.ª Conferencia Internacional de Protección de datos y Comisarios de Intimidad, octubre de 2016. *Artificial Intelligence, Robotics, Privacy and Data Protection.*

Referencias bibliográficas

Masnick afirma que cuanto más rápido aprenden las máquinas, más difícil es entender las razones que hay tras sus decisiones.

Mike Masnick (2016). *Activists Cheer On EU's Right To An Explanation For Algorithmic Decisions, But How Will It Work When There's Nothing To Explain?* [en línea] <https://www.techdirt.com/articles/20160708/11040034922/activists-cheer-eus-right-to-explanation-algorithmic-decisions-how-will-it-work-when-theres-nothing-to-explain.shtml>. Acceso: 10 de enero de 2016.

Metz señala que «las redes neuronales profundas dependen, en gran medida, de los datos y generan algoritmos complejos que pueden resultar opacos incluso para aquellos que pusieron en marcha dichos sistemas». Véase:

Cade Metz (2016). *Artificial Intelligence Is Setting Up the Internet for a Huge Clash With Europe* [en línea] <https://www.wired.com/2016/07/artificial-intelligence-setting-internet-huge-clash-europe/https://www.wired.com/2016/07/artificial-intelligence-setting-internet-huge-clash-europe/>. Acceso: 10 de enero de 2016.

Compárese también:

Bryce Goodman, **Seth Flaxman** (2016). *European Union Regulations on Algorithmic Decision-making and a "Right to Explanation"* [en línea] <https://arxiv.org/

abs/1606.08813v3https://arxiv.org/abs/1606.08813v3>. Acceso: 1 de septiembre de 2016.

Para una explicación matemática, véase:

Stuart J. Russell, **Peter Norvig** (eds.) (2010). *Artificial Intelligence. A Modern Approach. Pearson* (3.ª ed., pág. 728 y sigs.).

1.5.2. Obstáculos relacionados con la propiedad intelectual, ¿un tigre de papel?

Hasta cierto punto, los derechos de propiedad intelectual también pueden obstaculizar la transparencia algorítmica. Sostenemos que no ocurre así en relación con las patentes y el *copyright*. De un modo distinto, los secretos comerciales o la información confidencial sí pueden interponerse a la transparencia algorítmica.

Aunque la Convención Europea de Patentes (EPO, European Patent Office) permite patentar «invenciones implantadas por ordenador»[37], no basta con que el software sea creativo, también debe permitir una aplicación industrial[38]. Junto a esta línea de razonamiento, la EPO no permite patentar un algoritmo informático, puesto que el «programador tiene que haber tenido consideraciones técnicas más allá de encontrar "simplemente" un algoritmo informático»[39]. Aun así, incluso si el algoritmo estaba sujeto a una patente, ello no plantearía un obstáculo para la transparencia algorítmica, puesto que tener una patente obligaría a quien la ostenta a dar a conocer la composición y las modalidades de funcionamiento del algoritmo.

La protección que otorga el derecho de autor a un programa informático lleva a un resultado similar: si bien tanto el acuerdo ADPIC[40] como el Tratado relativo al derecho de autor de la OMPI[41] (WTC) otorgan la protección del derecho de autor al programa desde el momento de su creación, no queda del todo claro si los algoritmos por sí mismos pueden ser objeto de protección del derecho de autor. Parece ser que la UE no permite tal protección y esta aproximación está en línea con algunas otras jurisdicciones no europeas (por ejemplo, en Japón no se permite la protección propia del derecho de autor a los algoritmos; véase Karjala, 1991). En cualquier caso, conviene mencionar que la directiva de la UE relativa a programas informáticos[42] permite que el usuario de un programa de ordenador «observe, estudie o pruebe el funcionamiento del programa para determinar las ideas y los principios que subyacen a cualquier elemento del programa». En la práctica, esto significa que un usuario de un programa puede determinar el funcionamiento del algoritmo y, si es técnicamente posible, revelar la importancia de factores concretos implicados en la toma de una decisión algorítmica. Ello conllevaría que, incluso cuando el programa y/o el algoritmo estuvieran protegidos por el derecho de autor, dicha protección no impediría la transparencia algorítmica.

(37)Véase el artículo 52(2)(c) de la Convención Europea de Patentes en combinación con las Directrices relativas al Examen, punto 3.6, Programas para ordenadores, disponibles en <https://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_ii_3_6.htm>. Acceso: 20 de diciembre de 2016.

(38)Véase el artículo 52(1) de la Convención Europea de Patentes.

(39)Véase la Opinión de EPO G 0003/08 (Programas para ordenadores) de 12 de mayo de 2010, ECLI:EP:BA:2010:G000308.20100512, punto 13.5.

(40)Acuerdo sobre los Aspectos de la Propiedad Intelectual relacionados con el Comercio. Véase su artículo 10(1), en virtud del cual «los programas informáticos, ya sean en código fuente u objeto, deben estar protegidos como obras literarias».

