

¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?*

Agustí Cerrillo i Martínez

(Catedrático de Derecho Administrativo en la Universitat Oberta de Catalunya)

ABSTRACT Public administrations are already using artificial intelligence in decision-making as support to public officials but also to automate decision-making. Despite the benefits that this may entail, throughout this article it becomes clear that the use of artificial intelligence in public decision-making cannot be done in any way since this can generate several problems such as opacity, legal uncertainty, errors, biases or breaches of personal data protection. The mechanisms already provided by law are not enough to avoid these problems as they have not been designed to face the use of artificial intelligence in public administrations. Therefore, the article concludes that public administration should strengthen the reliability of algorithms through the guarantee of their legality, security and ability to obtain quality public decisions.

1. Introducción

La inteligencia artificial puede ser útil en la toma de decisiones públicas. En efecto, el uso de la inteligencia artificial permite que las Administraciones públicas puedan tener un mayor conocimiento del contexto en el que toman sus decisiones, valorar de manera más rápida los antecedentes y los precedentes, anticipar el impacto de las decisiones, tomar decisiones de manera automatizada o evaluar su impacto o efectividad. De este modo, la utilización de la inteligencia artificial puede dotar de mayor eficacia o calidad a las decisiones que adoptan las Administraciones públicas¹.

Sin embargo, a pesar de los potenciales beneficios del uso de la inteligencia artificial en la toma de decisiones públicas, no podemos desconocer los riesgos que también puede entrañar debido no solo a las limitaciones actuales de la tecnología que pueden traducirse en errores, sino también a la incidencia que el uso de algoritmos por las Administraciones públicas puede tener en los principios que guían su actuación y funcionamiento (por ejemplo, la seguridad jurídica, la igualdad o la transparencia) o en los derechos de las personas (en particular, la protección de datos personales, la intimidad o la igualdad).

* Article submitted to double-blind peer review.

Este artículo ha sido elaborado en el marco del proyecto "Retos jurídicos del uso de datos masivos para el fomento de la innovación y la buena administración a través de la inteligencia artificial", Programa Estatal de I+D+i Orientada a los Retos de la Sociedad (DER2017-85595-R).

¹ Cobbe afirma que la toma de decisiones automatizada será prevalente en el futuro (J. Cobbe, *Administrative law and the machines of government: judicial review of automated public-sector decision-making*, en *Legal Studies*, núm. 39, 4, 2019). De hecho, en la actualidad, las Administraciones públicas ya están utilizando la inteligencia artificial para tomar decisiones. Véase una muestra de diversos usos de la inteligencia artificial en la toma de decisiones automatizada en Autoritat Catalana de Protecció de Dades, *Inteligència Artificial. Decisions Automatitzades a Catalunya*, 2020.

De este modo, no cualquier algoritmo ni cualquier uso del mismo por las Administraciones públicas debe considerarse adecuado para la satisfacción del interés general. El desplazamiento de las personas por los algoritmos en la toma de decisiones públicas exige una valoración de los efectos que ello puede tener en la gestión del interés general con el fin de evitar que puedan generarse errores, opacidad o falta de fiabilidad en las decisiones o de confianza o credibilidad en las instituciones.

A lo largo de este trabajo nos proponemos identificar los riesgos y problemas que puede implicar el uso de la inteligencia artificial en la toma de decisiones por las Administraciones públicas y analizaremos los mecanismos que ya prevé el ordenamiento jurídico para hacerles frente así como las interpretaciones que se están haciendo de la normativa vigente para hacer frente a esta nueva realidad. A la vista de todo ello, identificaremos los elementos que deben concurrir para garantizar la fiabilidad de las decisiones adoptadas a través de algoritmos por las Administraciones públicas.

Como punto de partida y con carácter previo a nuestro análisis, debemos concretar qué entendemos por inteligencia artificial puesto que, hoy en día, no existe un único concepto de inteligencia artificial. Como se ha puesto de manifiesto en diversas ocasiones, la expresión *inteligencia artificial* se utiliza para hacer referencia a tecnologías muy diversas que derivan de tradiciones, aproximaciones y tecnologías distintas.

Desde un punto de vista general, la inteligencia artificial es aquella capacidad de los ordenadores de realizar tareas típicamente humanas, es decir, que entrañan el uso de inteligencia². A través de la inteligencia artificial,

² J. Kaplan, *Inteligencia artificial. Lo que todo el mundo debe saber*, Zaragoza, Teell, 2017.

los ordenadores emulan facultades intelectuales humanas para realizar tareas propias de las personas.

De una manera más precisa, en este trabajo utilizaremos una noción más restringida de inteligencia artificial para referirnos únicamente a aquellos algoritmos que tienen la capacidad de aprender de los datos y de la experiencia y, a partir de ellos, formar nuevos criterios o tomar decisiones autónomas, es decir, ceñimos la noción de inteligencia artificial al aprendizaje automático o *machine learning*³. Los algoritmos de aprendizaje automático identifican patrones y correlaciones entre los datos que son utilizados posteriormente para generar modelos cuya calidad es evaluada estadísticamente y, en su caso, mejorada a partir de los resultados que se van obteniendo⁴. Posteriormente, los algoritmos son utilizados sobre nuevos datos, por ejemplo, para tomar decisiones en otros casos. El aprendizaje automático puede tener una mayor (aprendizaje supervisado) o menor intervención humana (aprendizaje no supervisado, aprendizaje profundo o *deep learning*) lo que determina una mayor o menor autonomía del algoritmo.

El aprendizaje automático ya está demostrando resultados positivos en diversos ámbitos. Sin embargo, a pesar del potencial que tiene el aprendizaje automático, no podemos desconocer que en el estadio actual de desarrollo la inteligencia artificial únicamente consigue realizar tareas específicas como traducir textos, conducir vehículos sin conductor o reconocer imágenes (es lo que se conoce como inteligencia artificial limitada) estando aún lejos conseguir que los ordenadores puedan tener una conducta tan avanzada como una persona respecto a un amplio conjunto de tareas cognitivas (es decir, inteligencia artificial general)⁵.

A partir de este punto, el artículo se organiza en las siguientes secciones. En primer lugar, exponemos los beneficios y las limitaciones del uso de la inteligencia artificial en las Administraciones públicas. En segundo lugar, analizamos la regulación del uso de la inteligencia artificial en la toma de decisiones públicas automatizada. En tercer lugar, valoramos el impacto del uso de la inteligencia artificial en el ejercicio de potestades discrecionales. En cuarto lugar, estudiamos cómo puede incidir el uso de la inteligencia artificial en

el principio de buena administración que debe guiar el ejercicio de estas potestades. En quinto lugar, exponemos los mecanismos para la supervisión del uso de los algoritmos por las Administraciones públicas. Finalmente, concluimos el trabajo reflexionando sobre la fiabilidad de las decisiones adoptadas por algoritmos y reflexionamos sobre la necesidad y oportunidad de definir el principio de fiabilidad en el uso de la inteligencia artificial en las Administraciones públicas.

2. El uso de algoritmos en la toma de decisiones públicas

La inteligencia artificial tiene diferentes aplicaciones en la toma de decisiones públicas.

Actualmente, los cargos y empleados públicos pueden contar con la inteligencia artificial en las Administraciones públicas como ayuda o apoyo a la toma de decisiones públicas. En esta dirección, cada vez con mayor frecuencia, las Administraciones públicas utilizan los algoritmos para analizar datos necesarios para adoptar decisiones o para automatizar algunas fases del proceso de toma de decisiones de manera que se aumenten las capacidades de los empleados públicos sin sustituir por completo su implicación o participación sino complementándola o aumentándola. En estos casos, son las personas las que, en última instancia, adoptan la decisión. Desde esta perspectiva, los algoritmos pueden facilitar información o conocimiento relevante para la toma de decisiones a partir del análisis de los datos obtenidos de documentos, de imágenes o de sensores examinados automáticamente y a partir de aquí identificar patrones o perfiles que permitan recomendar o adoptar decisiones más precisas o ajustadas a las necesidades de sus destinatarios y, en última instancia, al interés general. Incluso pueden proponer decisiones que posteriormente son asumidas por el decisor. El uso de la inteligencia artificial permite llevar a cabo tareas que con frecuencia sería imposible que fuesen realizadas por empleados públicos o tendría un coste muy elevado si fuesen desarrolladas por ellos o simplemente permite simplificar estas tareas, por ejemplo, al analizar grandes cantidades de datos o realizar de manera rápida o sencilla un gran volumen de actividades rutinarias o repetitivas.

Pero la inteligencia artificial también se puede utilizar para tomar de manera automatizada decisiones públicas por un algoritmo sin que haya ninguna intervención humana. Estas decisiones elaboradas por los algoritmos pueden tener plenos efectos jurídicos. En estos casos, los algoritmos pueden basarse en el diseño que previamente haya hecho una persona para adoptar la decisión (aprendizaje supervisado). Pero también puede darse el caso en que el

³ En la misma dirección, entre otros, M. M. Young, J. B. Bullock y J. D. Lecy, *Artificial Discretion as a Tool of Governance: A Framework for Understanding the Impact of Artificial Intelligence on Public Administration*, en *Perspectives on Public Management and Governance*, núm. 2, 4, 2019.

⁴ Mayer-Schonberger y Cukier se refieren a *dejar hablar a los datos* (V. Mayer-Schonberger y K. Cukier, *Big data. La revolución de los datos masivos*, Madrid, Turner, 2013).

⁵ Executive Office of the President, *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights*, 2016.

¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?

algoritmo a partir de correlaciones e inferencias de los datos a su disposición adopte la decisión de manera automatizada (aprendizaje no supervisado).

En uno y otro caso la inteligencia artificial puede generar beneficios pero también puede entrañar limitaciones y riesgos para las Administraciones públicas.

2.1. Beneficios del uso de la inteligencia artificial en la toma de decisiones públicas

El uso de la inteligencia artificial en la toma de decisiones públicas puede reportar distintos beneficios para las Administraciones públicas⁶.

En primer lugar, los beneficios están directamente relacionados con el potencial de la inteligencia artificial que está estrechamente vinculado al desarrollo tecnológico y al incremento de la capacidad de computación de los ordenadores. Hoy en día, la inteligencia artificial permite el análisis de grandes cantidades de datos; la extracción de datos de textos, imágenes, vídeos o sensores; la identificación de patrones o perfiles; la elaboración de predicciones o recomendaciones o la adopción de decisiones de manera autónoma.

En segundo lugar, los beneficios pueden estar relacionados con la ausencia de la intervención de las personas. La inteligencia artificial puede dar respuesta a algunos de los problemas que puede entrañar la participación de las personas en la toma de decisiones como, por ejemplo, errores o sesgos, incoherencias causadas por la participación de distintas personas, por el trascurso del tiempo o por la situación personal del decisor (cansancio, problemas personales), conflictos de intereses y corrupción o costes laborales.

El uso de la inteligencia artificial puede suponer una oportunidad para mejorar estos problemas aportando mayor precisión en las decisiones, reducción de sesgos, menor corrupción y un ahorro vinculado a la automatización de decisiones rutinarias⁷. También puede hacer más visibles los errores o los sesgos de los humanos⁸.

De todos modos, estos beneficios no deben llevarnos a caer en el sesgo de la automatización, es decir, a considerar que las decisiones

adoptadas por máquinas son, simplemente por ello, más confiables que aquellas adoptadas por las personas⁹.

2.1.1. Limitaciones y riesgos del uso de la inteligencia artificial en la toma de decisiones públicas

Si bien, las ventajas de la automatización son evidentes, en la actualidad el uso de la inteligencia artificial no está exenta de limitaciones y riesgos que pueden llevar a considerar la necesidad de evitar tomar decisiones públicas o con efectos jurídicos sobre la ciudadanía a través de la inteligencia artificial. Algunas de ellas, pueden ser difíciles de prever, identificar o medir¹⁰.

En primer lugar, existen limitaciones tecnológicas que pueden desaconsejar el uso de la inteligencia artificial en la toma de decisiones públicas. A pesar del importante desarrollo experimentado por la inteligencia artificial en la última década, los algoritmos no son capaces de tomar cualquier decisión de manera autónoma y, con frecuencia, cuando lo hacen pueden generar daños o errores. En otras ocasiones, los algoritmos no disponen de los datos suficientes, adecuados o con la calidad suficiente para que los algoritmos puedan tomar decisiones de calidad.

En segundo lugar, existen limitaciones organizativas. Es habitual que las Administraciones públicas adquieran los algoritmos del mercado y no dispongan de las capacidades internas necesarias para valorar adecuadamente el impacto del uso de la inteligencia artificial en la toma de decisiones públicas o ni siquiera para determinar qué usos son adecuados para la satisfacción del interés general. También es frecuente que las Administraciones públicas carezcan de una gobernanza de datos con la que poder garantizar la calidad de los datos utilizados por los algoritmos en la toma de decisiones.

En tercer lugar, existen limitaciones jurídicas. Estas limitaciones pueden ser de muy distinto orden, como iremos viendo a lo largo de estas páginas, y pueden exigir que intervengan personas en la toma de decisiones públicas.

Además de estas limitaciones, el uso de la inteligencia artificial en las Administraciones públicas puede entrañar diversos riesgos que pueden poner a su vez en entredicho los

⁶ La sentencia del Consejo de Estado italiano número 8472 de 13 de diciembre de 2019 identifica algunas de las ventajas del uso de las decisiones automatizadas como, por ejemplo, la reducción del tiempo en la realización de tareas repetitivas, la exclusión de interferencias debidas a negligencias de los funcionarios públicos y la mayor garantía de imparcialidad.

