



Decisiones automatizadas y transparencia administrativa: nuevos retos para los derechos fundamentales

Inmaculada Jiménez-Castellanos Ballesteros¹

Universidad de Sevilla
España

ORCID: [0000-0001-9514-2869](https://orcid.org/0000-0001-9514-2869)

RECIBIDO: 4 de octubre de 2022
ACEPTADO: 11 de febrero de 2023

RESUMEN: En la sociedad tecnológica actual, las Administraciones Públicas, al igual que el sector privado, han generalizado el uso de la inteligencia artificial para la toma de decisiones. Los algoritmos, que pueden procesar cantidades ingentes de datos y ayudar por la precisión de sus resultados, no obstante, plantean nuevos retos para los derechos fundamentales. En primer lugar, por la escasa información que se tiene sobre su funcionamiento, lo que supone una quiebra del principio de transparencia. En segundo lugar, por su complejidad, lo que dificulta su sujeción a medidas de control y complica la motivación de los actos administrativos automatizados, limitando la capacidad de los ciudadanos para recurrir las decisiones de la Administración. Y, por último, por lo influyentes que son: su empleo inadecuado puede desembocar en situaciones discriminatorias, así como lesionar otros derechos fundamentales de los ciudadanos.

PALABRAS CLAVE: Algoritmos, Administración, transparencia, control, discriminación.

CONTENIDOS: 1.- Introducción. — 2.- Concepto de algoritmo. — 3.- La opacidad de los algoritmos utilizados en las decisiones automatizadas: la transparencia administrativa. — 4.- El control de los algoritmos empleados por las Administraciones Públicas. — 4.1.- *La respuesta europea.* — 4.2.- *El control judicial de las decisiones automatizadas en el sector público.* — 5.- Los efectos discriminatorios de los algoritmos. — 6.- Conclusiones. — Bibliografía.

¹ Profesora en el departamento de Derecho Constitucional de la Facultad de Derecho de la Universidad de Sevilla. Es licenciada en Derecho, Master en Derecho Constitucional y doctora en Derecho. Su principal línea de investigación es el derecho fundamental a la protección de datos personales, materia sobre la que versan varias de sus publicaciones en artículos indexados, obras colectivas y en una monografía.



Automated decisions and administrative transparency: new challenges for fundamental rights

ABSTRACT: In today's technological society, the use of artificial intelligence for decision making, both in public and private sectors, has become widespread. Algorithms can help process large amounts of information, but they also pose new challenges to fundamental rights. Firstly, because of the limited information available on how they work, which undermines the principle of transparency. Secondly, because of their complexity, which makes it difficult to subject them to control measures, as well as complicating the motivation of automated administrative acts, limiting the ability of citizens to appeal against the administration's decisions. And finally, because of their influence: their inappropriate use can lead to discriminatory situations, as well as infringe other fundamental rights of citizens.

KEYWORDS: Algorithms, Administration, transparency, control, discrimination.

1.- Introducción

Las Administraciones Públicas están experimentando un cambio paralelo al desarrollo de las nuevas tecnologías. Desde este punto de vista, el primer paso, la implantación de la administración electrónica, ha sido fundamental para mejorar la calidad y la eficacia de los servicios públicos dirigidos al ciudadano. Actualmente, el uso de la inteligencia artificial (IA) y el manejo masivo de datos (Big data) para lograr resultados muy precisos y objetivos en la toma de decisiones es un fenómeno imparable, que ha tenido un papel relevante en la respuesta a la pandemia provocada por el Covid-19. Al mismo tiempo, el recurso a la IA debe llevarse a cabo con prevención, como se ha puesto de manifiesto en otros ámbitos².

² El uso de la inteligencia artificial en las redes sociales por el sector privado ha puesto en peligro los pilares del Estado democrático al interferir en el proceso de toma de decisiones políticas como se puso



El Big data puede analizarse para diseñar políticas públicas o resolver solicitudes formuladas por los ciudadanos; e incluso puede servir para ejercer directamente potestades administrativas de forma automatizada (Velasco Rico, 2022: 230). Estos sistemas algorítmicos de procesamiento de datos pueden resultar muy útiles. En efecto, el uso de la IA permite que las Administraciones Públicas puedan tener un mayor conocimiento del contexto, valorar de manera más rápida los antecedentes y los precedentes, anticipar el impacto, tomar decisiones de manera automatizada o evaluar su efectividad. De este modo, la utilización de la IA puede dotar de mayor eficacia o calidad a las decisiones que aquellas adoptan (Cerrillo Martínez, 2020: 17).

Sin embargo, estos sistemas tecnológicos entrañan importantes riesgos y retos para los derechos fundamentales y para los valores democráticos. Hasta ahora, habían sido las normas sobre protección de datos las encargadas de la tutela del ciudadano. Como bien apunta la Agencia Española de Protección de Datos (AEPD), resulta necesario tener en cuenta que el uso y el tratamiento de datos personales en las Administraciones Públicas mediante la implementación de la IA implican riesgos distintos frente a aquellos derivados de un tratamiento que pueda realizar cualquier otro responsable. En este sentido, tales amenazas resultan del volumen de sujetos afectados, de la extensión de los datos recogidos, de la imposibilidad, en muchos casos, de oponerse al tratamiento, y del poder o asimetría inherente que existe en las relaciones entre la Administración Pública y los ciudadanos o interesados cuyos datos se tratan³.