(41)Organización Mundial de la Propiedad Intelectual. Véase su artículo 4, por el que «los programas informáticos están protegidos como obras literarias según el significado de […] la Convención de Berna.»

(42)Directiva 2009/24/CE del Parlamento Europeo y del Consejo, de 23 de abril de 2009, sobre la protección jurídica de programas de ordenador, DOUE n.º 111, 5 de mayo de 2009, pág.16.

Referencias bibliográficas

Dennis S. Karjala (1991). «Japanese Courts Interpret the "Algorithm" Limitation on the Copyright Protection of Programs». *Jurimetrics Journal* (n.º 31, pág. 233).

En cuanto a los autores norteamericanos sobre este tema (así como sobre patentes relativas a algoritmos) véase:

Richard H. Stern (1995). «On Defining the Concept of Infringement of Intellectual Property Rights in Algorithms and Other Abstract Computer-Related Ideas». *AIPLA Quarterly Journal* (n.º 23, pág. 401).

John Swinson (1991). «Copyright or Patent or Both: An Algorithmic Approach to Computer Software Protection». *Harvard Journal of Law & Technology* (n.º 5, pág. 145).

En cambio, un derecho de PI que obstaculiza la transparencia algorítmica es un secreto comercial (información confidencial o reservada) (más concretamente, véase Malgieri, 2016, en relación al equilibrio entre secretos comerciales y datos personales). Según el ADPIC, el secreto comercial permite a las personas físicas y jurídicas evitar que la información se revele o sea utilizada por terceros de forma que vaya en contra de las prácticas comerciales honestas si dicha información es secreta, tiene valor comercial y se han tomado medidas para mantenerla secreta[43]. En la directiva europea sobre secretos comerciales[44], que sigue esta definición, la UE prevé una excepción que permite la suspensión de un secreto comercial «con el fin de proteger un interés legítimo reconocido por la Unión o por la legislación nacional»[45]. El hecho de explicar una decisión algorítmica a un interesado podría encajar en esta excepción, puesto que está previsto en el RGPD y busca proteger un interés legítimo.

[43]Véase el artículo 39(2) del ADPIC.

[44]Directiva (UE) 2016/943 del Parlamento Europeo y del Consejo, de 8 de junio de 2016, relativa a la protección de los conocimientos técnicos y la información empresarial no divulgados (secretos comerciales) contra su obtención, utilización y revelación ilícitas, DOUE n.º 157, 15 de junio de 2016, pág.1.

[45]Véase el artículo 5(d) de la directiva de secretos comerciales.

Referencia bibliográfica

Gianclaudio Malgiere (2016). «Trade Secrets v. Personal Data: a possible solution for balancing rights». *International Data Privacy Law* (vol. 6, n.º 2, págs. 102-116).

Aunque esa excepción no sea aplicable, consideramos que la transparencia algorítmica no va necesariamente en contra de los secretos comerciales. Para proporcionar dicha transparencia en el marco del RGPD, el código fuente o incluso la forma en que funciona el algoritmo no tienen porqué revelarse. La «lógica tras la decisión» que emana del RGPD apunta hacia una herramienta (argumentativa) que se despliega mediante un algoritmo sin necesidad de revelar totalmente dicha herramienta. Por analogía con el periodismo computacional, Diakopoulos señala que la transparencia algorítmica podría gobernar, tan solo, «la revelación de determinadas piezas clave de información, incluyendo los resultados agregados y los *benchmarks*» (Diakopoulos, 2016).

Referencia bibliográfica

Nicholas Diakopoulos (2016). «Accountability in Algorithmic Decision Making». *Communications of the ACM* (n.º 59, págs. 56, 58-59).

1.5.3. Secretos estatales y otra información confidencial

Los mayores obstáculos a la transparencia algorítmica los constituyen los secretos de Estado y otras informaciones en posesión de las autoridades públicas que no se pueden revelar al público. Es en interés del Estado y de dichas autoridades el motivo por el que no se revela exactamente las razones por las que se adoptan determinadas decisiones[46]. Por ejemplo, una autoridad fiscal no revelará el algoritmo que escoge a contribuyentes cuya (ir)regularidad fis-

[46]Por ejemplo, el 95% de las decisiones de la autoridad fiscal holandesa son automáticas y las autoridades aduaneras también se basan en gran medida en decisiones automatizadas. Doy las gracias al profesor Sjir Nijssen por esta perspectiva.

cal debe revisarse. Del mismo modo, las autoridades aduaneras no revelarán el sistema de concordancia de patrones que elige qué empresa debe pasar por control aduanero. Asimismo, la autoridad policial no dará a conocer la norma que marca la elección del barrio o de las personas que tienen que monitorizar, por ejemplo, con fines de prevención del terrorismo o del tráfico de drogas.

La cuestión sobre si el algoritmo se tiene que facilitar al interesado en estas situaciones depende de la ponderación entre el derecho a la privacidad y otros derechos concurrentes. Es obvio que los secretos de Estado o la información confidencial interferirán con la transparencia algorítmica, pero esta no es un valor último de nuestra sociedad que deba prevalecer siempre por encima de otros intereses. En dichos casos, la revelación de la lógica que hay tras la decisión automatizada dependerá de la ponderación que se lleve a cabo en cada caso en particular.