⁷ M. M. Young, J. B. Bullock y J. D. Lecy, *Artificial Discretion as a Tool of Governance*.

⁸ Y. Hermstrüwer, *Artificial Intelligence and Administrative Decisions*, en *Regulating Artificial Intelligence*, T. Wischmeyer y T. Rademacher (eds.), Cham, Springer, 2020, 199-224.

⁹ J. Cobbe, *Administrative law and the machines of government: judicial review of automated public-sector decision-making*; J. Goldenfein, *Algorithmic Transparency and Decision-Making Accountability: Thoughts for Buying Machine Learning Algorithms*, en *Closer to the Machine: Technical, Social, and Legal aspects of AI*, Office of the Victorian Information Commissioner, 2019.

¹⁰ Grupo de expertos de alto nivel sobre inteligencia artificial, *Directrices éticas para una IA fiable*, 2019.

principios que rigen el funcionamiento de las Administraciones públicas.

Tal vez el principal riesgo que puede generar el uso de la inteligencia artificial por las Administraciones públicas son los errores en las decisiones que adoptan. Un segundo riesgo, que ha sido objeto de una mayor atención cuando las Administraciones públicas utilizan la inteligencia artificial, es la opacidad¹¹. Otro riesgo frecuente que puede derivarse del uso de la inteligencia artificial en las Administraciones públicas son los sesgos y las discriminaciones¹². Un cuarto riesgo, con frecuencia el que genera una mayor preocupación ciudadana, está relacionado con el impacto de la inteligencia artificial en la protección de los datos personales y la privacidad de las personas¹³. A todos ellos nos iremos refiriendo en las próximas páginas.

Ante la existencia de estas limitaciones y riesgos, las Administraciones públicas deberían realizar análisis de riesgos que eviten que estos se puedan llegar a materializar gracias a la adopción de las medidas necesarias para evitarlos¹⁴.

3. La automatización de la toma de decisiones públicas a través de la inteligencia artificial

En los últimos años, a medida que se han ido desarrollando las tecnologías de la información y la comunicación, se ha entendido que los medios electrónicos también podían ser utilizados por las Administraciones públicas para la toma automatizada de decisiones. Esta cuestión se ha vuelto a plantear en el momento en que la inteligencia artificial se está extendiendo para la automatización de decisiones a nivel económico y social.

En la actualidad, el ordenamiento jurídico prevé la posibilidad de que las Administraciones públicas puedan adoptar decisiones automatizadas utilizando para ello los medios electrónicos a su alcance.

Sin embargo, el uso de la inteligencia artificial en la toma de decisiones públicas

entraña algunas cuestiones específicas tales como su alcance y uso, la vinculación de la decisión adoptada por un algoritmo por parte de las Administraciones públicas, las fases del procedimiento administrativo en las que pueden intervenir los algoritmos o la motivación de las decisiones públicas que se toman a través de algoritmos.

Como punto de partida para responder a estas cuestiones, podemos traer a colación la Ley 40/2015, de 1 de octubre, de régimen jurídico del sector público (LRJSP) que regula la actuación administrativa automatizada entendiendo como tal “cualquier acto o actuación realizada íntegramente a través de medios electrónicos por una Administración Pública en el marco de un procedimiento administrativo y en la que no haya intervenido de forma directa un empleado público”¹⁵.

Así, desde un punto de vista general, cuando una Administración pública utilice la inteligencia artificial para la adopción de actos administrativos estaremos ante una actuación administrativa automatizada siempre y cuando concurren los siguientes elementos que dispone la LRJSP:

a) Actuación realizada a través de medios electrónicos.

La LRJSP utiliza de manera genérica la expresión *medios electrónicos* a la hora de definir la actuación administrativa automatizada. Si bien es probable que el legislador no tuviese en mente la posibilidad de utilizar algoritmos de aprendizaje automático en la toma de decisiones, como ya hemos avanzado, hoy en día se reconoce unánimemente que la inteligencia artificial encaja dentro de esta expresión y que, de este modo, no hay obstáculo alguno a su uso por las Administraciones públicas¹⁶.

b) Actuación en el marco de un procedimiento administrativo.

La LRJSP se refiere de este modo a actuaciones administrativas formalizadas realizadas durante la tramitación de un procedimiento administrativo¹⁷. Ello no es

¹¹ R. Brauneis y E. P. Goodman, *Algorithmic transparency for the smart city*, en *Yale Journal of Law & Technology*, núm. 20, 2017.

¹² K. Yeung, *Algorithmic regulation: a critical interrogation*, en *Regulation & Governance*, núm. 12, 4, 2017.

¹³ K. Crawford y J. Schultz, *Big data and due process: Toward a framework to redress predictive privacy harms*, en *Boston College Law Review*, núm. 55, 201.

¹⁴ R. Martínez Martínez, *Inteligencia artificial, derecho y derechos fundamentales*, en *Sociedad digital y derecho*, T. De la Quadra-Salcedo y J. L. Piñar Mañas (eds.), Madrid, Boletín Oficial del Estado-Ministerio de Industria, Comercio y Turismo-RED.ES, 2018, 259-278. En esta línea, el *Libro Blanco obre la Inteligencia Artificial de febrero 2020* señala que es necesario calificar los riesgos de los usos concretos y cuando, se dé un alto riesgo, determinar las garantías adecuadas y, en su caso, regular un procedimiento de aprobación.

¹⁵ Artículo 41.1 LRJSP.

¹⁶ La LRJSP no define en la actualidad qué debe entenderse por medio electrónico como sí hacía la Ley 11/2007, de 22 de junio, de acceso electrónico de los ciudadanos a los Servicios Públicos que definía el medio electrónico de una manera muy amplia.

¹⁷ Valero pone de manifiesto que “la regulación de 2015 se sigue centrande en la actividad administrativa formalizada y, en particular, en el acto y el procedimiento administrativos como paradigmas de las garantías subyacentes” lo que, advierte es “un enfoque manifiestamente insuficiente desde la perspectiva de la transformación digital que se demanda del sector público” (J. Valero Torrijos, *Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración*, en *Revista Catalana de Dret Públic*, núm. 58, 2019). También critica las insuficiencias de esta norma L. Cotino Hueso, «SyRI, ¿a

¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?

obstáculo para que la Administración pública pueda utilizar la inteligencia artificial en el desarrollo de otras actuaciones materiales que no se verán sometidas a lo previsto en la Ley 39/2015, de 1 de octubre, de procedimiento administrativo común de las Administraciones públicas (LPACAP). Este sería el caso, por ejemplo, del uso de los algoritmos para la detección de incendios o de zonas de alta inseguridad ciudadana.

En el marco de un procedimiento administrativo, la actuación administrativa automatizada puede llevarse a cabo en relación tanto a los actos de ordenación e instrucción del procedimiento como a la resolución del mismo. Asimismo, los algoritmos pueden ser utilizados en el marco de un procedimiento para la emisión de cualquier declaración de voluntad, de juicio, de conocimiento o de deseo, siguiendo la clásica definición de Zanobini.

Es evidente que el impacto de la inteligencia artificial en cada uno de estos supuestos es distinto y sus efectos jurídicos también.

c) Actuación sin intervención directa de empleado público.

La ausencia de intervención humana es el elemento determinante a la hora de decidir si nos encontramos ante una actuación administrativa automatizada¹⁸.

De todos modos, esta ausencia no es absoluta ni general. En efecto, tal y como dispone la LRJSP, la actuación administrativa automatizada exige la intervención humana a través del órgano u órganos competentes en dos momentos.

En primer lugar, al definir las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente. En relación a este aspecto, aparece la dificultad de determinar el alcance de la definición de las especificaciones y de la programación, en particular, en relación a los algoritmos de aprendizaje automático. Asimismo, desde un punto de vista más general, surge el riesgo de trasladar a los técnicos decisiones de carácter jurídico y normativo¹⁹.

quién sanciona?» Garantías frente al uso de inteligencia artificial y decisiones automatizadas en el sector público y la sentencia holandesa de febrero de 2020, en *Diario La Ley*, 2020.

¹⁸ Más allá de lo previsto en la legislación vigente, no podemos desconocer que algunos autores niegan la posibilidad de que un acto administrativo pueda dictarse sin la intervención de una persona puesto que estos “manifestaciones de voluntad y las máquinas, al no tenerla, carecen de posibilidad de producir actos jurídicos” (R. Parada Vázquez, *Régimen Jurídico de las Administraciones Públicas y Procedimiento Administrativo Común (Estudio, comentarios y texto de la Ley 30/1992, de 26 de noviembre)*, II ed., Madrid, Marcial Pons, 1999).

¹⁹ Esta cuestión ya fue advertida hace años por Fernández Rodríguez quien apuntaba que “hay que convenir que la programación del ordenador y el suministro de material al

En segundo lugar, al resolver los recursos contra las decisiones automatizadas. Al respecto, la LRJSP exige que la Administración pública defina el órgano que debe ser considerado responsable a efectos de impugnación.

Las actuaciones automatizadas adoptadas por un algoritmo se pueden imputar a la Administración pública²⁰. Para ello, lo más deseable sería que se determinase normativamente la idoneidad de un algoritmo para expresar la voluntad de la Administración pública que lo utiliza. Esta determinación podría ser específica para cada procedimiento pero también general²¹. Si bien como acabamos de ver de la lectura de la LRJSP ya se puede observar la obligación de que se definan determinados aspectos relativos al algoritmo, para garantizar la máxima seguridad jurídica sería oportuno que se impulsase una modificación de la legislación vigente con el fin de dar una respuesta a esta cuestión y concretar desde un punto de vista general las características y condiciones que deberían concurrir para que las Administraciones públicas puedan utilizar los algoritmos en la toma de decisiones automatizadas. Además, también sería deseable que las Administraciones

mismo constituyen una especie de codificación, realizada internamente a nivel de técnicos, sin las garantías externas de la codificación formal” (T.-R. Fernández Rodríguez, *Jurisprudencia y computadores*, en *Revista de Administración Pública*, núm. 64, 1971).

En efecto, la atribución de estas competencias a los analistas de datos y desarrolladores de tecnología (S. Zouridis, M. van Eck y M. Bovens, *Automated Discretion*, en *Discretion and the Quest for Controlled Freedom*, T. Evans y P. Hupe (eds.), Cham, Palgrave Macmillan, 2020) entraña una nueva generación de problemas como hasta qué punto están legitimados los técnicos, qué mecanismos de control democrático existen o si son responsables por las decisiones que toman. Ello genera problemas como ha puesto de relieve (M. Bovens y S. Zouridis, *From street-level to system-level bureaucracies: how information and communication technology is transforming administrative discretion and constitutional control*, en *Public Administration Review*, núm. 62, 2, 2002, 174-184, citados por J. B. Bullock, *Artificial intelligence, discretion, and bureaucracy*, en *The American Review of Public Administration*, núm. 49, 7, 2019, 751-761), lo que exige buscar un sistema de imputación de responsabilidad (S. Vaccari, *Note minime in tema di Intelligenza Artificiale e decisioni amministrative*, en *Giustamm - Rivista di diritto amministrativo*, núm. 10, 1-10, 2019).

²⁰ No podemos desconocer la advertencia que formula Valero quien señala que la falta de intervención humana quiebra del elemento subjetivo (J. Valero Torrijos, *El régimen jurídico de la e-Administración*, II ed., Granada, Comares, 2007). No obstante, Martín Delgado opina que “la teoría del órgano, por tanto, sigue vigente y no hay quiebra alguna del elemento subjetivo del acto administrativo” (I. Martín Delgado, *Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada*, en *Revista de Administración Pública*, núm. 180, 2009).

²¹ En esta dirección, véase, como ejemplo, el artículo 17 Ley 22/2018, de 6 de noviembre, de la Generalitat, de Inspección General de Servicios y del sistema de alertas para la prevención de malas prácticas en la Administración de la Generalitat y su sector público instrumental.

públicas adoptasen formalmente la decisión de utilizar algoritmos en la toma de decisiones automatizadas²².

A partir de aquí, el uso de la firma electrónica en la decisión adoptada por el algoritmo permitiría garantizar la vinculación entre el resultado obtenido por la inteligencia artificial y la Administración pública²³. En esta dirección, debemos tener presente que la LRJSP dispone que en el ejercicio de la competencia en la actuación administrativa automatizada, las Administraciones públicas podrán utilizar el sello electrónico de Administración Pública, órgano, organismo público o entidad de derecho público o el código seguro de verificación vinculado a la Administración Pública, órgano, organismo público o entidad de Derecho Público²⁴. De este modo, la decisión adoptada por el algoritmo no se le imputará a él sino que lo será al órgano administrativo que ejerce la potestad atribuida por la legislación²⁵.

El uso de la inteligencia artificial para automatizar la toma de decisiones de las Administraciones públicas debe ser proporcional, es decir, debe ser adecuado, necesaria para mejorar significativamente la eficacia o la calidad de las decisiones que no sea alcanzable con la intervención humana y que tenga un menor coste o impacto en los derechos de las personas o en el funcionamiento de la Administración pública que la participación

exclusiva de las personas²⁶. A *contrario*, los riesgos y los costes que el uso de la inteligencia artificial en las Administraciones públicas pueda entrañar deben ser proporcionales a los beneficios que pueda generar el uso de los algoritmos respecto a la toma de decisiones por las personas²⁷.