El tema de la igualdad de trato y la no discriminación en el ámbito de la IA y toma de decisiones automatizadas se ha plasmado en el articulado de la ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación (BOE de 13 de julio), que regula su uso por el sector público y también por las personas físicas y jurídicas de carácter privado. Según esta norma⁴, las Administraciones Públicas que utilicen algoritmos para la toma de decisiones favorecerán la puesta en marcha de mecanismos para que dichos algoritmos tengan en cuenta criterios de minimización de sesgos, transparencia y rendición de cuentas, al tiempo que abordarán su potencial impacto discriminatorio.

2.- Concepto de algoritmo

Los algoritmos son una manifestación de la IA. El Grupo de Expertos de Alto Nivel sobre Inteligencia Artificial (AIHLEG), creado por la Comisión Europea en 2018, define los sistemas de inteligencia artificial como: "aquellos sistemas de software (y posiblemente también de hardware), diseñados por seres humanos que, dado un

de manifiesto en el escándalo de Cambridge Analytica, el referéndum sobre el Brexit y la elección presidencial de Donald Trump.

³ AEPD (2020), Tecnologías y protección de datos en las Administraciones Públicas. Puede consultarse en: <https://www.aepd.es/es>

⁴ Art. 23 Ley 15/2022 de 12 de julio integral de igualdad de trato y de no discriminación.



objetivo complejo, actúan en la dimensión física o digital percibiendo su entorno a través de la adquisición de datos, interpretando los datos estructurados o no estructurados recopilados, razonando sobre el conocimiento o tratamiento de la información fruto de estos datos y decidiendo las mejores acciones que se llevarán a cabo para alcanzar el objetivo fijado"⁵.

La IA que se está desarrollando actualmente se nutre de datos y aprende de ellos, usando algoritmos para dar respuesta a una pregunta formulada con anterioridad, lo que reemplaza al razonamiento humano. Según Huergo Lora (2020:27), el concepto de algoritmo nos remite a "cualquier procedimiento formalizado en una serie de pasos para solucionar un problema o conseguir un resultado". El elemento clave de estos sistemas es el empleo de datos masivos con los que operan y de los que dependen.

El concepto de toma de decisiones asistida por algoritmos en el ámbito administrativo, siguiendo a Velasco Rico (2022:232), se refiere al uso de estas herramientas por parte de actores públicos en el ámbito de sus competencias para la adopción de decisiones de carácter general (nueva reglamentación, nueva política pública, etc.), o bien para la toma de decisiones de carácter individual (resoluciones particulares, actos administrativos, etc.). La toma de decisiones asistida por algoritmos incluye, pero no se identifica exclusivamente, con aquellos supuestos en los que el software o el sistema actúa y decide al margen de la intervención humana. Además, dicho concepto integra los modelos computacionales y los programas informáticos de apoyo a la toma de decisiones, sin que el sistema sea autónomo. Todas estas herramientas analíticas comparten procesos de acopio y tratamiento de datos, la búsqueda de relaciones (de patrones) –muchas veces no obvias– entre los datos y el análisis sistemático de esas relaciones, para responder a una pregunta.

Hoy en día, los algoritmos más destacados son aquellos que, de forma automática, crean un modelo sobre la base de lo que aprenden de los datos con los que han sido entrenados para acabar tomando decisiones de forma autónoma. Estos son los denominados sistemas de machine learning o de aprendizaje automático. En este tipo de algoritmos, los programadores no introducen las reglas para la toma de decisiones, sino que es el propio sistema el que las crea. Como apunta Cerrillo (2019:4) "a partir de los datos disponibles, el algoritmo identifica patrones y correlaciones entre los datos que son utilizados posteriormente para generar un modelo cuya calidad es evaluada estadísticamente y, en su caso, mejorada a partir de los resultados que se van obteniendo. Posteriormente, el algoritmo es utilizado sobre nuevos datos".

⁵ AI HLEG. 2018. *A definition of AI: main Capabilities and Disciplines*. Brussels: European Commission, p. 6. Disponible en: <https://digital-strategy.ec.europa.eu/es>. La definición que maneja la propuesta de Reglamento de Inteligencia artificial es mucho más restrictiva. Se recoge en el artículo 3 al que veremos más adelante.



En España se ha generalizado el uso de sistemas algorítmicos de apoyo para la toma de decisiones en la Administración Pública en los más diversos sectores, como en los procesos selectivos, en la sanidad o en el ámbito policial así los sistemas de predicción del riesgo de violencia de género (VioGén). También se están utilizando algoritmos en la detección de fraudes y casos de corrupción (Agencia Tributaria, Inspección de Trabajo) y en procesos de asignación de ayudas sociales (bono social eléctrico).

3.- La opacidad de los algoritmos utilizados en las decisiones automatizadas: la transparencia administrativa

Se entiende por actuación administrativa automatizada 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. En estos supuestos, deberá establecerse previamente el órgano u órganos competentes, según los casos, para la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente. Asimismo, se indicará el órgano que debe ser considerado responsable a efectos de impugnación⁶. Como aclara Martín Delgado (2009:360), el hecho de que la decisión haya sido adoptada por una máquina de forma automatizada no implica que la actuación se le impute a ella, sino que la autoría del acto recaerá sobre el órgano administrativo que ejerce la potestad y tiene encomendada la competencia.