2. Inteligencia artificial: asuntos jurídicos restantes

La sociedad europea y la global son testigos de un avance tecnológico exponencial en el campo del *big data* y de la inteligencia artificial. En los últimos años, los desarrollos tecnológicos en el ámbito de la robótica y de los programas asociados a ella han sido testigos de un progreso inimaginable, desde los robots humanoides, autónomos y de atención personal, los vehículos autónomos, los robots de atención, los robots canguros y juguetes, los asistentes robóticos, hasta los agentes de IA utilizados para una labor policial predictiva, la diagnosis médica y otros campos. Los robots ayudan a las personas paralizadas a caminar y, en 2016, se llevó a cabo la primera intervención quirúrgica robotizada autónoma (Orphanides, 2017). Las últimas tendencias en la innovación robótica y en el sector tienen un potencial empresarial extraordinario, al igual que limitaciones éticas y jurídicas. El BD y la IA ocupan el primer lugar de la agenda de la UE para digitalizar la economía europea mediante el Mercado Único Digital,[47] y las instituciones de la UE están siendo pioneras en el establecimiento de directrices jurídicas y éticas claras para la IA. El Parlamento Europeo, con la aprobación de la Resolución sobre normas de derecho civil relativas a la robótica[48], busca formular estándares jurídicos y éticos para los robots. Además, la UE ya ha construido un marco jurídico sólido para la protección de datos y la ciberseguridad que podría aplicarse al BD y a la IA, incluído el RGPD, la Directiva de seguridad de redes[49] y la Propuesta de Reglamento de privacidad y comunicaciones electrónicas[50].

(47)Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. Una estrategia para el Mercado Único Digital de Europa (COM/2015/192 final).

(48)Resolución del Parlamento Europeo de 16 de febrero de 2017 con recomendaciones para la Comisión sobre normas de derecho civil relativas a la robótica (2015/2103(INL)).

(49)Directiva (UE) 2016/1148 del Parlamento Europeo y del Consejo, de 6 de julio de 2016, relativa a las medidas destinadas a garantizar un elevado nivel común de seguridad de las redes y los sistemas de información en la Unión (DOUE n.º 194, 19 de julio de 2016, pág.1).

(50)Propuesta de Reglamento del Parlamento Euroepo y del Consejo sobre el respeto de la vida privada y la protección de los datos personales en el sector de las comunicaciones electrónicas y por el que se deroga la Directiva 2002/58/CE (Reglamento sobre la privacidad y las comunicaciones electrónicas) (COM/2017/010 final).

Referencia bibliográfica

K. G. Orphanides (2017). *Robot Carries Out First Autonomous Soft Tissue Surgery* [en línea] <http://www.wired.co.uk/article/autonomous-robot-surgeonhttp://www.wired.co.uk/article/autonomous-robot-surgeon>. Acceso: 25 de julio de 2017.

Sin embargo, el papel crucial que la UE tiene en este ámbito plantea muchas cuestiones pendientes de resolver. Mientras las empresas europeas incrementan el uso del BD y la IA de forma exponencial, este enfoque no solo tiene que incluirse en un marco jurídico claro y conciso de la UE que proporcione privacidad (Calo, 2010), transparencia (Goodman, 2016) y responsabilidad (Bertolini, 2013), sino que también debe respetar normas éticas (Mittelstadt, 2016). La transformación en una auténtica «sociedad europea del BD y la IA» con un uso más amplio de la explotación del BD y de los agentes de IA solo es

posible si los usuarios confían en estas herramientas. Esto nos lleva de nuevo a cuestiones tecnológicas, económicas y éticas: si la tecnología se diseña de forma que genere confianza y que permita cumplir con los requisitos legales de transparencia y responsabiliad, respetando a la vez las exigencias éticas, su uso en las empresas aumentará, así como la competitividad en el mercado único digital de la UE.

Referencias bibliográficas

Ryan Calo (2010). «Peeping HALs: Making Sense of Artificial Intelligence and Privacy». *European Journal of Legal Studies* (n.º 2, pág. 3) <http://www.ejls.eu/6/83uk.htm>

Bryce Goodman, **Seth Flaxman** (2016). «European Union Regulations on Algorithmic Decision-making and a "Right to Explanation"». *ICML Workshop on Human Interpretability in Machine Learning*. Nueva York.

Andrea Bertolini (2013). «Robots as Products. The Case for a Realistic Analysis of Robotic Applications and Liability Rules». *LIT* (vol. 5, n.º 2, págs. 214–247).

Brent D. Mittelstadt, **Patrick Allo**, **Maria R. Taddeo**, **Sandra Wachter**, **Luciano Floridi** (2016). «The Ethics of Algorithms. Mapping the Debate». *Big Data & Society* (vol. 3, n.º 1, págs. 1–21).