4. La adopción de decisiones discrecionales a través de la inteligencia artificial

Según se desprende de lo dicho en la sección anterior, la legislación vigente no dispone ninguna limitación general respecto a qué resoluciones pueden adoptarse a través de la inteligencia artificial ni tampoco qué potestades pueden ejercerse utilizando esta tecnología²⁸.

No obstante, en los últimos años distintos autores se han manifestado de manera contraria o han mostrado sus dudas respecto a la posibilidad de que a través de actuaciones administrativas automatizadas puedan adoptarse decisiones discrecionales.

En efecto, autores como Valero ya se había manifestado respecto al uso de los medios electrónicos en el ejercicio de potestades discrecionales al afirmar que “sólo en el caso de las actuaciones regladas cabría aceptar su completa automatización cuando la decisión a adoptar dependa exclusivamente de los resultados obtenidos a través de un tratamiento automatizado y se implementen las medidas de control y supervisión antes aludidas. En cambio, la naturaleza discrecional del acto administrativo obliga a una respuesta negativa”²⁹. Igualmente,

²² En la academia existen ya algunas propuestas en esta dirección. Así, Valero propone que “la decisión de poner en funcionamiento un sistema de inteligencia artificial debería ser adoptada formalmente mediante un acto administrativo” (J. Valero Torrijos, *Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración*). Por su parte, Boix considera que “las reglas tradicionales ya plenamente asentadas en nuestro Derecho en torno a las exigencias de calidad normativa de los reglamentos, parece no solo muy conveniente que sean plenamente aplicables también respecto de los algoritmos” (A. Boix Palop, *Los algoritmos son reglamentos: La necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la administración para la adopción de decisiones*, en *Revista de Derecho Público: Teoría y Método*, núm. 1, 2020).

²³ Según Martín Delgado “la conexión entre actividad automatizada y órgano se produce a través del uso de un sistema para firmar y autenticar el ejercicio de la competencia” (I. Martín Delgado, *Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada*). Posteriormente, I. Martín Delgado, *Automazione, intelligenza artificiale e pubblica amministrazione: vecchie categorie concettuali per nuovi problemi?*, en *Istituzioni del federalismo. Rivista di studi giuridici e politici*, núm. 3, 2019.

²⁴ Artículo 42 LRJSP.

²⁵ La Ley 11/2007, de 22 de junio, de acceso electrónico de los ciudadanos a los servicios públicos era más clara al efecto al disponer que “la resolución de un procedimiento utilizando medios electrónicos garantizará la identidad del órgano competente mediante el empleo de alguno de los instrumentos previstos” (artículo 38).

²⁶ A. Roig Batalla, *Las garantías frente a las decisiones automatizadas. Del Reglamento General de Protección de Datos a la gobernanza algorítmica*, Barcelona, J.M. Bosch, 2020; Y. Hermstrüwer, *Artificial Intelligence and Administrative Decisions*, en *Regulating Artificial Intelligence*, T. Wischmeyer y T. Rademacher (eds.), Cham, Springer, 2020, 199-224.

²⁷ En esta dirección, el Tribunal de Distrito de La Haya en su sentencia de 5 de febrero de 2020 en la que ha anulado la elaboración de perfiles para luchar contra el fraude en la Seguridad Social (SyRI) es de la opinión que “legislación no alcanza el ‘equilibrio justo’ requerido por el CEDH” (fundamento 6.83) y que “la legislación de SyRI es insuficientemente transparente y verificable para concluir que la interferencia con el derecho al respeto a la vida privada que puede implicar el uso de SyRI sea necesaria, proporcional y proporcionada en relación con las finalidades que persigue la norma” (fundamento 6.86). Véase al respecto L. Cotino Hueso, «SyRI, ¿a quién sanciono?»; C. B. Fernández, *Primera sentencia europea que declara ilegal un algoritmo de evaluación de características personales de los ciudadanos*, en *Diario La Ley*, núm. 13 de febrero, 2020.

²⁸ Huergo afirma que “típicamente, este tipo de aplicaciones, algoritmos o modelos algorítmicos sirven para el ejercicio de potestades administrativas regladas” (A. Huergo Lora, *Una aproximación a los algoritmos desde el derecho administrativo*, en *La regulación de los algoritmos*, A. Huergo Lora (ed.), Cizur Menor, Aranzadi-Thomson-Reuters, 2020).

²⁹ J. Valero Torrijos, *El régimen jurídico de la e-Administración*, II ed., Granada, Comares, 2007.

¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?

Bauzá señala que “la actuación automatizada será residual, reservada a aquellos actos sencillos que no requieran una labor de ponderación”³⁰. En una línea parecida, Alamillo y Urios consideran que “la automatización es más viable en el ámbito de las potestades regladas que en el ámbito de las potestades discrecionales”, lo que “llevaría a excluir la actuación administrativa automatizada de todos aquellos supuestos en que hay discrecionalidad por parte de la Administración y a aplicarla solos a supuestos claramente reglados”³¹. Por último, podemos traer a colación a Ponce quien recomienda, específicamente hablando sobre la inteligencia artificial, “la prohibición del uso de la IA en relación con ejercicio de potestades discrecionales”³².

En cambio, otros autores a la vista de los beneficios que pueden derivarse del uso de la inteligencia artificial han aceptado su utilización en las Administraciones públicas. En este sentido, Martín Delgado se ha mostrado partidario de utilizar los medios electrónicos para el ejercicio de potestades discrecionales siempre y cuando sea de baja intensidad, por ejemplo, en el caso de la conocida como discrecionalidad técnica³³. Asimismo, Huergo considera que “es perfectamente posible y válido que la Administración utilice, entre otros factores, predicciones algorítmicas” en el caso de potestades discrecionales “siempre que esté adecuadamente motivada y se hayan respetado las normas procedimentales”³⁴. Otros autores incluso han acuñado la expresión discrecionalidad digital para hacer referencia al

uso de los medios electrónicos en el ejercicio de potestades discrecionales ofreciendo alternativas o reemplazando al decisor a través de la automatización con lo que incrementar significativamente la calidad de las decisiones³⁵. Boix incluso considera que será en las decisiones discrecionales donde la inteligencia artificial puede suponer una ganancia cualitativa superior³⁶.

A nuestro entender, y sin profesar una *vulgar fe en la informática y los ordenadores*³⁷, no cabe una respuesta general a esta cuestión que está vinculada a dos aspectos que trataremos en detalle a continuación. En primer lugar, los elementos que caracterizan la potestad discrecionalidad y la capacidad de la inteligencia artificial de llevar a cabo de manera adecuada el proceso inherente a su ejercicio. En segundo lugar, la garantía de la buena administración cuando se utiliza la inteligencia artificial en el ejercicio de potestades discrecionales. En última instancia que se pueda garantizar la fiabilidad de los algoritmos.

4.1. Discrecionalidad e inteligencia artificial

Se han vertido ríos de tinto a la hora de definir la discrecionalidad tanto desde la perspectiva de la teoría del derecho como desde la perspectiva del derecho administrativo. La simple reproducción de las principales definiciones desbordaría la finalidad de este trabajo.

Por ello, en esta ocasión, simplemente queremos poner de manifiesto que cuando hablamos de decisiones discrecionales estamos ante entornos complejos y de incertidumbre en los que la norma únicamente fija fines atribuyendo a la Administración pública la

³⁰ F. J. Bauzá Martorell, *Identificación, autenticación y actuación automatizada de las administraciones públicas*, en *Tratado de Procedimiento Administrativo Común y Régimen Jurídico Básico del sector público*, E. Gamero Casado, S. Fernández Ramos y J. Valero Torrijos (eds.), Vol. 1, Valencia, Tirant lo Blanch, 2017

³¹ I. Alamillo Domingo y X. Urios Aparisi, *L'actuació administrativa automatitzada en l'àmbit de les administracions públiques. Anàlisi jurídica i metodològica per a la construcció i explotació de tràmits automàtics*, Barcelona, Escola d'Administració Pública de Catalunya, 2011.

³² J. Ponce Solé, *Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico*, en *Revista General de Derecho Administrativo*, núm. 50, 2019. La Ley alemana de procedimiento administrativo prohíbe explícitamente el uso de la inteligencia artificial para dictar actos administrativos cuando exista discrecionalidad o margen de apreciación (*Verwaltungsverfahrensgesetz* § 35a).

³³ I. Martín Delgado, *Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada*. Igualmente, la sentencia del Consejo de Estado número 8472 de 13 de diciembre de 2019 reconoce la posibilidad de utilizar algoritmos en el ejercicio de potestades discrecionales al constatar que no hay razones concretas ni de principio para limitar dicho uso.

³⁴ A. Huergo Lora, *Una aproximación a los algoritmos desde el derecho administrativo*.

³⁵ P. A. Busch y H. Z. Henriksen, *Digital discretion: A systematic literature review of ICT and street-level discretion*, en *Information Polity*, núm. 23(1), 2018.

Sobre esta cuestión, previamente, T. J. Barth y E. Arnold, *Artificial intelligence and administrative discretion: Implications for public administration*, en *The American Review of Public Administration*, núm. 29, 4, 1999; M. Bovens y S. Zouridis, *From street-level to system-level bureaucracies: how information and communication technology is transforming administrative discretion and constitutional control*.

³⁶ A. Boix Palop, *Los algoritmos son reglamentos: La necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la administración para la adopción de decisiones*.

³⁷ Traemos a colación las palabras de García de Enterría y Fernández Rodríguez al considerar que “la existencia de potestades discrecionales es una exigencia indeclinable del gobierno humano: éste no puede ser reducido a una pura monocracia objetiva y neutral, a un simple juego automático de normas, contra lo que en su tiempo espero la entelequia social y política de la Ilustración (y como hoy, en cierto modo, alimenta la más vulgar fe en la informática y los ordenadores)” (E. García de Enterría y T. R. Fernández Rodríguez, *Curso de Derecho Administrativo*, XV ed., Vol. I, Madrid, Civitas, 2008).

facultad de completar o integrar el supuesto de hecho normativo imperfecto³⁸. Para ello, la Administración pública puede tener en cuenta criterios políticos, técnicos o de mera oportunidad o conveniencia³⁹. Este proceso puede exigir a la Administración pública realizar un juicio y una ponderación⁴⁰.

La pregunta que nos surge en este punto es si los algoritmos pueden llevar a cabo estos procesos. Si bien, generalmente se considera que no, antes de dar por zanjada de manera taxativa esta cuestión, es necesario tener presentes algunos elementos.

Hoy por hoy, la inteligencia artificial ya es mucho más eficiente que las personas en entornos de baja complejidad. En cambio, no lo está siendo por el momento en situaciones complejas. Asimismo, la inteligencia artificial ha demostrado ya buenos resultados en entornos de cierta incertidumbre siempre y cuando la complejidad sea baja. Sin embargo, cuando más incertidumbre exista menos espacio existe para la inteligencia artificial⁴¹. Estos entornos se caracterizan por una escasez de datos o porque estos son de baja calidad cuando precisamente es necesaria una gran cantidad de datos de calidad para que los algoritmos puedan obtener resultados óptimos. De este modo, a medida que se incrementa la complejidad y la incertidumbre que define la potestad pública, menos puede aportar la inteligencia artificial, siendo la intervención humana más eficiente.

Por otra parte, los algoritmos aún no son capaces de llevar a cabo, adecuadamente y de manera generalizada, procesos racionales como realizar juicios o ponderaciones complejos si bien ya lo hacen en determinados entornos (por ejemplo, los coches automatizados ya son capaces de realizar algunos de estos procesos)⁴².

³⁸ M. Bacigalupo, *La discrecionalidad administrativa (estructura normativa, control judicial y límites constitucionales de su atribución)*, Madrid, Marcial Pons, 1997.

³⁹ M. Sánchez Morón (ed.), *Discrecionalidad administrativa y control judicial*, Madrid, Tecnos, 1994; M. Beltrán de Felipe, *Discrecionalidad administrativa y constitución*, Madrid, Tecnos, 1995.

⁴⁰ Como recuerda Ponce Solé, “la discrecionalidad consiste en una ponderación de los hechos, intereses y normas relevantes en cada caso y en la elección que se deriva de la misma, elección que es realizada en base a un criterio, el cual no está prefijado la mayor parte de las veces en el ordenamiento, de naturaleza extrajurídica” (J. Ponce Solé, *La lucha por el buen gobierno y el derecho a una buena administración mediante el estándar jurídico de diligencia debida*, Madrid, Universidad de Alcalá-Defensor del Pueblo, 2019).

⁴¹ J.B. Bullock, *Artificial intelligence, discretion, and bureaucracy*, en *The American Review of Public Administration*, núm. 49, 7, 2019.

⁴² Noto La Diega considera que es imposible diseñar un algoritmo que pueda interpretar la norma como lo haría un juez por la ubicuidad y la complejidad de esta tarea relacionada con numerosos factores, algunos de ellos de carácter psicológico y no meramente cognitivo (G. Noto La

En cambio sí que es habitual que desarrollen de manera adecuada procesos analíticos o de razonamiento (informar el procedimiento de toma de decisiones al aportar conocimiento basado en datos a través del uso de algoritmos) o en la tarea de decidir⁴³.