En términos generales, hay algoritmos que simplemente sirven para facilitar la actuación administrativa y vuelcan el contenido de las normas en un programa para que sea más fácil aplicarlas. Se utilizan en decisiones regladas. En estos casos, el programa informático es instrumental y no aporta nada al contenido de la decisión administrativa. Por ejemplo, el otorgamiento de certificados o la jubilación de un funcionario por razón de su edad. Por el contrario, es muy frecuente que el titular del órgano adopte la decisión sobre la base -entre otros elementos- de predicciones algorítmicas. La inteligencia artificial, entendida como predicciones algorítmicas, no necesariamente supone actuación automatizada (Huergo Lora.2021:87). Los sistemas de inteligencia artificial, cuando se aplican por la Administración, se mueven en los campos menos controlados jurídicamente, como ocurre, por ejemplo, con las decisiones por las que se selecciona a los sujetos que van a ser inspeccionados. Se trata de actos de trámite que normalmente no son recurribles, que no es necesario motivar y que muchas veces se toman por razones subjetivas. Otra cosa muy distinta será utilizar estos sistemas para la toma de auténticas resoluciones (Huergo Lora 2021:86). Para Martín Delgado (2009:369-370), la

⁶ Artículo 41 de la L40/2015, de 1 de octubre de régimen jurídico del Sector Público desarrollado por el artículo 13 del Real Decreto 203/2021, de 30 de marzo en el ámbito de la Administración general del Estado y el artículo 44 de la Ley 26/2010, de 3 de agosto de régimen jurídico y de procedimiento de las Administraciones Públicas de Cataluña.



aplicación de la actuación administrativa automatizada exclusivamente al campo de las potestades regladas no ha lugar a la duda, aunque también cabría la automatización de la decisión en el ejercicio de una potestad administrativa discrecional cuando la libertad decisoria se concreta en criterios técnicos y no políticos. Para Ponce Solé (2019:29-30), cuando se ejerce discrecionalidad administrativa y se toman en consideración hechos, intereses y derechos de personas, la empatía entendida como la capacidad para ponerse en el lugar del otro debería estar presente y parece que sólo podría ejercerla un ser humano. Estaríamos ante el sabio ejercicio del poder, basado en una ponderación cuidadosa de las circunstancias pertinentes en cada toma de la decisión⁷.

En primer término, acudiendo al principio de transparencia administrativa⁸ y al derecho a una buena administración del artículo 41 de la Carta de los Derechos Fundamentales de la Unión Europea (CDFUE), habría que comenzar dando a conocer qué aplicaciones de inteligencia artificial utilizan las Administraciones Públicas. Uno de los problemas más acuciantes en el uso de los sistemas algorítmicos es el de su opacidad. La Administración utiliza instrumentos técnicos cuya razón última resulta desconocida para el funcionario medio, quien además ha de desarrollar nuevas destrezas, así como para el ciudadano. Ello se acrecienta con el uso de los algoritmos opacos o de caja negra (De la Sierra Morón.2020:7). Los algoritmos son considerados como cajas negras por tres razones: la opacidad, fruto de su complejidad técnica, al emplearse lenguaje de programación; la confidencialidad o el secreto sobre el contenido de los mismos; y la ausencia de acceso a su código fuente⁹ (Baz Lomba.2021:128). Habría que profundizar en dar publicidad activa a los algoritmos utilizados por los organismos públicos¹⁰. Por ello,

⁷En esta línea la ley catalana 26/2010, de 3 de agosto de régimen jurídico y de procedimiento de las administraciones públicas de Cataluña que en su artículo 4.4.2 señala: "Sólo son susceptibles de actuación administrativa automatizada los actos que puedan adoptarse con una programación basada en criterios y parámetros objetivos".

⁸Si bien la transparencia no aparece recogida en la Constitución Española (CE), si se recoge el derecho fundamental de todos los ciudadanos a participar en los asuntos públicos (artículo 23 CE) y el deber de los poderes públicos de fomentar la participación de todos los ciudadanos en la vida política (artículo 9.2 CE). Este principio supone dar a conocer a los ciudadanos asuntos relacionados con la actividad de los poderes públicos para fomentar el control de la gestión de la vida pública con apoyo constitucional en el principio democrático.

⁹ El código fuente de un algoritmo se define por la Comisión de Garantía del Derecho de Acceso a la Información Pública de Cataluña (GAIP) en su Resolución 123-124/2016 acumuladas, de 21 de septiembre de 2016, como el conjunto de instrucciones escritas-en lenguaje de programación que corresponda- del programa informático empleado para implementar el algoritmo.

¹⁰ La fundación Ethicas ha desarrollado un observatorio de algoritmos de impacto social (OASI). Este contiene un Registro de algoritmos que están siendo desarrollados e implementados por diferentes instituciones y organizaciones en todo el mundo. El Registro OASI recopila, clasifica y hace que los algoritmos se puedan buscar, y se actualiza periódicamente con nuevos contenidos. Puede accederse a dicho registro a través del enlace: <https://eticasfoundation.org/es/>



sería deseable para los algoritmos empleados por las Administraciones Públicas la puesta en marcha de un registro público que los identifique en base al principio de transparencia¹¹.