Esta parte del mòdulo aborda brevemente los asuntos jurídicos restantes relativos a la inteligencia artificial, es decir, la cuestión de la responsabilidad (civil) por las acciones de los agentes autónomos dotados de IA, en particular los robots, y la personalidad jurídica de dichos agentes.

El uso creciente de los robots también plantea problemas jurídicos en relación con la rendición de cuentas y la responsabilidad civil. Es cuestionable que las normas generales relativas a la responsabilidad civil sean suficientes y/o apropiadas para aplicarlas a los robots. Uno de los mayores dilemas a este respecto es el tema de la responsabilidad civil de los vehículos autónomos, en especial los coches sin conductor, que tomaremos como ejemplo para ilustrar este asunto. Si un coche sin conductor provoca un accidente, la cuestión preliminar que hay que responder es si el coche estaba en modo totalmente automatizado y se conducía solo o si bien el conductor le asistía en la conducción. En la actualidad, hay seis niveles de conducción automatizada: no automatización, asistencia al conductor, automatización parcial, automatización condicional, automatización alta y automatización total (Pillath, 2016). Otras clasificaciones utilizan tres niveles: automatización parcial, media y alta (Schellekens, 2015). Es relevante hacer una clarificación previa del nivel de autonomía, por ejemplo, cuando el coche está preparado para el aprendizaje y cuando podría haber adoptado determinadas prácticas de conducción del propio conductor. La primera opción de regulación es adoptar la *responsabilidad del conductor*. En la primera generación de coches sin conductor, el conductor todavía tendrá una función activa de supervisión del coche cuando este está conduciendo. El conductor deberá estar atento y en posición de recuperar el control del vehículo en cualquier momento, ya sea porque el coche así se lo indica o porque el conductor detecta que el coche no puede conducir de forma autónoma. Se puede argumentar que, si el conductor no cumple con este deber, bien de forma intencionada bien por negligencia, será responsable en

(51)Véase la Directiva del Consejo 85/374/CEE, de 25 julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados miembros en materia de responsabilidad por los daños causados por productos defectuosos (DOUE n.º 210, 7 de agosto de 1985, pág. 29).

(52)Según el artículo 1 de la Directiva relativa a la responsabilidad por los productos, «el productor será responsable de los daños causados por un defecto en su producto.»

caso de accidente. En cambio, esta situación varía si el conductor cumple con su obligación de atención y supervisión del coche, pero el accidente se produce de todas formas. En este caso, el esquema de responsabilidad podría ser el de la *responsabilidad objetiva del propietario* del coche o el de la *responsabilidad del fabricante*, semejante al concepto de responsabilidad propio de la Directiva de responsabilidad por los productos[51]. El razonamiento que hay detrás de la responsabilidad del propietario es que el coche constituye un riesgo o un peligro para los otros participantes del tráfico en el momento en que circula por la carretera (Shah, 2014) y que, por consiguiente, el propietario debe asumir las consecuencias de dicho riesgo. La responsabilidad del fabricante se basa en la presunción de que el coche es un producto peligroso en sí mismo y que, por ese mero hecho, debería ser responsable[52]. A pesar de ello, la doctrina plantea dudas sobre si el uso de esta directiva es una solución apropiada (de Bruin, 2016; Gurney, 2013). En cualquier caso, es importante asegurar un equilibrio entre, por un lado, no hacer recaer demasiada carga de responsabilidad en el fabricante, lo cual podría frenar la innovación, y, por otro, reducirla tanto que las «funciones de las reglas propias de la responsabilidad civil, en especial la prevención de accidentes y la compensación a la víctima» queden menoscabadas (Shah *et al.*, pág. 63). La situación es diferente si el accidente se ha producido a causa de un mal funcionamiento técnico del coche sin conductor, en cuyo caso se podría justificar más fácilmente la responsabilidad del fabricante (Marchant, 2012).

Referencias bibliográficas

Susanne Pillath (2016). «Briefing January 2016: Automated Vehicles in the EU». European Parliamentary Research Service (PE 573.902, 4) [en línea] <http://www.europarl.europa.eu/regdata/etudes/brie/2016/573902/eprs_bri/2016/573902_EN.pdf>. Acceso: 27 de julio de 2017.

Maurice Schellekens (2015). «Self-driving Cars and the Chilling Effect of Liability Law». *Computer Law & Security Review* (vol. 31, n.º 4, pág. 507).

Melinda F. Lohmann (2016). «Liability Issues Concerning Self-driving Vehicles». *European Journal of Risk Regulation* (n.º 7, págs. 335-340).

Huma Shah, **Kevin Warwick**, **Frederica Lucivero**, **Maurice Schellekens** (2014). «Self-driving Cars». En: Azzari, Frederico *et al. Guidelines on Regulating Robotics*. *RoboLaw* (pág. 65) [en línea] <http://www.robolaw.eu/robolaw_files/documents/robolaw_d6.2_guidelinesregulatingrobotics_20140922.pdf>. Acceso:26 de julio de 2017.