Por todo lo anterior, la inteligencia artificial podría ser utilizada para desarrollar tareas con un bajo nivel de discrecionalidad o cuando el ejercicio de la potestad discrecional supone el uso de criterios técnicos y no políticos⁴⁴.

En cambio, en situaciones con un mayor nivel discrecionalidad o y, por ende, mayor complejidad, incertidumbre o juicios o ponderaciones más complicados o que exijan el uso de criterios no estrictamente técnicos, la inteligencia artificial puede ser útil como apoyo en la toma de decisiones, por ejemplo, analizando grandes volúmenes de datos, previendo sus efectos o recomendando posibles decisiones que puedan resultar relevantes para que el decisor pueda tomar la decisión de mayor calidad.

A medida que se incrementa el nivel de discrecionalidad, las personas son hoy en día decisores más efectivos. En estos supuestos, la inteligencia artificial constituye un aporte para amplificar la inteligencia humana pero no una alternativa válida al decisor humano. Por ejemplo, cuando el decisor debe analizar grandes volúmenes de datos, la inteligencia artificial puede serle de gran ayuda, porque como han advertido Young *et alii*, puede “reducir la dimensionalidad de los datos y el volumen en estructuras, patrones y otras informaciones entendibles por los decisores humanos”⁴⁵. El uso

Diega, *Against the Dehumanisation of Decision-Making – Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information*, en *Journal of Intellectual Property, Information Technology and E-Commerce Law*, núm. 9, 1, 2018). En la misma dirección, existen estudios empíricos que muestran como la inteligencia artificial no es capaz de realizar satisfactoriamente una interpretación jurídica (J. Karaganis y J. Urban, *The rise of the robo notice*, en *Communications of the ACM*, núm. 58, 9, 2015).

⁴³ K. Yeung, *Why worry about decision-making by machine?*, en *Algorithmic Regulation*, K. Yeung y J. Lodge (eds.), Oxford, Oxford University Press, 2019. De hecho, Boix concluye que “los algoritmos empleados por la Administración pública de modo no puramente instrumental producen materialmente los mismos efectos que cualquier reglamento, al pre-ordenar la decisión final del poder público y limitar el ámbito de discreción o de capacidad de determinación de quienes los han de aplicar a partir de los postulados contenidos en la programación. Son por ello, no ya código 2.0, que también, sino más en concreto reglamentos en un sentido jurídico material” (A. Boix Palop, *Los algoritmos son reglamentos*).

⁴⁴ M. M. Young, J. B. Bullock y J. D. Lecy, *Artificial Discretion as a Tool of Governance*; I. Martín Delgado, *Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada*.

⁴⁵ M. M. Young, J. B. Bullock y J. D. Lecy, *Artificial Discretion as a Tool of Governance*.

¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?

de algoritmos también puede evitar que existan sesgos que puedan derivarse de la intervención de las personas y que se puede concretar, por ejemplo, en la existencia de conflictos de intereses.

De todos modos, a pesar de la utilidad de los algoritmos no podemos dejarnos seducir por la inteligencia artificial y sus utilidades y beneficios sin desconocer que los algoritmos también cometen errores y que en ocasiones estos tienen mayores consecuencias o consecuencias más difícilmente identificables que cuando son las personas las que toman decisiones.

En cualquier caso, la decisión sobre qué decisiones discrecionales pueden ser adoptadas por algoritmos estará vinculada al estado de la técnica y debe venir necesariamente presidida por el principio de precaución con el que se persigue dar una respuesta a los riesgos que, aunque inciertos o incompletos, pueda entrañar el uso de la inteligencia artificial por las Administraciones públicas⁴⁶. De este modo, el uso de la inteligencia artificial en la automatización del ejercicio de potestades discrecionales debe ser limitado o rechazado hasta que se pueda probar que no van a causar daño a personas, grupos o normas⁴⁷.

4.2. La intervención humana y el ejercicio de potestades discrecionales a través de algoritmos

El principio de precaución al que acabamos de hacer referencia nos lleva a señalar que a pesar de que determinadas decisiones discrecionales pueden ser adoptadas por

algoritmos, ello no es óbice para que sea fiable.

En esta dirección, en este punto nos interesa valorar si desde la perspectiva del decisor, la intervención humana permite adoptar decisiones de mayor calidad porque permite aportar elementos al proceso de decisión que no puedan aportar los algoritmos (desde empatía y compasión hasta sentido común o experiencia pasando por la capacidad de entender, razonar o ponderar)⁴⁸. Todo ello sin desconocer que los algoritmos pueden tener una mayor capacidad de análisis que las personas o no tener conflictos de intereses que puedan comportar riesgos para la integridad o la imparcialidad.

Una de ellas es la empatía. La empatía es la capacidad de ponerse en el lugar de otras personas y de preocuparse por ellas respondiendo a sus emociones y experiencias⁴⁹.

Según Ponce en el ámbito de decisiones discrecionales debería establecerse “una reserva para la toma de ciertas decisiones a humanos, lo que denominaremos aquí una reserva de humanidad”⁵⁰. Según este autor, el ejercicio de potestades discrecionales implica la necesidad de empatía en la consideración de los hechos, intereses y derechos relevantes a considerar y sopesar y esta solo puede ser ejercida por un ser humano. Para Ponce, el decisor debe tener “la capacidad de ponerse en la piel de los interesados, participantes y posibles afectados por la futura decisión, sea ésta normativa o no, pues solo así podrá ponderar correctamente para el buen ejercicio de la discrecionalidad”⁵¹.

⁴⁶ C. Cierco Seira, *El principio de precaución: reflexiones sobre su contenido y alcance en los Derechos comunitario y español*, en *Revista de Administración Pública*, núm. 163, 2004. Boix señala que “quizás no solo haya que atender al principio de precaución en su declinación más tradicional, sino ir más allá e integrar la propia noción de precaución en una formulación más radical adaptada a todas las necesidades y exigencias del nuevo entorno” (A. Boix Palop, *Los algoritmos son reglamentos*).

De todos modos, no podemos desconocer las dificultades que entraña la aplicación de este principio como ha puesto de manifiesto (G. Doménech Pascual, *Derechos fundamentales y riesgos tecnológicos. El derecho del ciudadano a ser protegido por los poderes públicos*, Madrid, Centro de Estudios Políticos y Constitucionales, 2006).

⁴⁷ A. Thierer y A. Castillo, *Preparing for the Future of Artificial Intelligence*, en *Public Interest Comment*, Arlington, VA, Mercatus Center at George Mason University, 2016.

Así ya se ha reconocido, por ejemplo, en el Código de conducta ética para los ingenieros en robótica anexo a la Resolución del Parlamento Europeo incorpora entre otros el principio de precaución según el cual “Las actividades de investigación en el ámbito de la robótica deben llevarse a cabo de conformidad con el principio de precaución, anticipándose a los posibles impactos de sus resultados sobre la seguridad y adoptando las precauciones debidas, en función del nivel de protección, al tiempo que se fomenta el progreso en beneficio de la sociedad y del medio ambiente”.

⁴⁸ Además, como ha sido ampliamente reconocido, la intervención humana es una garantía de la autonomía de la persona. Al respecto, desde una perspectiva general, téngase en cuenta el *Libro blanco sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza* de 19 de febrero de 2020 [COM(2020) 65 final] en el que se señala que “la supervisión humana ayuda a garantizar que un sistema de IA no socave la autonomía humana o provoque otros efectos adversos”. En la misma dirección se había manifestado previamente la Comisión Europea en su Comunicación *Generar confianza en la inteligencia artificial centrada en el ser humano* de 8 de abril de 2019 [COM(2019) 168 final] al afirmar que “la supervisión humana ayuda a garantizar que un sistema de IA no socave la autonomía humana ni cause otros efectos adversos”. Véase un análisis del proceso de definición de los aspectos éticos de la inteligencia artificial en L. Cotino Hueso, *Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho*, en *Revista Catalana de Dret Públic*, núm. 58, 2019, 29-41.

⁴⁹ S. Dolamore, *Toward a Framework for Detecting Empathy in Public Sector Organizations*, Baltimore, MD, Robert L. Bogomolny Library, 2018.

⁵⁰ J. Ponce Solé, *Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico*, en *Revista General de Derecho Administrativo*, núm. 50, 2019.

⁵¹ J. Ponce Solé, *La lucha por el buen gobierno y el derecho a una buena administración mediante el estándar jurídico de diligencia debida*. En la misma línea, Jensen y Pedersen quienes concluyen que la toma de decisiones discrecionales

Hoy en día, los algoritmos carecen de empatía tal y como se entiende en general⁵². Sin embargo, no podemos desconocer que ya se han diseñado algoritmos que son capaces de analizar emociones por lo que el desarrollo de la inteligencia artificial puede traernos en un plazo más o menos largo algoritmos capaces de *ponerse en la piel de las personas* al adoptar decisiones automatizadas. De todos modos, como advierte Yeung “son substitutos artificiales e inferiores de la auténtica empatía, compasión e interés por aquellos con los que compartimos los vínculos comunes de la experiencia humana”⁵³.

Sin poder en esta ocasión profundizar en otras capacidades propiamente humanas, la exposición que hemos realizado respecto a la empatía nos permite considerar la oportunidad de apostar por la intervención humana como complemento de la automatización de las decisiones públicas. De este modo, aquello que no sea capaz de realizar el algoritmo pueda ser realizado por las personas⁵⁴.

5. La protección de datos personales como límite a la actuación administrativa automatizada

está influida por las habilidades empáticas de los burócratas (D. C. Jensen y L. B. Pedersen, *The impact of empathy explaining diversity in street-level decision-making*, en *Journal of Public Administration Research and Theory*, núm. 27, 3, 2017). Asimismo, se pronuncia Noto La Diega para quien la empatía no es un defecto del proceso de decisión humano sino que tiene una función positiva (G. Noto La Diega, *Against the Dehumanisation of Decision-Making*). Igualmente, Edlins considera que la empatía puede funcionar como un mecanismo de control de calidad de la toma de decisiones discrecionales garantizando que los empleados públicos conectan con las historias de los interesados yendo más allá de los hechos (M. Edlins, *Developing a model of empathy for public administration*, en *Administrative Theory & Praxis*, 2019, 1-20).

⁵² De todos modos, ya se han ido produciendo algunos avances en esta dirección como se recoge en: <https://www.bbvaopenmind.com/humanidades/sociologia/e-n-busca-de-la-nueva-empatia-humano-maquina/> (última consulta: agosto de 2020).

Desde un punto de vista general, Benítez, Escudero, Kanaan, y Masip Rodó observan que una de las cuestiones para las que la neurociencia no ha encontrado aún respuesta es si procesos como la empatía o la consciencia están intrínsecamente ligados al sistema nervioso y si por ello no son reproducibles por la inteligencia artificial (R. Benítez, G. Escudero, S. Kanaan y D. Masip Rodó, *Inteligencia artificial avanzada*, Barcelona, EdiUOC, 2013).

⁵³ K. Yeung, *Why worry about decision-making by machine?*

⁵⁴ En esta dirección, por ejemplo, López de Mántaras observa que “el médico ve cosas que el algoritmo no ve. Y el algoritmo encuentra patrones que el ojo humano no ve. El algoritmo mira los árboles, y el médico el bosque. Por eso, cuando trabajan juntos disminuye mucho el error. En la detección del cáncer de mama, hay estudios que demuestran que el mejor médico tiene un error del 5% o 6% con las mamografías y la ADA, del 6% o 7%, pero que, trabajando juntos, el error es sólo del 0,5%” (Autoritat Catalana de Protecció de Dades, *Inteligència Artificial. Decisions Automatitzades a Catalunya*).

La toma de decisiones automatizadas puede generar diversos riesgos desde el punto de vista de la protección de los datos personales en la medida en que no haya ninguna intervención humana que pueda supervisar o controlar el tratamiento de los datos que comporte el proceso decisional o la misma decisión o que se puedan inferir datos personales a través del tratamiento automatizado de datos personales⁵⁵.

Para evitar las consecuencias negativas que se pueden derivar por el interesado de la toma de decisiones automatizadas, cuando la toma de decisiones discrecionales se base en el tratamiento automatizado de datos personales debe contemplarse lo dispuesto en el RGPD⁵⁶.

En particular, por lo que respecta a la toma de decisiones automatizadas a través de la inteligencia artificial y su impacto en los datos personales, el RGPD dispone que “todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar”⁵⁷. Este derecho se ha reconocido como una prohibición de que se lleven a cabo decisiones automatizadas cuando concurren los elementos a los que nos referiremos posteriormente y no exista una habilitación para ello⁵⁸.

En efecto, el propio RGPD prevé algunas excepciones a dicha prohibición en determinados

⁵⁵ De todos modos, como advierte la Agencia Española de Protección de Datos “en muchos casos no es sencillo determinar si durante una etapa del ciclo de vida de un sistema basado en IA se tratan o no datos personales” (Agencia Española de Protección de Datos, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, 2020).

⁵⁶ Asimismo, resulta de interés lo dispuesto en el artículo 9 Convenio 108+ del Consejo de Europa adoptado por el Consejo de Ministros en su sesión de 18 de mayo de 2018.