En segundo lugar y como consecuencia de lo anterior, resulta evidente que, si se pretenden usar estos sistemas como herramientas de apoyo a la toma de decisiones, ya sea en procedimientos reglados o en aquellos otros de naturaleza discrecional¹² (Martínez Gutiérrez.2021:11), o como sistemas que acaben sustituyendo toda intervención humana, la Administración debería estar obligada a facilitar al ciudadano una explicación sobre el funcionamiento del algoritmo en un lenguaje comprensible. Y ello, particularmente, cuando el acceso al código fuente del algoritmo no sea posible, o sea tan complejo que lo haga inaccesible para un ciudadano medio (Velasco Rico.2022:235).

Hasta ahora, las normas sobre protección de datos han sido la primera respuesta normativa a este fenómeno de las decisiones automatizadas cuando se empleen datos personales, si bien los datos de entrenamiento para las predicciones algorítmicas suelen estar anonimizados. En términos generales, a todos los tratamientos automatizados de datos personales les resulta aplicable íntegramente el régimen de protección de datos, el Reglamento General de Protección de Datos (RGPD) y la Ley Orgánica 3/2018 de Protección de datos y garantía de derechos digitales (LOPDGDD), que establece la necesidad de que los sistemas se configuren de conformidad a la privacidad desde el diseño (Martínez Gutiérrez.2021:14). Las exigencias generales del principio de la transparencia contemplado en el artículo 5.1 RGPD son los derechos de información de los artículos 13 y 14 RGPD y el artículo 11 de la LOPDGDD y el derecho de acceso del artículo 15 RGPD y artículo 13 LOPDGDD, que forman parte del contenido esencial del derecho fundamental a la protección de datos personales del artículo 18.4 CE. Este último derecho se encuentra consagrado en el artículo 8.2 CDFUE, cuando dispone que toda persona tiene derecho a acceder a los datos recogidos que le conciernan.

En cuanto a las decisiones individuales automatizadas, se prohíben salvo que se cumplan determinados requisitos. En primer término, el artículo 22 RGPD contiene como regla general una prohibición de las decisiones puramente automatizadas, es decir, tomadas sin intervención humana, y ello tanto si se basan en una elaboración

¹¹ De acuerdo con el artículo 157.2 de la Ley 40/2015, de 1 de octubre de régimen jurídico del Sector Público, las Administraciones podrán declarar como fuentes abiertas las aplicaciones, desarrolladas por sus servicios o que hayan sido objeto de contratación y de cuyos derechos de propiedad intelectual sean titulares, cuando de ello se derive una mayor transparencia en el funcionamiento de la Administración Pública.

¹² Sentencia del Tribunal Supremo de 1 de junio de 1987: "la potestad que (...) tiene la Administración de elegir entre varias alternativas legalmente indiferentes, ya que la decisión discrecional se basa en criterios extrajurídicos, de oportunidad o de conveniencia que la Ley no predetermina, sino que deja a su libre consideración y decisión, pudiendo en consecuencia optar según su subjetivo criterio".



de perfiles como en otras fuentes de datos. Con esta disposición se pretende dotar al interesado de un control sobre sus datos personales¹³. Ahora bien, esta prohibición se aplica únicamente cuando se dan unas condiciones muy estrictas: la falta de intervención humana y la producción de efectos jurídicos o la afectación significativa, de forma similar, a un interesado.

La falta de intervención humana determina la aplicación del art. 22 RGPD, con sus garantías reforzadas. Por ello, no sería extraño que algunos sistemas simulen intervenciones de expertos para eludir este precepto. Siguiendo al Grupo de Trabajo del artículo 29, la intervención humana debe ser cualitativamente significativa -meaningful-, autorizada y competente para modificar la decisión. Para Palma Ortigosa (2019:9), si la decisión que toma el algoritmo es la que finalmente acaba adoptándose por el humano sin realizar un análisis de tal decisión o simplemente existe una mera supervisión, hay que entender que tal decisión es plenamente automatizada y, por tanto, se incluiría en el ámbito de aplicación del artículo 22 RGPD. Esta será una tarea que corresponderá acreditar en la evaluación de impacto de protección de datos (art.35 RGPD y art.28 LOPDGDD).

Por otra parte, resulta difícil precisar qué se considera efectos jurídicos o afectación suficientemente significativa para un interesado. Podrían ajustarse a esta categoría las siguientes decisiones: "las que afecten a las circunstancias financieras de una persona, como su elegibilidad para un crédito; aquellas que afecten al acceso de una persona a los servicios sanitarios; o que denieguen a una persona una oportunidad laboral o que la coloquen en gran desventaja; o que afecten al acceso de una persona a la educación, por ejemplo, su ingreso en la universidad".¹⁴

Hay que tener en cuenta que la prohibición del párrafo primero del artículo 22 RGPD no es absoluta y admite excepciones cuando se cumplan las condiciones contempladas en el apartado segundo. El precepto viene a indicar cuáles son las únicas bases jurídicas que justifican este tratamiento. En el caso de las Administraciones Públicas, la única que legitima este tratamiento es la habilitación normativa del artículo 22.2 b) RGPD, que permite al Derecho de los Estados miembros que autorice a los responsables a llevar a cabo aquel. Así, la ley de régimen jurídico del sector público, ley 40/2015 de 1 de octubre (LRJSP), antes mencionada, hace referencia a la actuación administrativa automatizada, previéndose la posibilidad de realizar auditorías a los códigos fuente en los que se