Roeland de Bruin (2016). «Autonomous Intelligent Cars on the European Intersection of Liability and Privacy - Regulatory Challenges and the Road Ahead». *European Journal of Risk Regulation* (vol. 7, n.º 3, pág. 490 y sigs.).

Jeffrey K. Gurney (2013). «Sue My Car not Me: Products Liability and Accidents Involving Autonomous Vehicles». *University of Illinois Journal of Law* (n.º 2, pág. 271 y sigs.).

Gary E. Marchant, **Rachel A. Lindor** (2012). «The Coming Collision Between Autonomous Vehicles and the Liability System». *Santa Clara Law Review* (vol. 52, n.º 4, pág. 1328 y sigs.).

Para los coches sin conductor será crucial el seguro obligatorio, regulado actualmente por la Directiva 2009/103/CE relativa al seguro de responsabilidad civil respecto al uso de vehículos a motor[53]. Aunque los modelos de seguro actuales todavía se basan en la responsabilidad para determinar qué seguro

debe cubrir los daños (el del propietario o el del conductor), Suecia introdujo un modelo de seguro diferente por el que la víctima está asegurada frente al daño que se le cause en un accidente de tráfico (Shah *et al.*, 64). En este modelo de no responsabilidad[54] no debe establecerse la responsabilidad y la «víctima de un accidente de tráfico» recibe compensación de un asegurador de «primera parte», es decir, se trata de un seguro frente al daño, no frente a la responsabilidad (Shah et. al., pág. 64). A pesar de que este modelo parece tener la ventaja de cubrir a la víctima de forma consecuente por daños y perjuicios, es esta quien asume los costes del seguro y no quien causa los daños (Shah et. al., pág. 64). Otro modelo potencialmente posible sería el basado en otrogar al agente autónomo dotado de IA (el coche sin conductor) una personalidad jurídica y establecer, así, la propia *«responsabilidad» del coche auto-conducido*. Esta personalidad jurídica implicaría un seguro obligatorio para dicho coche que permitiera compensar los daños causados en caso de accidente por dicho coche. El problema de este modelo de responsabilidad es que cumple con el papel compensatorio por daños, pero no la función punitiva o, no necesariamente, la de prevención de accidentes. Además, es muy probable que en este modelo el propietario del coche tenga que pagar los costes del seguro, lo cual, en última instancia, hace que este modelo sea igual que el de responsabilidad del propietario.

(53)Directiva 2009/103/CE del Parlamento Europeo y del Consejo, de 16 de septiembre de 2009, relativa al seguro de responsabilidad civil que resulta de la circulación de vehículos automóviles así como al control de la obligación de asegurar esta responsabilidad (DOUE n.º 263, 7 de octubre de 2009, pág. 11).

(54)Este modelo se denomina también «seguro de primera parte» o «esquema de no culpa». Véase: Andrea Bertolini y Erica Palmerini (2014). «Regulating robotics - A challenge for Europe». En: *European Parliament Directorate General for Internal Policies Policy Department C: Citizens' Rights and Constitutional Affairs: Workshop on upcoming issues of EU law*. Strasbourg: Policy Department C: Citizens' Rights And Constitutional Affairs, 187.

Resumen

Este módulo analiza las reglas contenidas en el RGPD y en la Directiva relativas a la protección en asuntos penales (Directiva 2016/680) en lo que respecta a la adopción de decisiones automatizadas. Se considera que, aunque estas normas claramente otorgan al interesado el derecho a no ser objeto de una decisión totalmente automatizada, incluída la elaboración de perfiles, las excepciones a este derecho lo vacían de contenido hasta el punto que las propias excepciones se convierten en norma. A este respecto, la excepción más importante es la posibilidad de que tanto la legislación de un Estado miembro como la de la UE, respaldadas ambas por el RGPD y la directiva, permitan la toma de decisiones automatizadas.

En este mòdulo también se afirma que el interesado, desde la perspectiva de la búsqueda intensa de una gran transparencia en la legislación de la UE relativa a la protección de datos, debería tener derecho a una explicación sobre la decisión automatizada y recibir «información significativa sobre la lógica aplicada» en dicha decisión. A pesar de las voces que consideran que dichas normas, especialmente el RGPD, no reconoce este derecho a la explicación (Wachter, 2017), en este módulo se sostiene que una interpretación teniendo en cuenta la finalidad de las disposiciones del RGPD, junto con una búsqueda más amplia de un gran nivel de transparencia en el tratamiento de los datos, debería llevar al establecimiento de ese derecho para un interesado. No obstante, será el Tribunal de Justicia de la UE quien tendrá la última palabra en el momento de interpretar las disposiciones relevantes del RGPD.