⁵⁷ Artículo 22.1 RGPD. La Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPDGDD) no ha desarrollado este derecho. Lo único que ha hecho ha sido integrarlo en el derecho de oposición al tratamiento al disponer que “el derecho de oposición, así como los derechos relacionados con las decisiones individuales automatizadas, incluida la realización de perfiles, se ejercerán de acuerdo con lo establecido, respectivamente, en los artículos 21 y 22 del Reglamento (UE) 2016/679” (artículo 18).

⁵⁸ Como ha reconocido el Grupo de trabajo del artículo 29 (2018), el hecho de que el artículo 22 RGPD se refiera a un derecho no quiere decir que “se aplique solo cuando se invoque de forma activa por parte del interesado”. De ser así, haría esta disposición dependiente de la acción del interesado y quedaría en la mayoría de los casos sin efecto ante su inacción, a menudo determinada por su desconocimiento de lo dispuesto en el artículo 22. Esta posibilidad debilitaría sustancialmente el derecho por lo que hay que entenderlo en última instancia como una prohibición (I. Mendoza y L. A. Bygrave, *The right not to be subject to automated decisions based on profiling EU Internet Law*, Springer, 2017, 77-98; Privacy International, *Data Is Power: Profiling and Automated Decision-Making in GDPR*, 2017).

¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?

supuestos. En particular, el RGPD dispone que no es de aplicación, si la decisión automatizada⁵⁹: 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 y que establezca asimismo medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, o c) se basa en el consentimiento explícito del interesado.

Así pues, cuando concurra alguno de estos supuestos, el responsable del tratamiento podrá adoptar la decisión de manera automatizada sin que el interesado tenga reconocido el derecho a no ser objeto de dicho tratamiento automatizado.

Sin embargo, si bien el RGPD contempla tres posibles excepciones al derecho a no ser objeto de tratamientos automatizados, en el caso de los tratamientos realizados por las Administraciones públicas para la toma de decisiones automatizadas únicamente será de aplicación el segundo. En efecto, en el caso de que el tratamiento lo lleve a cabo una Administración pública y este sea necesario para la tramitación de un procedimiento administrativo es poco probable que el interesado pueda manifestar de manera completamente libre su consentimiento ante el perjuicio que su eventual negativa pueda llegar a tener en la actuación pública⁶⁰. Tampoco es aplicable con carácter general el supuesto relativo a la celebración o la ejecución de un contrato entre el interesado y la Administración pública como responsable del tratamiento.

Ello nos lleva a apuntar la necesidad de que el uso de la inteligencia artificial para la adopción de decisiones públicas cuando suponga el tratamiento de datos personales debe estar autorizado por el Derecho de la Unión o de los Estados miembros que se aplique al responsable del tratamiento. Al respecto, debe tenerse presente que el artículo 8.1 LOPDGGD exige una norma con rango de ley que, en el caso de las decisiones automatizadas adoptadas por las Administraciones públicas, encuentra su acomodo en el artículo 41 LRJSP⁶¹.

Asimismo, el RGPD prevé otras limitaciones a la toma de decisiones automatizadas como que no puede basarse en los datos personales especialmente protegidos relativos al origen étnico o racial, las opiniones políticas, las

convicciones religiosas o filosóficas o la afiliación sindical, y el tratamiento de datos genéticos, datos biométricos dirigidos a identificar de manera unívoca una persona física, datos relativos a la salud o datos relativos a la vida sexual o la orientación sexual de una persona física⁶².

Dados los riesgos que puede generar la toma de decisiones automatizadas, se debe valorar la oportunidad de llevar a cabo un análisis utilizando la evaluación de impacto en la protección de datos personales⁶³. El objetivo es que el responsable del tratamiento pueda identificar, evaluar y, en su caso, minimizar los riesgos que la toma de decisiones automatizada puede generar antes de que surjan⁶⁴.

Hecha esta primera aproximación, a continuación, debemos analizar brevemente el alcance de la prohibición de tomar decisiones automatizadas que impliquen el tratamiento de datos personales. Como se desprende de la lectura del artículo 22 RGPD debemos tener en cuenta tres elementos:

1. Decisiones basadas únicamente en el tratamiento automatizado.

El primer elemento que se debe tener en cuenta es que las decisiones se basen únicamente en un tratamiento automatizado de datos personales. Sin embargo, el RGPD no nos da ningún elemento para determinar que deben limitarse y cuando ésta se considera que se basa únicamente en el tratamiento automatizado⁶⁵.

Los tratamientos únicamente automatizados serían aquellos que no cuentan con la

⁵⁹ Artículo 22.2 RGPD.

⁶⁰ Como ha puesto de manifiesto el Grupo de Trabajo del Artículo 29 (2018) “siempre hay un claro desequilibrio de poder en la relación entre el responsable del tratamiento y el interesado”. Véase al respecto A. Cerrillo i Martínez, *Las características del consentimiento del interesado y su incidencia en el tratamiento de datos en las administraciones públicas*, en *El Consultor de los Ayuntamientos*, núm. III (julio), 2019b.

⁶¹ A. Boix Palop, *Los algoritmos son reglamentos*.

⁶² Artículo 22.3 RGPD en relación al artículo 9, apartado 1. No obstante, cuando el interesado haya dado su consentimiento explícito para el tratamiento de estos datos personales para una o más de las finalidades o el tratamiento sea necesario por razones de un interés público esencial de acuerdo con lo dispuesto el artículo 9, apartado 2, letra a) o g) RGPD y se hayan tomado las medidas adecuadas para proteger los derechos y las libertades y los intereses legítimos del interesado, el tratamiento automatizado sí se podrá realizar tratando datos especialmente protegidos cuando sea necesario para formalizar o ejecutar un contrato entre el interesado y un responsable del tratamiento; esté autorizado por el derecho de la Unión o de los Estados miembros y se base en el consentimiento explícito del interesado (artículo 9.4 RGPD).

⁶³ Artículo 35 RGPD. Véase al respecto (Grupo de Trabajo del artículo 29, 2017c); Agencia Española de Protección de Datos, *Guía para una evaluación de impacto en la protección de datos*, 2014.

⁶⁴ En la lista de los tratamientos susceptibles de ser evaluados de acuerdo con lo dispuesto en el artículo 35.4 RGPD elaborada por la Agencia Española de Protección de Datos se ha incluido, entre otros tratamientos, aquellos que impliquen la toma de decisiones automatizadas o que contribuyan en gran medida a la toma de tales decisiones. Accesible en: <https://www.aepd.es/media/criterios/listas-dpia-es-35-4.pdf> (última consulta: septiembre de 2020).

⁶⁵ L. Edwards y M. Veale, *Enslaving the Algorithm: From a “Right to an Explanation” to a “Right to Better Decisions”?*, en *IEEE Security & Privacy*, núm. 16, 3, 2018.

intervención humana. Ahora bien, ¿cuál es el grado de participación que debe tener un empleado público para que no se pueda considerar que el tratamiento ha sido realizado únicamente de manera automatizada? ¿Ello excluye cualquier intervención humana o si una intervención mínima de una persona, como la mera confirmación de la decisión adoptada por un algoritmo, permitiría superar esta prohibición?

En relación a este aspecto, el Grupo de Trabajo del artículo 29 ha advertido que el responsable del tratamiento no puede inventarse una participación humana. Por ello, si no hay una influencia real, no se podrá entender que ha habido una intervención humana. Es necesario que el responsable del tratamiento tenga capacidad de supervisión sobre la decisión que se toma de manera automatizada. El responsable del tratamiento deberá valorar con qué frecuencia la persona puede variar el sentido de la decisión adoptada por el algoritmo⁶⁶. El grado de participación debe ser significativo y no puede ser simbólico. Por otra parte, también considera que es necesario asegurar que la persona tiene la autoridad y la competencia para cambiar la decisión⁶⁷.

2. Decisiones que produzcan efectos jurídicos en el interesado o la afecten significativamente de manera similar.

El segundo requisito que deben cumplir las decisiones es que produzcan efectos jurídicos en el interesado o le afecten significativamente de manera similar.

En relación a este elemento, surge la duda a la

hora de determinar cuándo se puede considerar que los efectos del afectan significativamente que debe entenderse como equiparable a un efecto jurídico⁶⁸.

Este sería el caso de decisiones que afecten a las circunstancias financieras de una persona; decisiones que afecten el acceso de una persona a los servicios sanitarios; decisiones que denieguen a una persona una oportunidad laboral o que la coloquen en una gran desventaja o decisiones que afecten el acceso de una persona a la educación; decisiones que denieguen una ayuda al que se tenga derecho⁶⁹.

3. Las garantías de los interesados ante una decisión automatizada.

El responsable del tratamiento deberá adoptar las medidas adecuadas para proteger los derechos y las libertades y los intereses legítimos del interesado⁷⁰.

Estas medidas deben incluir como mínimo las siguientes⁷¹:

- el derecho a obtener intervención humana por parte del responsable. Según el Grupo de Trabajo del Artículo 29 “la intervención humana es un elemento clave. Toda revisión debe ser llevada a cabo por una persona con la autorización y capacidad adecuadas para modificar la decisión. El revisor debe llevar a cabo una evaluación completa de todos los datos pertinentes, incluida cualquier información adicional facilitada por el interesado”⁷². El reconocimiento de este derecho es una manera de asegurar la dignidad humana en la medida en que las personas y no los algoritmos son las responsables de protegerse⁷³.

- el derecho a expresar su punto de vista⁷⁴. En

⁶⁶ En esta dirección la sentencia del Tribunal de distrito de La Haya considera que la simple eliminación de falsos positivos no puede entenderse como intervención humana significativa (6.57).

⁶⁷ Grupo de trabajo del Artículo 29 (2018). Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679. En la misma línea se manifestó la Information Commissioner’s Office del Reino Unido al observar que el artículo 22 se aplica a aquellas decisiones automatizadas en que las personas no ejercen una influencia real en la decisión. En palabras de la ICO “una intervención mínima humana sin una influencia real en el resultado no puede ser suficiente para excluir la aplicación del artículo 22.1” (Information Commissioner’s Office, *Big data, artificial intelligence, machine learning and data protection*, 2017). Igualmente, Veale y Edwards afirman que esta limitación se aplicará a algunas decisiones que tengan algún nivel de intervención humana (M. Veale y L. Edwards, *Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling*, en *Computer Law & Security Review*, núm. 34 (2), 2018). Malgieri y Comandé son más precisos al considerar que la acción humana debe ser relevante y como manifestación de las funciones típicas humanas, por ejemplo, implicando capacidades fundamentales de sensibilidad y juicio) (G. Malgieri y G. Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation*, en *International Data Privacy Law*, núm. 7, 4, 2017).

⁶⁸ Para el Grupo de trabajo del artículo 29 (2018), “el límite de importancia debe ser similar al de una decisión que produzca un efecto jurídico” que sean dignas de atención”.

⁶⁹ Grupo de trabajo del Artículo 29 (2018).

⁷⁰ Artículo 22.3 RGPD.

⁷¹ El artículo 9 Convenio 108+ del Consejo de Europa es ilustrativa sobre el alcance que deben tener estas medidas de garantía.

En cualquier caso, es necesario adoptar las medidas suficientes. Al respecto, resulta de interés traer de nuevo a colación la sentencia del Tribunal de instancia de La Haya que entre otros motivos considera que las garantías adoptadas no eran suficientes (6.81 y 6.92).

⁷² Grupo de trabajo del Artículo 29 (2018). Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679.

⁷³ I. Mendoza y L. A. Bygrave, *The right not to be subject to automated decisions based on profiling EU Internet Law*; M. Hildebrandt, *Law as computation in the era of artificial legal intelligence: Speaking law to the power of statistics*, en *University of Toronto Law Journal*, núm. 68, supplement 1, 2018.

⁷⁴ Para Roig, el derecho a expresar su punto de vista tiene un efecto reducido si se tiene en cuenta la dificultad de refutar una decisión automatizada sin una clara explicación de cómo se ha obtenido la decisión y con los conocimientos técnicos que pueda tener la mayoría de las personas (A. Roig Batalla, *Safeguards for the right not to be subject to a decision based solely on automated processing (Article 22*

¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?

esta dirección, en la medida en que la decisión automatizada se inserte en un procedimiento administrativo, el interesado tendrá derecho a formular alegaciones⁷⁵. Para ello, el interesado debe disponer de elementos suficientes para formarse un punto de vista sobre el proceso como el algoritmo ha tomado la decisión. Asimismo, debe garantizarse que este sea tenido en consideración por la Administración pública cuando lo manifieste a través de los canales establecidos.

- el derecho a impugnar la decisión. Contra cualquier resolución o acto de trámite automatizado, el interesado podrá interponer recurso de alzada o potestativo de reposición⁷⁶. Para que este derecho sea efectivo, con carácter previo a su ejercicio es necesario que el interesado conozca los motivos en los que se ha basado la decisión y comprenda cómo se ha tomado.

Además, de estas garantías, de acuerdo con lo que dispone el RGDP, los interesados podrán ejercer su derecho de acceso para poder conocer información significativa sobre la lógica aplicada en la toma de decisiones automatizada, así como la importancia y las consecuencias previstas de dicho tratamiento para el interesado⁷⁷. Asimismo, podrán ejercer el derecho de rectificación cuando los datos personales utilizados en la elaboración del perfil sean inexactos⁷⁸. También podrán ejercer el derecho de supresión de los datos personales⁷⁹.