¹³ El Grupo de trabajo del artículo 29 al abordar si el concepto derecho del artículo 22.1 RGPD implica la necesidad de invocación activa del interesado de acción por su parte o establece una prohibición general se decanta por esta última interpretación pues refuerza la idea de que sea el interesado quien tenga el control sobre sus datos personales, lo cual corresponde con los principios fundamentales del RGPD. Grupo de Trabajo para la protección de datos del artículo 29 (2017) *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*. p. 22. Puede accederse a dicho informe a través de : <https://www.aepd.es/es>

¹⁴ *Ibidem*, p.24.



basan los programas que dictan los actos administrativos, así como el órgano responsable a efectos de impugnación (art.41.2 LRJSP). Y en el mismo sentido, el artículo 96.3 de la Ley General Tributaria, Ley 58/2003, de 17 de diciembre (LGT), dispone que cuando la Administración tributaria actúe de forma automatizada se garantizará la identificación de los órganos competentes para la programación y supervisión del sistema de información y de los órganos competentes para resolver los recursos que puedan interponerse.

Conforme al artículo 22.3 RGPD, los responsables del tratamiento deben proporcionar las garantías apropiadas, incluido "el derecho a obtener la intervención humana... a expresar su punto de vista y a impugnar la decisión". Estas facultades solo se prevén cuando el tratamiento sea necesario para la celebración o ejecución de un contrato o cuando esté basado en el consentimiento del particular (art. 22.3 RGPD). Por tanto, en relación con las Administraciones Públicas, los particulares no gozarán en principio de tales derechos, salvo que las normas nacionales los prevean. En este supuesto, compartimos la opinión de Palma Ortigosa (2019:32) de que, si las normas que desarrollen el artículo 22.2 b) RGPD no contemplan unas garantías adecuadas, estas deberían ser como mínimo las mismas que se prevén en el número 3 del mismo artículo, es decir, el derecho a obtener la intervención humana por parte del responsable, el derecho del interesado a expresar su punto de vista y el derecho a impugnar la decisión automatizada, entendida como la posibilidad de revertir la decisión en un contexto prejudicial. En definitiva, un conjunto de facultades que suponen para el particular todo un sistema de garantías sobre las decisiones tomadas por una máquina que le afecten.

A los anteriores cabe añadir el derecho a obtener una explicación sobre la decisión automatizada del Considerando 71 RGPD. Estas facultades se materializan en los artículos 13.2 f) y 14.2 g) RGPD, que obligan al responsable del tratamiento a informar al interesado sobre la existencia de decisiones automatizadas, incluida la elaboración de perfiles a la que se refiere el artículo 22, apartados 1 y 4 RGPD, y al menos en tales casos, a otorgarle *información significativa sobre la lógica aplicada, así como la importancia y las consecuencias previstas de dicho tratamiento para el interesado* (la cursiva es nuestra). Y en términos idénticos se pronuncia el RGPD, sobre el derecho del interesado a acceder a sus datos personales y a determinada información del artículo 15.1 h) RGPD.

La doctrina ha puesto énfasis en este último precepto, como referencia general para una nueva garantía: un derecho a la explicación de las decisiones automatizadas (Roig, 2020:28). No obstante, el Grupo de Trabajo del artículo 29 ha sostenido que el RGPD exige que el responsable del tratamiento ofrezca información significativa sobre la lógica aplicada, no necesariamente una completa explicación de los algoritmos utilizados o la revelación de todo el algoritmo¹⁵. Sin olvidar que el derecho

¹⁵ Ibidem, p.28.



a la protección de datos no es un derecho absoluto, debiendo ponderarse con otros bienes jurídicos en juego, como los derechos de propiedad intelectual o los secretos comerciales a los que puede estar sometido el algoritmo. En opinión de Medina Guerrero (2022:165), de los preceptos involucrados en el tema (entre ellos el considerando 71 y los problemas de su carácter vinculante o no) resulta muy difícil sostener que del RGPD se derive inmediatamente el derecho a la explicación algorítmica como norma armonizada para toda la Unión. Dicho esto, en el ámbito de las relaciones Administración/administrados no deja de percibirse un deber de explicación o motivación de las decisiones concretas, constitucionalmente exigible no ya como una exigencia derivada del derecho fundamental a la protección de datos personales, sino como una obligación impuesta por manifestaciones esenciales del Estado de Derecho, como las contenidas en los artículos 9.3 CE en relación con el artículo 24 CE y el artículo 106 CE. En este sentido, se pronuncia el artículo 35 de la ley 39/2015 de procedimiento administrativo común de las Administraciones Públicas (LPAC).

Más adecuado nos parece sostener que el fundamento constitucional del derecho a la explicación de las decisiones automatizadas se encuentra en el derecho de acceso a la información pública recogido en el artículo 105 b) CE, como opción para conocer los algoritmos que emplean las Administraciones Públicas en la toma de decisiones. Este derecho está desarrollado en el artículo 13 d) LPAC y, fundamentalmente, en los artículos 12 y siguientes de la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno (LTBG), cuyo artículo 12 dispone específicamente que:

"Todas las personas tienen derecho a acceder a la información pública, en los términos previstos en el artículo 105.b) de la Constitución Española, desarrollados por esta Ley".