Asimismo, el hecho de eliminar los obstáculos a la transparencia algorítmica no solo permitiría al individuo tener una mejor visión de las razones por las que se tomó una decisión, sino que también sería importante en otros dos aspectos (sobre la discriminación relacionada con la toma de decisiones automatizada, véase Newell, 2015). Por un lado, ayudaría a eliminar los sesgos discriminatorios en el propio proceso de decidir. La discriminación se puede producir en decisiones automatizadas algorítmicas porque los conjuntos de datos con los que opera el algoritmo pueden estar influenciados y, por lo tanto, el propio algoritmo aprende a estar sesgado. Sin embargo, algunos autores afirman que eliminar los datos sesgados del proceso automatizado puede dar lugar, no obstante, a resultados sesgados (para más información sobre discriminación algorítmica, véase Goodman, 2016). Por otra parte, la responsabilidad está íntimamente ligada a la transparencia algorítmica. En este módulo se afirma que la responsabilidad debería estar repartida entre o atribuida a tanto al programador como al usuario del algoritmo. El programador asumiría la responsabilidad de los aspectos relativos a cómo se diseñó el algoritmo, mientras que el usuario, al alimentar el algoritmo con datos, podría ser responsable del aprendizaje del algoritmo en base a dichos datos. No obstante, también

advierte acerca de los problemas de establecer una responsabilidad objetiva tanto del programador como del usuario en lo que respecta a las decisiones de un algoritmo, puesto que este no debería ser considerado como un producto peligroso. También se rechaza la posibilidad de la responsabilidad del propio agente autónomo, puesto que este enfoque podría llevar a la elusión total de dicha responsabilidad.

Referencias bibliográficas

Sue Newell, **Marco Marabelli** (2015). «Strategic Opportunities (and Challenges) of Algorithmic Decision-making: A Call for Action on the Long-term Societal Effects of "Datification"». *Journal of Strategic Information Systems* (n.º 24, págs. 3-14).

Indré Žliobaité, **Bart Custers** (2016). «Using Sensitive Personal Data May Be Necessary for Avoiding Discrimination in Data-driven Decision Models». *Artif Intell Law* (n.º 24, págs. 183–201).

Toon Calders, **Indré Žliobaité** (2013). «Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures». En: Bart Custers *et al.* (eds.). *Discrimination and Privacy in the Information Society. Data Mining and Profiling in Large Databases* (págs. 43-57). Springer.

Bryce Goodman (2016). «Discrimination, Data Sanitisation and Auditing in the European Union's General Data Protection Regulation». *European Data Protection Law Review* (vol. 2, n.º 4, págs. 493-506).

Este módulo también analiza diferentes tipos de obstáculos a la transparencia algorítmica y, por consiguiente, al derecho a la explicación que tiene el individuo: obstáculos técnicos, relativos a la propiedad intelectual, secretos de Estado y otra información confidencial de las autoridades estatales. Aunque los obstáculos relacionados con las PI se consideran como meros tigres de papel, los secretos de Estado pueden y podrán dificultar, en muchos casos de forma justificada, la transparencia algorítmica. Puesto que este no es el valor último de la sociedad, se tendrá que ponderar con los intereses públicos y estatales, en especial en aquellos casos en los que esté en riesgo la seguridad pública. Por último, aunque los obstáculos técnicos representan la mayor dificultad para alcanzar un nivel elevado de transparencia algorítmica, de manera gradual los científicos están desarrollando medios para superarlos. No obstante, las decisiones autónomas tomadas con la ayuda de las redes neuronales podrían convertirse en «cajas negras», muchas de ellas no explicables a los individuos.

Finalmente, este módulo propone diferentes modelos de responsabilidad para los coches sin conductor con el fin de ilustrar el acuciante problema jurídico de la responsabilidad para/de los agentes autónomos dotados de IA. Otro asunto que requerirá mayor investigación es el altamente problemático impacto que tienen en los derechos fundamentales de la UE la industria que evoluciona hacia la automatización, el uso creciente de agentes dotados de IA inteligentes y autónomos, la compartición de BD, los sistemas ciberfísicos, la internet de las cosas y la computación en la nube. Son cada vez más relevantes aspectos como la menor protección de los derechos fundamentales de los ciudadanos de la UE debida al desarrollo tecnológico, y los derechos que los agentes autónomos y los robots deberían tener como resultado de la cuarta revolución industrial que estamos viviendo. La protección de los derechos fundamentales

(55)Véase el artículo 1 de la Carta de los derechos digitales fundamentales de la UE, disponible en <https://digitalcharta.eu/wp-content/uploads/2016/12/digital-charta-en.pdf>.

es uno de los pilares de la sociedad democrática moderna, pero si los derechos de las personas entran en conflicto con los de agentes inteligentes capaces de realizar un aprendizaje automático, la protección de los primeros estaría en peligro. Por supuesto, si a los robots se les dota de una personalidad electrónica, ¿también comporta esto garantizarles derechos fundamentales como el derecho a la vida (por ejemplo, el derecho de un robot a no ser apagado o destruido)? Con el objetivo de proteger mejor los derechos de los ciudadanos europeos, se ha planteado un debate sobre la necesidad de una nueva «Carta de los derechos digitales fundamentales de la UE» que incluye, por ejemplo, un derecho fundamental a la dignidad digital[55]¿Qué impacto tiene en nuestros derechos fundamentales el hecho de que nuestros datos tengan un valor monetario para las empresas? Nuestros derechos fundamentales tendrían que estar garantizados por la propia tecnología, por ejemplo, mediante altos estándares roboéticos incorporados en ella.