Junto a estas medidas, varios autores han ido proponiendo otras como las auditorías o las certificaciones para garantizar que los algoritmos no producen discriminaciones y garantizan un trato justo⁸⁰. También la aprobación de códigos de conducta, la creación de comités de ética para analizar el impacto que puede tener la elaboración de perfiles en los derechos de los interesados o el uso de técnicas de anonimización y pseudoanonimización. Incluso se ha propuesto reconocer un nuevo derecho, el derecho a inferencias razonables⁸¹.

GDPR).

⁷⁵ Artículo 53.1.e) LPACAP.

⁷⁶ Artículo 112.1 LPACAP.

Boix considera la naturaleza reglamentaria del algoritmo de lo que se deriva la posibilidad de interponer un recurso directo con el objeto de impugnar el algoritmo en sí y de las bases de la programación o indirecto impugnando el acto que sea consecuencia de su aplicación concreta (A. Boix Palop, *Los algoritmos son reglamentos*).

⁷⁷ Artículo 15 RGPD.

⁷⁸ Artículo 16 RGPD.

⁷⁹ Artículo 17 RGPD.

⁸⁰ B. Mittelstadt, *Auditing for Transparency in Content Personalization Systems*, en *International Journal of Communication*, núm. 10, 2016, 4991–5002; Grupo de trabajo del Artículo 29 (2018).

⁸¹ S. Wachter y B. Mittelstadt, *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI*, en *Columbia Business Law Review*, núm.

6. El uso de la inteligencia artificial y la buena administración

La buena administración ha adquirido carta de naturaleza en los ordenamientos jurídicos europeos bien como principio bien como derecho que debe guiar las actuaciones de las Administraciones públicas en sus relaciones con la ciudadanía⁸². Tal y como ha puesto de manifiesto Ponce, “el derecho a una buena administración introduce una nueva obligación jurídica, inexistente hasta el momento o al menos no explicitada con claridad, de utilidad, consistente en la *debida diligencia* o *el debido cuidado* que debe desplegar decisor administrativo en la ponderación de todos los hechos, intereses y derechos relevantes para la toma de la decisión y en la motivación de esta, con fundamento en el expediente, manteniendo una congruencia racional entre aquellas dos y este último”⁸³.

La buena administración ha de ser tenida en cuenta en el uso de la inteligencia artificial en el ejercicio de potestades discrecionales⁸⁴.

La buena administración tiene distintas manifestaciones en nuestro ordenamiento jurídico. En particular, cuando las Administraciones públicas utilizan la inteligencia artificial en la toma de decisiones deben hacerlo de acuerdo con el debido cuidado en la toma de decisiones, de manera transparente y motivada.

6.1. Inteligencia artificial y debido cuidado en la toma de decisiones

Los algoritmos deben garantizar que en las decisiones adoptadas automáticamente se han ponderado todos los hechos, intereses y derechos relevantes para la toma de la decisión, es decir, utilizando las palabras de Ponce, la *debida diligencia* o *el debido cuidado* en la toma de decisiones⁸⁵.

Para ello, los algoritmos deben cumplir con determinados estándares que permitan garantizar su fiabilidad y que, entre otros extremos, exigen el reconocimiento del derecho de audiencia o la garantía del principio de imparcialidad⁸⁶.

2019 (1), 2019.

⁸² Por todos, el artículo 41 de la Carta de los Derechos Fundamentales de la Unión Europea.

⁸³ J. Ponce Solé, *La lucha por el buen gobierno y el derecho a una buena administración mediante el estándar jurídico de diligencia debida*.

⁸⁴ J. Valero Torrijos, *Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración*.

⁸⁵ J. Ponce Solé, *La lucha por el buen gobierno y el derecho a una buena administración mediante el estándar jurídico de diligencia debida*.

⁸⁶ D. K. Citron y F. Pasquale, *The scored society: due process for automated predictions*, en *Washington Law Review*, núm. 89, 2014; K. Crawford y J. Schultz, *Big data and due process: Toward a framework to redress predictive privacy harms*, en *Boston College Law Review*, núm. 55,

Sin embargo, no se puede desconocer la dificultad de hacerlo porque el uso de los algoritmos puede poner en entredicho estos principios⁸⁷.

6.2. Inteligencia artificial y transparencia

La buena administración exige transparencia.

Sin embargo, cuando las Administraciones públicas utilizan la inteligencia artificial existe un riesgo de opacidad. De hecho, es habitual referirse a los algoritmos de aprendizaje automático como cajas negras (*black boxes*)⁸⁸.

Como hemos tenido oportunidad de analizar en otra ocasión, los algoritmos pueden ser opacos por distintas causas: por causas técnicas (los algoritmos de aprendizaje automático son dinámicos y complejos lo que hace difícil sino imposible conocer su diseño, los datos que han tenido en cuenta para obtener un determinado resultado o qué peso se les ha dado), por causas jurídicas (la legislación de transparencia y acceso a la información pública reconoce diversos límites para evitar perjudicar a determinados bienes o derechos como la seguridad ciudadana, la propiedad intelectual o los datos personales) o por causas organizativas (es habitual que las Administraciones públicas no formalicen las decisiones relativas al uso de los algoritmos o que no dispongan de información relativa a las especificaciones técnicas de un algoritmo o a su código fuente, que quieran evitar que pueda eludirse su uso o que hayan adquirido los algoritmos de terceros)⁸⁹.

Desde el punto de vista de la transparencia, no podemos desconocer que un algoritmo es información pública. Así se ha reconocido desde distintas instancias⁹⁰.

2014; M. Suksi, *Administrative due process when using automated decision-making in public administration: some notes from a Finnish perspective*, en *Artificial Intelligence and Law*, 2020.

En el ámbito anglosajón se ha traído a colación el derecho al proceso debido para asegurar que los interesados puedan participar en las decisiones que les afecten y garantizar que estas no contienen sesgos (K. Yeung, *Algorithmic regulation: a critical interrogation*).

⁸⁷ C. Coglianese y D. Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, en *The Georgetown Law Journal*, núm. 105, 2017.

De hecho, este fue uno de los aspectos analizados en la sentencia del Tribunal Supremo del Estado de Wisconsin de 13 de julio de 2016 *State v. Loomis*. En ella se afirma que “si se utiliza adecuadamente con conciencia de las limitaciones y las cautelas, la consideración de la valoración de riesgo por COMPAS por los tribunales de circuito al dictar sentencia no viola el derecho del defendido al proceso debido”.

⁸⁸ F. Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge, MA, Harvard University Press, 2015.

⁸⁹ A. Cerrillo i Martínez, *How can we open the black box of public administration? Transparency and accountability in the use of algorithms*, en *Revista Catalana de Dret Públic*, núm. 56, 2019a, 13-28.

⁹⁰ Por ejemplo, la *Comisión de Garantía del Derecho de*

Por ello, para conocer los algoritmos que utiliza la Administración pública para la toma de decisiones automatizadas son aplicables los mecanismos de transparencia⁹¹. En primer lugar, la difusión proactiva de los algoritmos a través de portales de transparencia⁹². En segundo lugar, el derecho de acceso a la información pública que deben facilitar las Administraciones públicas.

Sin embargo, no podemos desconocer las limitaciones de estos mecanismos⁹³. Entre otros motivos por la dificultad de determinar qué información en concreto debe difundirse o facilitarse sobre los algoritmos ¿el código fuente, las especificaciones técnicas o una sucinta descripción?

La legislación vigente no da una respuesta clara a esta cuestión por lo que es necesario desarrollar los mecanismos de transparencia en vigor para dar respuesta a los problemas de opacidad vinculados al uso de los algoritmos en la toma de decisiones automatizada⁹⁴.

Acceso a la Información Pública en su resolución de 21 de septiembre de 2016, de estimación de las Reclamaciones 123/2016 y 124/2016 en la que afirma que, aunque se encuentre en lenguaje matemático o informático, un algoritmo no deja de ser un tipo de información que estando en poder de la Administración pública constituye información pública. En Francia, la *Commission d'Accès aux Documents Administratifs* en su resolución 20144578 de 8 de enero de 2015. En Italia, el Tribunal administrativo regional de Lazio (sección tercera) en su sentencia de 14 de febrero de 2017.

⁹¹ Además de los mecanismos previstos en la legislación de transparencia, cabe tener presente lo dispuesto en los artículos 13, 14 y 15 RGPD en relación a las decisiones automatizadas que supongan el tratamiento de datos personales que exigen cuando exista dicho tratamiento que el responsable del tratamiento facilite “información significativa sobre la lógica aplicada, así como la importancia y las consecuencias previstas de dicho tratamiento para el interesado”.

⁹² Si bien no es habitual que las Administraciones públicas difundan proactivamente a través de sus portales de transparencia los algoritmos que utilizan, algunos lo hacen a través de repositorios como GitHub. Véase, como ejemplo, <https://github.com/AjuntamentdeBarcelona> (última consulta: septiembre de 2020).

⁹³ R. Brauneis y E. P. Goodman, *Algorithmic transparency for the smart city*, en *Yale Journal of Law & Technology*, núm. 20, 2017, 104-176; A. Boix Palop, *Los algoritmos son reglamentos*.

⁹⁴ En esta dirección, por ejemplo, podemos recordar que la Ley 30/1992, de 26 de noviembre, de régimen jurídico de las administraciones públicas y del procedimiento administrativo común preveía que las administraciones públicas debían aprobar previamente los programas y aplicaciones electrónicos, informáticos y telemáticos y difundir sus características. No obstante, esta previsión general no se mantuvo en la Ley 11/2007, de 22 de junio, de acceso de los ciudadanos a los servicios públicos, ni tampoco se ha incorporado en la LRJSP. Esta involución ha sido criticada por diversos autores. Así, para Boix “el retroceso producido en 2007 en este punto, y consolidado definitivamente en 2015, es manifiesto” (A. Boix Palop, *Los algoritmos son reglamentos*). Asimismo, J. Valero Torrijos, *Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena*

¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?

La importancia de la transparencia de los algoritmos ya ha sido reconocida por los órganos de garantía de la transparencia por ejemplo el Consejo de Transparencia y Buen Gobierno o la Comisión de Garantía del Derecho de Acceso a la Información Pública⁹⁵.

También, fuera de nuestro país, por algunos tribunales.

Así, en Italia el Consejo de Estado en su sentencia número 2270 de 8 de abril de 2019 afirma se debe garantizar la cognoscibilidad del algoritmo, el procedimiento usado para su elaboración, el mecanismo de decisión, el valor de los datos en el procedimiento de valoración y decisión. Todo ello con la finalidad de poder verificar que la decisión adoptada por el algoritmo sea conforme a las prescripciones y finalidades previstas por la norma o por la Administración pública⁹⁶. Posteriormente, la sentencia número 8472 de 13 de diciembre de 2019 reitera que el conocimiento del algoritmo “debe garantizarse en todos sus aspectos: desde sus autores hasta el procedimiento utilizado para su elaboración, el mecanismo de decisión, incluidas las prioridades asignadas en el procedimiento de valoración y decisión de los datos seleccionados como relevantes”⁹⁷.

Asimismo, en los Países Bajos, el Tribunal de Distrito de La Haya en su sentencia de 5 de febrero de 2020 concluye que la “legislación SyRI es insuficientemente transparente y verificable”⁹⁸. Asimismo, reconoce que a pesar de la información facilitada por el Estado no ha explicado en qué información objetivamente verificable se basan los ejemplos facilitados y “no proporciona información sobre el funcionamiento del modelo de riesgo, como por ejemplo el tipo de algoritmos utilizados en el modelo, ni tampoco facilita información sobre el método de análisis de riesgo”⁹⁹. Concluye el tribunal que es imposible verificar cómo está diseñado el árbol de decisión que utiliza el algoritmo y en qué pasos consiste. Una circunstancia que dificulta que una persona afectada por el mismo pueda defenderse contra el hecho de que se haya realizado un informe de riesgos con respecto a él o ella¹⁰⁰.

administración.

⁹⁵ Además, de las resoluciones citadas de la Comisión de Garantía del Derecho de Acceso a la Información Pública, también se ha pronunciado al respecto el Consejo de Transparencia y Buen Gobierno en su resolución 701/2018 de 18 de febrero de 2019.

⁹⁶ Fundamento 8.3.

⁹⁷ Fundamento 13.1. Véase al respecto E. Carloni, *IA, algoritmos y Administración pública en Italia*, en *Revista Internet, Derecho y Política*, núm. 30, 2020, 1-12.

⁹⁸ Fundamento 6.86. Véase L. Cotino Hueso, «SyRI, ¿a quién sanciono?».

⁹⁹ Fundamentos 6.88 y 6.89.

¹⁰⁰ Fundamento 6.91.

6.3. Inteligencia artificial y motivación. Rendición de cuentas y derecho a la explicación

La motivación es el tercer de los pilares del derecho a la buena administración.

Como es bien conocido, el ordenamiento jurídico dispone la obligación de motivar determinadas decisiones con una sucinta referencia de hechos y fundamentos de derecho entre las que se encuentran aquellas que se dicten en el ejercicio de potestades discrecionales¹⁰¹. De este modo, cuando las Administraciones públicas automatizan la toma de decisiones discrecionales deben poder explicar cómo ha funcionado el algoritmo, qué datos ha utilizado para obtener aquel resultado.

Sin embargo, el uso de algoritmos de aprendizaje automático plantea problemas por lo que respecta a la motivación de las decisiones en la medida en que sea difícil sino imposible conocer el razonamiento que ha seguido el algoritmo para obtener un determinado resultado, los datos que ha utilizado o el peso que ha dado a cada uno de ellos en la decisión adoptada.