Partiendo de un concepto amplio de información pública, que el artículo 13 define como "los contenidos o documentos, *cualquiera que sea su formato o soporte*, que obren en poder de alguno de los sujetos incluidos en el ámbito de aplicación de este título y que hayan sido elaborados o adquiridos en el ejercicio de sus funciones" (la cursiva es nuestra), podría interpretarse que los algoritmos son información pública. En base a ese derecho de acceso a la información pública, las autoridades de Transparencia han reconocido que dicho concepto incluye también los algoritmos que haya utilizado la Administración¹⁶.

Esta necesidad de explicación de las decisiones automatizadas acabará imponiéndose en la práctica por el desarrollo de la normativa sobre transparencia. En este sentido, ha sido pionera la nueva ley 1/2022, de 13 de abril de Transparencia

¹⁶ Resolución 0093/2019 de 22 de febrero de la Comisión de Garantía del derecho de acceso a la información pública de Cataluña (GAIP) y Resoluciones RT 0748/2021 y RT 0253/2021 de 19 de noviembre del Consejo de Transparencia y Buen Gobierno.



y Buen Gobierno de la Comunidad Valenciana, que en su artículo 161.1 l) impone a las Administraciones Públicas, la obligación de publicar la relación de sistemas algorítmicos o de inteligencia artificial que tengan impacto en los procedimientos administrativos o la prestación de los servicios públicos, con la descripción de manera comprensible de su diseño y funcionamiento, el nivel de riesgo que implican y el punto de contacto al que poder dirigirse en cada caso, de acuerdo con los principios de transparencia y explicabilidad.

4.- El control de los algoritmos empleados por las Administraciones Públicas

4.1.- La respuesta europea

La ausencia de un marco jurídico específico de la IA no es obstáculo, como es obvio, para la aplicación en este terreno de las garantías constitucionales para la protección de los derechos fundamentales. La Unión Europea (UE) ha sido pionera en explorar respuestas legales a los desafíos que plantea el uso de la IA. Desde 2017, ha estado trabajando en el establecimiento de una política común sobre la IA y analizando los riesgos de estas nuevas tecnologías¹⁷. Entre los documentos más relevantes podemos citar la Comunicación COM (2019)168 de 8 de abril, Generar confianza en la inteligencia artificial centrada en el ser humano (Gamero Casado. 2021:271). Asimismo, el 19 de febrero de 2020, la Comisión Europea presentó el "Libro Blanco sobre inteligencia artificial- un enfoque europeo orientado a la excelencia y a la confianza"¹⁸, en el que considera conveniente mejorar el marco normativo europeo. A partir de sus postulados se desarrollarán las líneas de actuación de la UE en esta materia. Entre sus propuestas, se incluye la creación de una red de autoridades de control algorítmico para supervisar el uso de estos sistemas, tanto en el sector privado como en el público. Como señala Soriano Arnanz (2021:118), estas instituciones deberían estar integradas por personal procedente de diferentes disciplinas en la protección de los derechos fundamentales y en el campo de la ética algorítmica, ya que se echa en falta que a la hora de su diseño no se tengan en cuenta la protección de la igualdad y otros valores y principios democráticos.

Paralelamente a la Comisión, también el Parlamento Europeo ha impulsado varias iniciativas, entre la que cabe destacar la Resolución de 20 de enero de 2021 sobre

¹⁷ El 25 de abril de 2018 la Comisión Europea expuso una iniciativa sobre inteligencia artificial en su Comunicación Inteligencia Artificial para Europa en la que propuso un enfoque que coloca a las personas en el centro del desarrollo de la inteligencia artificial y entre cuyos fines figuraba el de garantizar el establecimiento de un marco ético y jurídico apropiado basado en los valores de la Unión y en consonancia con la CDFUE. Vid. COM(2018)237 final ; El 4 de diciembre de 2018 la Comisión propuso elaborar un Plan Coordinado con los Estados miembros entre el que se abordarían las cuestiones de carácter ético y jurídico Vid. Comunicación sobre Plan coordinado sobre inteligencia artificial COM/2018/795 final.

¹⁸ COMISIÓN EUROPEA, "Libro Blanco sobre inteligencia artificial- un enfoque europeo orientado a la excelencia y a la confianza", COM(2020) 65 final. Disponible en: <https://eur-lex.europa.eu>



inteligencia artificial: cuestiones de interpretación y de aplicación del Derecho internacional¹⁹.

Algunas de las líneas de acción han pasado a convertirse en mandatos normativos mediante el Reglamento (UE) 2021/694 del Parlamento Europeo y del Consejo, de 29 de abril de 2021, por el que se establece el Programa Europa Digital y por el que se deroga la Decisión (UE) 2015/2240. Esta norma regula en su artículo 5 los objetivos específicos de la Unión en materia de IA: "Las soluciones basadas en la IA y los datos que se faciliten deberán respetar el principio de privacidad y seguridad desde el diseño y deberán respetar plenamente la normativa sobre protección de datos"²⁰.