Bibliografía

Auditoría sobre el Reglamento de protección general de datos de la Unión Europea (2016). 2 Revisión de la legislación europea sobre protección de datos (págs. 493-506).

Bench-Capon, T.; Gordon, T. (2015). «Tools for Rapid Prototyping of Legal Cased-Based Reasoning». ULCS (15-005). United Kingdom: University of Liverpool.

Sartor, G.; Branting, L. (eds.) (1998). *Judicial Applications of Artificial Intelligence* Springer.

Bertolini, A. (2013). «Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules». *LIT* (vol. 5, n.º 2, págs. 214–247).

Burrell, J. (2016). «How the Machine "thinks": Understanding Opacity in Machine Learning Algorithms». *Big Data & Society*(1-12).

Bertolini, A; Palmerini, E. (2014). *Regulating robotics: A challenge for Europe*. In EU Parliament, Workshop on «Upcoming issues of EU law» for the IURI Committee, Compilation of in-depth analyses. Brussels: Publications Office of the EU Parliament.

Calders, T.; Žliobaité, I. (2013). «Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures». En: Custers *et al*. (eds.). *Discrimination and Privacy in the Information Society. Data Mining and Profiling in Large Databases* (págs. 43-57). Springer.

Calo, R. (2010). «Peeping HALs: Making Sense of Artificial Intelligence and Privacy». *European Journal of Legal Studies* (n.º 2, pág. 3) [en línea] <http://www.ejls.eu/6/83UK.htm>.

Christin, A.; Rosenblat, A.; Boyd, D. (2015). «Courts and Predictive Algorithms». *Data & Civil Rights: A New Era of Policing and Justice* [en línea] <http://www.law.nyu.edu/sites/default/files/upload_documents/Angele%20Christin.pdf. Acceso: 16 de enero de 2017.

Coormen, T. H. (2013). *Algorithms Unlocked*. MIT Press 1.

Datta, A.; Shayak, S.; Zick, Y. ((2016)). «Algorithmic Transparency via Quantitative Input Influence» [en línea] <https://www.andrew.cmu.edu/user/danupam/datta-sen-zick-oakland16.pdf>. Acceso: 10 de diciembre de 2016.

de Bruin, R. (2016). «Autonomous Intelligent Cars on the European Intersection of Liability and Privacy - Regulatory Challenges and the Road Ahead». *European Journal of Risk Regulation* (vol. 7, n.º 3, págs. 485-501).

Diakopoulos, N. (2016). «Accountability in Algorithmic Decision Making». *Communications of the ACM* (vol. 59, n.º 56, págs. 58-59).

Edwards, L.; Veale, M. (2017). «Slave to the Algorithm? Why a "Right to an Explanation" Is Probably not the Remedy You Are Looking For» [en línea] <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2972855>.

Goodman, B. (2016). «Discrimination, Data Sanitisation and Auditing in the European Union's General Data Protection Regulation». *European Data Protection Law Review* (n.º 493).

Goodman, B.; Flaxman, S. (2016). «European Union Regulations on Algorithmic Decision-making and a "Right to Explanation"». *ICML Workshop on Human Interpretability in Machine Learning*. New York [en línea] <https://arxiv.org/abs/1606.08813v3>.

Gurney, J. K. (2013). «Sue My Car Not Me: Products Liability And Accidents Involving Autonomous Vehicles». *University of Illinois Journal of Law* (n.º 2, págs. 247-277).

Karjala, D. S. (1991). «Japanese Courts Interpret the "Algorithm" Limitation on the Copyright Protection of Programs». *Jurimetrics Journal* (n.º 31, pág. 233).

Klimas, T.; Vaiciukaitè, J. (2008). «The Law of Recitals in European Community Legislation». *ILSA Journal of International & Comparative Law* (n.º 15, pág. 26).

Kroll, J. *et al*. (2017). «Accountable Algorithms». *University of Pennsylvania Law Review* (n.º 165, 1, pág. 18) [en línea] <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2765268>.

Lohmann, M. F. (2016). «Liability Issues Concerning Self-Driving Vehicles». *European Journal of Risk Regulation* (n.º 7, págs. 335-340).

Loveless, J.; Stoikov, S; Waeber, R. (2013). «Online Algorithms in High-frequency Trading. The challenges faced by competing HFT algorithms». *Communications of the ACM* (vol. 56, nº. 10, págs. 50-56)

Mantelero, A. (2016). «Personal Data for Decisional Purposes in the Age of Analytics: From an Individual to a Collective Dimension of Data Protection». *Computer Law & Security Review* (vol. 32, n.º 2, págs. 238-255).