La motivación de las decisiones que adoptan las Administraciones públicas utilizando inteligencia artificial debe ser suficiente y congruente. Además, como observa Ponce, según el tipo de decisión será exigible una mayor diligencia y detalle¹⁰². De este modo, cuanto mayor sea el nivel de autonomía del algoritmo y menor la intervención humana más se deberá poder garantizar la motivación de las decisiones adoptadas.

La relevancia de la motivación de los algoritmos ha sido puesta de manifiesto en la sentencia del Consejo de Estado italiano número 2270 de 8 de abril de 2019 donde se considera que la imposibilidad de comprender cómo se ha

¹⁰¹ Artículo 35 LPACAP. Más allá de esta previsión, Capdeferro sugiere que “en el futuro no debería incluirse entre los actos que deben ser motivados aquellos que se aparten del criterio o solución resultante de la aplicación de un sistema de inteligencia artificial avanzado y sofisticado” Ó. Capdeferro Villagrasa, *Las herramientas inteligentes anticorrupción: entre la aventura tecnológica y el orden jurídico*, en *Revista General de Derecho Administrativo*, núm. 50, 2019; Ó. Capdeferro Villagrasa, *La inteligencia artificial del sector público: desarrollo y regulación de la actuación administrativa inteligente en la cuarta revolución industrial*, en *Revista Internet, Derecho y Política*, núm. 30, 2020, 1-14.

¹⁰² J. Ponce Solé, *La lucha por el buen gobierno y el derecho a una buena administración mediante el estándar jurídico de diligencia debida*. Como señala Valero, “la explicación que se proporcione no solo ha de ser razonable, sino que ha de adaptarse a la singularidad de la tecnología; lo que puede llevar a ser más exigentes al requerir elementos de contraste, criterios de decisión relevantes una justificación ex post sino como un criterio inherente al diseño de los sistemas de inteligencia artificial” (J. Valero Torrijos, *Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración*).

tomado una decisión a través de un algoritmo “constituye en sí misma un defecto que invalida el trámite”. En esta dirección insiste posteriormente la sentencia del Consejo de Estado número 8472 de 13 de diciembre de 2019 que exige “pleno conocimiento previo del módulo utilizado y de los criterios aplicados”¹⁰³. Según esta sentencia “este conocimiento del algoritmo debe garantizarse en todos los aspectos: desde sus autores hasta el procedimiento utilizado para su elaboración, el mecanismo de decisión, las prioridades asignadas en el procedimiento de evaluación y toma de decisiones y los datos seleccionados como relevantes”¹⁰⁴. Todo ello “a fin de poder verificar que los criterios, condiciones y resultados del procedimiento robótico cumplen con las prescripciones y las finalidades establecidas por la ley o por la propia administración sobre dicho procedimiento y para que queden claras —y, por lo tanto, sean cuestionables— las modalidades y reglas a partir de las que se haya programado”¹⁰⁵.

La falta de motivación impide a la persona interesada tener elementos para poder impugnar la decisión que le afecta lo que incide negativamente en su derecho a la defensa y le puede ocasionar indefensión¹⁰⁶. Pero más allá del propio interesado, como puso de relieve en su día Fernández Rodríguez, “exigir a la Administración que dé cuenta de sus actos (...) es algo que no sólo interesa al justiciable, sino que importa decisivamente a la comunidad entera”¹⁰⁷. A través de la motivación de la decisión adoptada la Administración pública debe ser capaz de explicar que el algoritmo “tiene razones que sustentan lo decidido”¹⁰⁸.

En los casos en los que la Administración pública no motive su decisión no solo se incurrirá en arbitrariedad sino se estará dando cumplimiento a la buena administración¹⁰⁹. En aras a la buena administración, cuando la Administración pública no pueda motivar la decisión adoptada por un algoritmo no debería adoptarse automáticamente dicha decisión.

En última instancia, la motivación es un mecanismo de rendición de cuentas que fortalece la confianza ciudadana en las Administraciones públicas cuando utilizan algoritmos en la toma

de decisiones. Además, también permite la detección de errores y sesgos¹¹⁰.

La motivación de las decisiones adoptadas a través de algoritmos puede canalizarse mediante la transparencia bien de las decisiones bien de los algoritmos a través de los que estas se toman. En esta dirección, algunos autores han propuesto profundizar en los mecanismos de transparencia a través de los que la pueda Administración pública rendir cuentas de las decisiones adoptadas mediante algoritmos¹¹¹. Sin embargo, tal y como se ha apuntado en el apartado anterior, la transparencia tiene numerosas limitaciones¹¹².

Otros autores sugieren avanzar hacia una justificación de las decisiones adoptadas a través de algoritmos estableciendo una obligación de explicar el algoritmo, sus finalidades y funcionamiento, así como los datos utilizados que permitan su conocimiento por parte de los empleados públicos y, en última instancia, su control por parte de la ciudadanía¹¹³. En esta dirección, se ha propuesto reconocer un derecho a obtener una explicación del uso de los algoritmos para las Administraciones públicas¹¹⁴.

Así se prevé, por ejemplo, la Ley francesa n. 2016-1321, de 7 de octubre de 2016, por una República Digital que modifica el Código de las relaciones entre la ciudadanía y la administración exigiendo informar al interesado cuando se adopta una decisión individual a través de un algoritmo¹¹⁵. En particular, se prevé que cuando así lo solicite, se debe informar sobre las normas que definen el tratamiento y las principales características de su funcionamiento¹¹⁶. También

¹⁰³ Fundamento 12.

¹⁰⁴ Fundamento 13.1.

¹⁰⁵ Fundamento 13.1.

¹⁰⁶ M. M. Fernando Pablo, *La motivación del acto administrativo*, Madrid, Tecnos, 1993.

¹⁰⁷ T.-R. Fernández Rodríguez, *Juzgar a la Administración contribuye también a administrar mejor*, en *Revista Española de Derecho Administrativo*, 76, 1992.

¹⁰⁸ J. Igartua Salaverria, *Discrecionalidad, arbitrariedad y control judicial*, en *Revista Vasca de Administración Pública*, núm. 46, 1996.

¹⁰⁹ J. Igartua Salaverria, *Discrecionalidad, arbitrariedad y control judicial*; T.-R. Fernández Rodríguez, *De la arbitrariedad de la Administración*, Madrid, Civitas, 1994.

¹¹⁰ F. Doshi-Velez y M. Kortz, *Accountability of AI Under the Law: The Role of Explanation*, en *Berkman Klein Center Working Group on Explanation and the Law*, Berkman Klein Center for Internet & Society working paper, 2017; B. Lepri, N. Oliver, E. Letouze, A. Pentland y P. Vinck, *Fair, Transparent, and Accountable Algorithmic Decision-making Processes*, en *Philosophy & Technology*, núm. 31, 3, 2017.

¹¹¹ N. Diakopoulos, *Accountability in algorithmic decision making*, en *Communications of the ACM*, núm. 59, 2, 2016; B. Lepri, N. Oliver, E. Letouze, A. Pentland y P. Vinck, *Fair, Transparent, and Accountable Algorithmic Decision-making Processes*.

¹¹² En la misma dirección, M. Ananny y K. Crawford, *Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability*, en *New Media & Society*, núm. 20, 3, 2018.

¹¹³ D. Lehr y P. Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, en *University of California Davies Law Review*, núm. 51, 2017.

¹¹⁴ En la misma dirección, J. Ponce Solé, *Inteligencia artificial, Derecho administrativo y reserva de humanidad*. Este autor apunta que en caso de vulnerarse el estándar de debida diligencia en la comprensibilidad, se podría alegar la imposibilidad de cumplir con lo establecido y solicitar la anulación judicial de la decisión.

¹¹⁵ Artículo L. 311-3-1.

¹¹⁶ El Decreto 2017-330 de 14 de marzo relativo a los derechos de las personas sujetas a decisiones individuales

dispone que las Administraciones públicas deben difundir en Internet las reglas que definen los principales tratamientos a través de algoritmos utilizados en el cumplimiento de sus funciones sobre los que se basen las decisiones individuales¹¹⁷.

Otros autores proponen desarrollar una inteligencia artificial explicable que permita explicar el modelo, el resultado, examinar la caja negra y crear una caja blanca o transparente¹¹⁸.

Asimismo, se ha propuesto que las Administraciones públicas rindan cuentas sobre las decisiones adoptadas a través de algoritmos mediante la evidencia empírica que les permite medir el funcionamiento de los algoritmos y demostrar sus beneficios o impactos con los que justificar las decisiones tomadas sin la necesidad de utilizar una explicación del algoritmo¹¹⁹.

En última instancia, la supervisión humana del uso de algoritmos puede ser una buena manera de rendir cuentas¹²⁰.

7. La supervisión del uso de inteligencia artificial en el ejercicio de potestades discrecionales

A la vista de lo anterior, el uso de la inteligencia artificial en el ejercicio de potestades públicas debe llevarse a cabo con supervisión¹²¹.

adoptadas sobre la base de un tratamiento algorítmico prevé que la Administración pública debe informar de manera inteligible información relativa al grado y modo de contribución del tratamiento automatizado a la toma de decisiones; los datos tratados y su fuente; los parámetros del tratamiento y su ponderación y las operaciones efectuadas (artículo 1).

¹¹⁷ Artículo L. 312-1-3. El Consejo Constitucional francés en su decisión 2018-765 de 12 de junio, ha afirmado que “cuando los principios de funcionamiento de un algoritmo no puedan comunicarse sin infringir alguno de los secretos o intereses previstos en el apartado 2 del artículo L. 311-5 del código de relaciones entre el público y administración, no se puede tomar ninguna decisión individual sobre la base exclusiva de este algoritmo” (fundamento 70).

Véase al respecto P. Coppolani, *Le décret n° 2017-330 du 14 mars 2017 relatif aux droits des personnes faisant l'objet de décisions individuelles prises sur le fondement d'un traitement algorithmique: vers un victoire à la pyrrhus?*, en *I2D – Information, données & documents*, núm. 54, 4, 2017, 16-21.

¹¹⁸ R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti y D. Pedreschi, *A survey of methods for explaining black box models*, en *ACM Computing Surveys (CSUR)*, núm. 51, 5, 2018.

¹¹⁹ A. Cerrillo i Martínez, *Com obrir les caixes negres de les Administracions públiques?*; F. Doshi-Velez y M. Kortz, *Accountability of AI Under the Law: The Role of Explanation*, en *Berkman Klein Center for Internet & Society working paper*, publicación del Berkman Klein Center Working Group on Explanation and the Law, 2017.

¹²⁰ J. Goldenfein, *Algorithmic Transparency and Decision-Making Accountability*.

¹²¹ En esta dirección el Consejo Constitucional francés en su decisión 2018-765 de 12 de junio afirma que “el responsable del tratamiento debe asegurarse de que el tratamiento algorítmico y sus desarrollos están bajo control para poder explicar, en detalle y de forma inteligible, al interesado la forma en que se ha llevado a cabo el

La intensidad de la supervisión estará estrechamente vinculada a la concurrencia de los distintos elementos que hemos ido identificando a lo largo de estas páginas.

La supervisión humana puede realizarse por las personas que adoptan decisiones sobre la base de la propuesta elaborada por un algoritmo. Estaríamos ante un modelo de decisión con intervención humana que consistiría en que las Administraciones públicas utilizan la inteligencia artificial como apoyo o ayuda a los decisores públicos. En estos casos es necesario asegurar que la persona dispone de otros datos que puedan facilitarle elementos para cuestionar la propuesta realizada por el algoritmo. De no ser así, su intervención podría reducirse a ratificar la decisión propuesta. Es necesario que se garanticen las circunstancias que permitan a la persona no solo adoptar substancialmente la decisión sino también identificar los errores o los sesgos que pueda contener la decisión propuesta por el algoritmo.

La supervisión humana también se puede llevar a cabo en aquellos casos en los que la inteligencia artificial se utilizar para tomar de manera automatizada decisiones públicas con plenos efectos jurídicos pero las decisiones son monitorizadas por personas pero sin que estas puedan ser frenadas¹²².

La supervisión se puede llevar a cabo a través de distintos mecanismos¹²³.

En primer lugar, a través de la participación humana en el diseño del algoritmo. En segundo lugar, la supervisión puede canalizarse utilizando la inteligencia artificial únicamente como apoyo a la toma de decisiones. En tercer lugar, mediante el control humano de las decisiones que adopta el algoritmo y la evaluación del funcionamiento del algoritmo. Finalmente, la supervisión humana puede concretarse en el seguimiento de la actividad desarrollada por la inteligencia artificial en su conjunto para poder decidir cómo y cuándo se utilizan los algoritmos.

El nivel de supervisión debe estar estrechamente relacionado con la potestad que se

tratamiento con respecto a él. No podrán ser utilizados como fundamento exclusivo de una decisión administrativa individual, los algoritmos capaces de revisar por sí mismos las reglas que aplican, sin el control y la validación del responsable del tratamiento” (fundamento 71).

Asimismo, también se ha manifestado la UNESCO cuando exige que “siempre debería ser posible atribuir la responsabilidad ética y jurídica de la investigación, el diseño, el desarrollo, el despliegue y la utilización de los sistemas de IA a una persona física o a una entidad jurídica existente”.