Asimismo, en abril de 2021 se publicó la Propuesta de Reglamento del Parlamento y del Consejo por el que se establecen normas armonizadas sobre la Inteligencia Artificial de la Unión Europea, también conocida como Ley de IA, Artificial Intelligence Act (AIA). Se trata de una regulación novedosa en esta materia, a la que le queda un largo camino por recorrer hasta que finalmente entre en vigor. El objetivo que persigue es que los sistemas de IA que se utilizan en Europa sean seguros y respeten los derechos fundamentales. Ciertamente, supone un avance muy positivo y necesario para generar confianza en esta tecnología y representa la adopción por la UE de una posición de liderazgo a nivel mundial, como ya lo hizo en materia de protección de datos. Si bien hay que señalar que la regulación está pensada para proveedores, desarrolladores y usuarios de la IA, echándose en falta alguna mención expresa a los afectados por los sistemas de IA, a sus derechos y sus posibles mecanismos de garantía. La Administración Pública será muy posiblemente usuaria de servicios de inteligencia artificial cuyo uso adquiere.

En cuanto a su ámbito de aplicación, la definición de sistema de inteligencia artificial (sistema de IA) que maneja la Propuesta de Reglamento se encuentra en su artículo 3: el software que se desarrolla empleando una o varias de las técnicas y estrategias que figuran en el anexo I²¹ y que puede, para un conjunto determinado de objetivos definidos por seres humanos, generar información de salida como contenidos, predicciones, recomendaciones o decisiones que influyan en los entornos con los

¹⁹ PARLAMENTO EUROPEO, Resolución de 20 de enero de 2021, sobre inteligencia artificial: cuestiones de interpretación y de aplicación del Derecho internacional en la medida en que la UE se ve afectada en los ámbitos de los usos civil y militar, así como de la autoridad del Estado fuera del ámbito de la justicia penal (2020/2013(INI)), DOUE, C456, 10/11/202, pp.34-46.

²⁰ Reglamento (UE) 2021/694 del Parlamento Europeo y del Consejo de 29 de abril de 2021 por el que se establece el Programa Europa Digital y por el que se deroga la Decisión (UE) 2015/2240. Disponible en: <https://eur-lex.europa.eu>

²¹ ANEXO I TÉCNICAS Y ESTRATEGIAS DE INTELIGENCIA ARTIFICIAL mencionados en el artículo 3, punto 1: a) Estrategias de aprendizaje automático, incluidos el aprendizaje supervisado, el no supervisado y el realizado por refuerzo, que emplean una amplia variedad de métodos, entre ellos el aprendizaje profundo. b) Estrategias basadas en la lógica y el conocimiento, especialmente la representación del conocimiento, la programación (lógica) inductiva, las bases de conocimiento, los motores de inferencia y deducción, los sistemas expertos y de razonamiento (simbólico). c) Estrategias estadísticas, estimación bayesiana, métodos de búsqueda y optimización.



que interactúa. La definición de «sistema de IA» que figura en el marco jurídico pretende ser tecnológicamente neutra y resistir al paso del tiempo lo mejor posible, habida cuenta de la rápida evolución de la tecnología y del mercado en relación con la IA.

La norma establece un sistema de control basado en niveles de riesgo. Según el artículo 6 de esta propuesta y el Anexo III de la misma, se consideran sistemas de IA de alto riesgo y, por tanto, sometidos a unos requisitos específicos para poder ser utilizados en el ámbito de la Unión Europea, los utilizados en: a) la identificación biométrica y categorización de personas físicas; b) la gestión y funcionamiento de infraestructuras esenciales (como la gestión y funcionamiento del tráfico rodado y el suministro de agua, gas, calefacción y electricidad, etc.); c) la educación y formación profesional; d) el empleo, gestión de los trabajadores y acceso al autoempleo; e) el acceso y disfrute de servicios públicos y privados esenciales y sus beneficios (como los sistemas de IA destinados a ser utilizados por las autoridades públicas, o en su nombre, para evaluar la admisibilidad de las personas físicas para acceder a prestaciones y servicios de asistencia pública, así como para conceder, reducir, retirar o recuperar dichas prestaciones y servicios); f) los utilizados por las fuerzas policiales o judiciales (como polígrafos y herramientas similares, los sistemas para la evaluación de la fiabilidad de las pruebas durante la investigación o el enjuiciamiento de infracciones penales o los utilizados para predecir la frecuencia o reiteración de una infracción penal real o potencial); g) la gestión de la migración, el asilo y el control fronterizo. Como puede apreciarse, muchas de estas tareas se realizan por las Administraciones Públicas, por lo que su utilización deberá estar sujeta a los requisitos establecidos por el Reglamento.

La calidad de los datos utilizados para la programación es uno de los grandes desafíos de los sistemas de IA. La inadecuación de los bancos de datos puede producir sesgos o disfunciones del sistema. El artículo 10 de la propuesta establece garantías y requisitos específicos en relación con esta cuestión (Gamero Casado.2021:282).

No obstante, la normativa resulta un tanto abstracta e imprecisa y no aborda los problemas asociados al uso de estas nuevas tecnologías, derivados, entre otros, de la opacidad, la complejidad y la capacidad para un funcionamiento autónomo de los algoritmos. Así, no se refiere en ninguna ocasión a los derechos individuales de las personas afectadas por el funcionamiento de estos sistemas de IA. En concreto, la propuesta de "Ley de Inteligencia Artificial" europea trata de establecer una serie de requisitos relacionados con la transparencia, explicabilidad y comprensibilidad de los sistemas, con el objetivo de someterlos a fiscalización. El artículo 11 de la propuesta establece la obligación de preparar y mantener actualizada la documentación técnica relativa al sistema. El considerando 46 aclara que "para verificar si los sistemas de IA de alto riesgo cumplen los requisitos previstos en el presente Reglamento, resulta esencial disponer de información sobre el modo en