Marchant, G. E.; Lindor, R. A. (2012). «The Coming Collision between Autonomous Vehicles and the Liability System». *Santa Clara Law Review* (vol. 52, n.º 4, págs. 1321-1340).

Masnick, M. (2016). «Activists Cheer On EU's "Right To An Explanation" For Algorithmic Decisions, But How Will It Work When There's Nothing To Explain? [en línea] <https://www.techdirt.com/articles/20160708/11040034922/activists-cheer-eus-right-to-explanation-algorithmic-decisions-how-will-it-work-when-theres-nothing-to-explain.shtml>.

Mendoza, I.; Bygrave, L. A. (2017). «The Right not to be Subject to Automated Decisions based on Profiling». *University of Oslo Faculty of Law Legal Studies Research* Paper Series (n.º 20/2017). University of Oslo [en línea] <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2964855>.

Metz, C. (2016). «Artificial Intelligence Is Setting Up the Internet for a Huge Clash With Europe» [en línea] <https://www.wired.com/2016/07/artificial-intelligence-setting-internet-huge-clash-europe/>.

Metz, C. (2016). «The Sadness and Beauty of Watching Google's AI Play Go» [en línea] <https://www.wired.com/2016/03/sadness-beauty-watching-googles-ai-play-go/>.

Mittelstadt, B. D.; Allo, P.; Taddeo, M. R.; Wachter, S.; Floridi, L. (2016). «The ethics of algorithms: Mapping the debate». *Big Data & Society* (vol. 3, n.º 1, págs. 1–21).

Newell, Marco Marabelli (2015). «Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of "datification"». *24 Journal of Strategic Information Systems* (3, 14).

Opinión de la EPO G 0003/08 (Programas para ordenadores) of 12 de mayo de 2010, ECLI:EP:BA:2010:G000308.20100512, punto 13.5.

Orphanides, K. G. (2017). «Robot Carries Out First Autonomous Soft Tissue Surgery» [en línea] <http://www.wired.co.uk/article/autonomous-robot-surgeon>. Acceso: 25 de julio de 2017.

Palmerini, E.; Bertolini, A. (2016). «Liability and Risk Management in Robotics». En: Schulze, R.; Staudenmayer, D. (eds.). *Digital Revolution: Challenges for Contract Law in Practice* (págs. 225-260). Nomos.

Pasquale, F. (2016). *The Black Box Society. The Secret Algorithms That Control Money and Information*. Harvard University Press.

Perry, M. (2017). «iDecide: Administrative Decision-Making in the Digital World». *Australian Law Journal*

Pillath, S. (2016). «Briefing January 2016: Automated Vehicles in the EU». *European Parliamentary Research Service* (PE 573.902) [en línea] <http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/573902/EPRS_BRIE/2016/573902_EN.pdf>.

Russell, S. J.; Norvig, P. (eds.) (2010). *Artificial Intelligence. A Modern Approach* (3.ª ed.). Pearson.

Schellekens, M. (2015). «Self-driving Cars and the Chilling Effect of Liability Law». *Computer Law & Security Review* (vol. 31, n.º 4, págs. 506-517).

Scholz, L. H. (2017). «Algorithmic Contracts». *Stanford Technology Law Review* [en línea] <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=274770>. Acceso: 16 de noviembre de 2016.

Shah, H.; Warwick, K.; Lucivero, F.; Schellekens, M. (2014). «Self-Driving Cars». En: Azzari, F. *et al.*. *Guidelines on Regulating Robotics*. *RoboLaw* [en línea] <http://www.robolaw.eu/RoboLaw_files/documents/robolaw_d6.2_guidelinesregulatingrobotics_20140922.pdf>.

Stern, R. (1995). «On Defining the Concept of Infringement of Intellectual Property Rights in Algorithms and Other Abstract Computer-Related Ideas». *AIPLA Quarterly Journal* (n.º 23, pág. 401).

Swinson, J. (1991). «Copyright or Patent or Both: An Algorithmic Approach to Computer Software Protection». *Harvard Journal of Law & Technology* (n.º 5, pág. 145).

Vromen, S. (2015). «Ethnic profiling in the Netherlands and England and Wales: Compliance with international and European standards». *Public Interest Litigation Project (PILP-NJCM)*. Utrecht University [en línea]. <https://pilpnjcm.nl/wp-content/uploads/2015/06/Research-project-B-FINAL-version.pdf>.

Wachter, S.; Mittelstadt, B.; Floridi, L. (2017). «Why a Right to Explanation of Automated Decision-making Does not Exist in the General Data Protection Regulation» [en línea] <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2903469>. Acceso: 18 de mayo de 2017.

Zarsky, T. (2016). «The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making». *Science, Technology, & Human Values* (n.º 41, págs. 118-132).

Žliobaité, I.; Custers, B. (2016). «Using Sensitive Personal Data May Be Necessary for Avoiding Discrimination in Data-driven Decision ModelsV». *Artif Intell Law* (n.º 24, págs. 183–201).