¹²² Véase al respecto, M. Wieringa, *What to account for when accounting for algorithms: a systematic literature review on algorithmic accountability*, en *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020.

¹²³ Grupo de expertos de alto nivel sobre inteligencia artificial, *Directrices éticas para una IA fiable*.

deba ejercer y con el desarrollo de la inteligencia artificial. De este modo, las Administraciones públicas deben concretar el nivel de supervisión en función de aspectos como las decisiones que se automatizarán, el grado de automatización o los riesgos que pueda entrañar el algoritmo utilizado. Cuando menor sea la supervisión *ex ante* mayor deberá ser la capacidad de supervisión *ex post*.

Más allá de la supervisión inmediata del uso de la inteligencia artificial por parte de las personas, consideramos que sería oportuno impulsar una supervisión institucional del conjunto de la actividad que las Administraciones públicas lleven a cabo a través de algoritmos.

En esta dirección, tal y como hemos propuesto en otra ocasión, consideramos que sería oportuno crear una gobernanza de la inteligencia artificial¹²⁴. Este marco institucional debería estar diseñado adecuadamente para garantizar la fiabilidad, o sea, la seguridad – jurídica y tecnológica–, la transparencia y la rendición de cuentas en el uso de los algoritmos por las Administraciones públicas. Este marco institucional se podría concretar con la creación de un organismo que velara por la incorporación de la inteligencia artificial en las Administraciones públicas y por la transparencia de los algoritmos y el control de su diseño, funcionamiento y aplicación. Este organismo debería garantizar la participación de todos los implicados en el desarrollo de la inteligencia artificial en las Administraciones públicas¹²⁵.

¹²⁴ A. Cerrillo i Martínez, *Com obrir les caixes negres de les Administracions públiques?*

¹²⁵ Esta propuesta está en sintonía con las propuestas formuladas por distintas instituciones como el Parlamento Europeo, en su resolución de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica o la Office of Science and Technology Policy de la Casa Blanca norteamericana Executive Office of the President National Science and Technology Council Committee on Technology, *Preparing for the Future of Artificial Intelligence*, 2016, o la Science and Technology Committee. House of Commons, *Robotics and artificial intelligence*, 2016.

También por distintos autores, entre otros, S. Barocas y A. D. Selbst, *Big data's disparate impact*, en *California Law Review*, núm. 104, 2016; I. Martín Delgado, *Automazione, intelligenza artificiale e pubblica amministrazione: vecchie categorie concettuali per nuovi problemi?*, en *Istituzioni del federalismo. Rivista di studi giuridici e politici*, núm. 3, 2019; B. Mittelstadt, *Auditing for Transparency in Content Personalization Systems*, en *International Journal of Communication*, núm. 10, 2016 ; J. Ponce Solé, *Inteligencia artificial, Derecho administrativo y reserva de humanidad*; A. Roig Batalla, *Las garantías frente a las decisiones automatizadas. Del Reglamento General de Protección de Datos a la gobernanza algorítmica*, Barcelona, J.M. Bosch, 2020; J. Valero Torrijos, *Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración*, 91; C. Velasco Rico, *Vigilando al algoritmo. Propuestas organizativas para garantizar la transparencia*, en *El derecho ante la transformación digital. Oportunidades, riesgos y garantías*,

Este organismo debería desarrollar diversas funciones como impulsar y facilitar la incorporación de los algoritmos a las Administraciones públicas (por ejemplo, a través de la información, la formación y el asesoramiento sobre el uso de los algoritmos, del diseño de algoritmos para la toma de decisiones públicas o para la prestación de servicios públicos o de su reutilización); controlar el uso de los algoritmos por las Administraciones públicas para garantizar su seguridad y legalidad; auditar y/o certificar los algoritmos a utilizar por las Administraciones públicas con el fin de evitar errores o sesgos en los algoritmos o evaluar el impacto en derechos fundamentales como la protección de los datos personales; evaluar el impacto y eficacia de los algoritmos. Este organismo también podría controlar la legalidad de las decisiones tomadas por las Administraciones públicas a través de los algoritmos bien resolviendo directamente las reclamaciones formuladas por los interesados, bien emitiendo un informe en los recursos que los interesados puedan interponer contra las resoluciones adoptadas a través de algoritmos.

En última instancia, a medida que se vaya extendiendo el uso de la inteligencia artificial será más importante fortalecer los mecanismos de supervisión para, en última instancia, poder garantizar un control adecuado de la actividad administrativa a través de algoritmos¹²⁶.

8. Reflexiones finales: el principio de fiabilidad de los algoritmos como garantía

El uso de la inteligencia artificial en la toma de decisiones públicas no puede realizarse de cualquier manera puesto que ello puede generar problemas diversos como errores, sesgos o incumplimientos de los principios de protección de datos personales.

La mayoría de estos problemas ya han ido siendo identificados. A lo largo de estas páginas hemos podido describir distintos mecanismos

B. Puentes Cociña y A. Quintiá Pastrana (eds.), Barcelona, Atelier, 2019.

¹²⁶ En esta dirección, Auby también ha puesto de manifiesto que “el juez no estará en general más preparado que el ciudadano medio para comprender los algoritmos”. Además, observa que “los instrumentos que utiliza habitualmente para controlar los motivos de la decisión administrativa y la relación de la decisión con los motivos puedan mantener su eficacia habitual” porque se dirigen a racionalidades clásicas y causales (J.-B. Auby, *Contrôle de la puissance publique et gouvernance par algorithme*, en *Das öffentliche Recht vor den Herausforderungen der Informations- und Kommunikationstechnologien jenseits des Datenschutzes/ Information and Communication Technologies Challenging Public Law, Beyond Data Protection/ Le droit public au défi des technologies de l'information et de la communication, au-delà de la protection des données*, D.-U. Galetta y J. Ziller (eds.), Baden-Baden, Nomos Verlagsgesellschaft mbH & Co. KG, 2018).

¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?

previstos en nuestro ordenamiento jurídico para evitarlos y garantizar la eficacia, calidad o transparencia de las decisiones que se adoptan a través de algoritmos.

No obstante, no podemos ignorar las limitaciones que tiene hoy en día su regulación que, en términos generales, no ha sido concebida para dar respuesta a la incorporación de la inteligencia artificial en las Administraciones públicas.

Por ello, desde diferentes instancias ya se han hecho algunas propuestas para reforzar estos mecanismos se dar respuesta a los problemas identificados. Como hemos visto, algunas de ellas pasan por profundizar en la aplicación de la legislación vigente (por ejemplo, en relación a la transparencia de los algoritmos), por aplicar las garantías procedimentales existentes a los algoritmos (por ejemplo, para la elaboración de normas reglamentarias), por limitar el uso de la inteligencia artificial respecto a determinadas potestades (por ejemplo, prohibiendo su uso en el ejercicio de potestades discrecionales) o por reconocer nuevos derechos de la ciudadanía frente a la toma de decisiones a través de algoritmos.

De manera paralela a estas propuestas, a nuestro entender, el uso de la inteligencia artificial en la toma de decisiones en las Administraciones públicas debería ir acompañado del fortalecimiento de la fiabilidad de los algoritmos, es decir, de la garantía de su licitud y seguridad y de su capacidad para obtener buenos resultados, es decir, decisiones públicas de calidad¹²⁷.

La fiabilidad de los algoritmos exigiría garantizar la licitud de los algoritmos lo que se concretaría en la aprobación por la Administración pública de las especificaciones técnicas de los algoritmos identificando los mecanismos de supervisión y control y valorando su impacto en los principios de funcionamiento de las Administraciones públicas y en los derechos y libertades de las personas. Este proceso podría ser realizado por cada órgano administrativo competente para la toma de las decisiones con carácter previo a la utilización del algoritmo o podría articularse a través de la intervención de los organismos creado *ad hoc* para garantizar la seguridad -jurídica y

tecnológica-, la transparencia y la rendición de cuentas de las decisiones adoptadas a través de algoritmos tal y como se ha propuesto anteriormente.

Asimismo, la fiabilidad de los algoritmos se concretaría en la concurrencia de distintos elementos¹²⁸.

En primer lugar, la supervisión humana de la inteligencia artificial. La supervisión se podría realizar por la persona titular del órgano competente para tomar una decisión sobre la base de la propuesta formulada por un algoritmo o previendo el control *a posteriori* de la decisión adoptada por el algoritmo. La supervisión también podría realizarse a través de la participación de los interesados en las distintas fases del ciclo de vida del algoritmo (desde el diseño del algoritmo al control judicial de la decisión adoptada por un algoritmo pasando por la evaluación del algoritmo). En última instancia, la supervisión se podría articular a través del organismo al que ya nos hemos referido. La intensidad de la supervisión debería estar vinculada al tipo de procedimiento de decisión al que se incorporase la inteligencia artificial, al tipo de decisión que se adopte así como al nivel de automatización de la decisión.

En segundo lugar, la solidez de los algoritmos. Para ello, sería necesario que se previese un procedimiento para el diseño y adopción del algoritmo. Este procedimiento debería garantizar la solidez y exactitud del algoritmo, la ausencia de errores o sesgos, la seguridad del algoritmo y su adecuación a la legislación vigente. En particular, debería garantizar la transparencia del algoritmo y, entre otros extremos, el respeto a la normativa de protección de datos.

En tercer lugar, la transparencia de los algoritmos que utilizan las Administraciones públicas. En primera instancia, las Administraciones públicas deberían difundir de manera proactiva los algoritmos que utilizan facilitando una descripción suficientemente detallada de su diseño, los datos que utiliza o su contribución en la toma de decisiones. También se debería informar de las principales especificaciones técnicas y los datos utilizado se, incluso, el código fuente. En esta dirección, las Administraciones públicas deberían fomentar el uso de programario abierto o la declaración de fuentes abiertas en aquellos algoritmos que diseñen.

En cuarto lugar, la prohibición de discriminación. Para ello sería necesario que las

¹²⁷ La fiabilidad ha sido reconocida como principio ético de la inteligencia artificial. Así, según las Directrices éticas para una IA fiable la fiabilidad del algoritmo consiste en la licitud del uso de la inteligencia artificial; en su eticidad y en su robustez (Grupo de expertos de alto nivel sobre inteligencia artificial, 2019). Asimismo, entre otros, recoge el principio de fiabilidad de los algoritmos la Carta catalana para los derechos y las responsabilidades digitales que está en proceso de elaboración. Accesible en: https://participa.gencat.cat/uploads/decidim/attachment/file/790/Cartacatalanaperalsdretsilesresponsabilitatsdigitals_v1.pdf (última consulta: septiembre de 2020).

¹²⁸ Estas características se basan en Grupo de expertos de alto nivel sobre inteligencia artificial, *Directrices éticas para una IA fiable*, 2019, y Comisión Europea, *Libro Blanco sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza COM(2020) 65 final*, 2020.

Administraciones públicas adoptasen medidas para evitar los sesgos en los algoritmos, garantizar la calidad de los datos que utilizan los algoritmos, diseñar algoritmos que tengan particularmente en cuenta posibles discriminaciones, abrir el diseño de los algoritmos a la participación ciudadana, elaborar auditorías que persigan identificar sesgos u otros problemas que puedan estar en el código de los algoritmos o en los datos que estos usan o elaborar análisis de riesgos.

En quinto lugar, la fiabilidad comportaría la rendición de cuentas sobre el uso de los algoritmos. Las administraciones públicas deberían facilitar una explicación sobre el funcionamiento de los algoritmos y una motivación de los resultados obtenidos. En esta dirección, como se ha propuesto, sería recomendable avanzar hacia un derecho a obtener una explicación del uso de los algoritmos por las Administraciones públicas. En particular, las decisiones adoptadas por algoritmos o con el uso de algoritmos deberían informar de manera explícita e inteligible al interesado de esta circunstancia, indicando la finalidad perseguida y el derecho a conocer las reglas que definen el tratamiento y las principales características de su aplicación y cómo se ha de ejercer y los mecanismos de supervisión y control existentes.

Finalmente, cuando los algoritmos tratasen datos personales, la fiabilidad también requeriría la adopción de las medidas necesarias para la gestión de la privacidad y de los datos personales de acuerdo con lo dispuesto en el RGPD. Por un lado, el uso de algoritmos en la toma de decisiones automatizada que suponga un tratamiento de datos personales debería estar explícitamente autorizado por una norma con rango de ley y no podría basarse en datos personales especialmente protegidos excepto en los casos previsto en el RGPD y cuando se hubiesen tomado las medidas adecuadas para proteger los derechos, libertades e intereses legítimos del interesado. Por otro lado, se deberían articular las medidas para garantizar que el interesado puede expresar su punto de vista y puede impugnar la decisión adoptada. En cualquier caso, sería necesario que el interesado conociese la existencia de una decisión automatizada y que pudiese tener acceso a la información significativa información significativa sobre la lógica aplicada, así como la importancia y las consecuencias previstas de dicho tratamiento para el interesado.

La fiabilidad de los algoritmos debería incluirse entre los principios de funcionamiento de las Administraciones públicas y de este modo los elementos que lo integran podrían guiar el proceso de incorporación de la inteligencia artificial en las entidades públicas y dar una respuesta a aquellas lagunas que entrañe la

interpretación del ordenamiento jurídico ante el desarrollo vertiginoso de esta tecnología disruptiva. La falta de fiabilidad de un algoritmo debería determinar la nulidad de la decisión adoptada.