que se han desarrollado y sobre su funcionamiento durante todo su ciclo de vida. A tal fin, es preciso llevar registros y disponer de documentación técnica que contenga la información necesaria para evaluar si el sistema de IA en cuestión cumple los requisitos pertinentes. Dicha información debe incluir, en particular, *los algoritmos*" (la cursiva es nuestra). Y en cuanto a la obligación de transparencia, el considerando número 47 la explica con mayor claridad: "cierto grado de transparencia respecto de los sistemas de IA de alto riesgo para subsanar la opacidad que puede hacer a algunos de ellos incomprensibles o demasiado complejos para las personas físicas. Los usuarios deben ser capaces de interpretar la información de salida del sistema y de usarla adecuadamente. En consecuencia, los sistemas de IA de alto riesgo deben ir acompañados de la documentación y las instrucciones de uso oportunas e incluir información clara y concisa, en particular sobre los posibles riesgos para los derechos fundamentales y de discriminación, cuando corresponda". En este sentido, los artículos 13 y 52. No obstante, no se reconoce un derecho a la información como tal por parte de las personas físicas afectadas por el funcionamiento del sistema. No hay obligaciones de transparencia como garantía de cara al afectado.

Dada la capacidad de estos sistemas de evolucionar a medida que son utilizados, se establece la obligación de registrar los eventos del sistema (art.12), la cual tiene relevancia, puesto que estos funcionan de manera autónoma, modificando constantemente los parámetros de análisis, lo que dificulta la prueba en los casos de discriminación algorítmica o cuando se trate de acreditar la ilegalidad de la decisión automatizada.

Y, finalmente, se recoge la posibilidad de vigilancia humana, en el sentido de que los sistemas de IA se diseñarán y desarrollarán de modo que puedan ser supervisados de manera efectiva por personas físicas durante el periodo que estén en uso, para prevenir o reducir al mínimo los riesgos para los derechos fundamentales. Para Gamero Casado (2021:282), las medidas en este sentido son relevantes, pero insuficientes, pues considera que para ajustarse a esta finalidad, la supervisión humana debe llevarse a cabo por especialistas, de ahí que debería confiarse expresamente a equipos multidisciplinares, en el que tengan cabida tanto técnicos como juristas.

Entre los aspectos destacables, la propuesta prohíbe expresamente los sistemas de puntuación social (social score)²² y de vigilancia biométrica masiva, salvo

²² En el marco de la UE la propuesta de Reglamento contempla expresamente en el considerando 17 los llamados sistemas de puntuación social o social score, "sistemas de IA que proporcionan calificaciones sociales de personas físicas para su uso con fines generales por parte de las autoridades públicas o en representación de estas pueden tener resultados discriminatorios y abocar a la exclusión a determinados grupos. Pueden menoscabar el derecho a la dignidad y la no discriminación y los valores de igualdad y justicia. Dichos sistemas de IA evalúan o clasifican la fiabilidad de las personas físicas en función de su comportamiento social en múltiples contextos. La calificación social resultante de dichos sistemas de IA puede dar lugar a un trato perjudicial o desfavorable de personas físicas o colectivos enteros en contextos sociales que no guardan relación con el contexto donde se generaron o recabaron los datos originalmente, o a un trato perjudicial desproporcionado o injustificado en



excepciones, impidiendo en el primer caso a las autoridades públicas emplear la IA para evaluar los distintos aspectos de la vida de las personas, para, a partir de ello, determinar el acceso a servicios, ayudas y contratos o imponerles sanciones o penas.

4.2.- El control judicial de las decisiones automatizadas en el sector público

Hasta el momento, el control efectivo sobre el uso de los algoritmos ha recaído en los órganos jurisdiccionales. Principios básicos como el de transparencia resultan de aplicación en este nuevo ámbito, como se ha puesto de manifiesto en algunas resoluciones judiciales en el Derecho Comparado. Frente a la postura en el ámbito anglosajón que defiende la inexistencia de la obligación de informar cómo funciona el algoritmo, (sentencia del Tribunal Supremo de Wisconsin en el caso *State v. Loomis*, de 10 de marzo de 2017), en Europa se impone la necesidad de particulares garantías, en supuestos donde los algoritmos sirven de apoyo a la toma de decisiones por parte de la Administración o son fuente de decisiones automatizadas. Hemos seleccionado algunos ejemplos.

4.2.1.- Italia

El Consejo de Estado italiano se ha pronunciado en varias ocasiones sobre el uso de los algoritmos por las Administraciones Públicas. Traemos a colación una de ellas, la Sentencia núm. 8472 de 13 de diciembre de 2019²³. Se trataba de un procedimiento de movilidad nacional de personal docente, cuyo resultado, realizado sobre la base de un algoritmo desconocido y que no funcionó correctamente, dispuso los traslados sin tener en cuenta las preferencias expresadas por los solicitantes, a pesar de la existencia de plazas disponibles en las provincias indicadas.

En el recurso de apelación, el Consejo de Estado considera, que "la utilización de procedimientos informáticos no puede ser motivo de elusión de los principios que conforman nuestro ordenamiento jurídico y que rigen la conducta de las actividades administrativas. En este contexto, en efecto, el uso del algoritmo debe enmarcarse correctamente en términos de módulo organizativo, de instrumento procesal y preliminar, sujeto a los controles propios de todo procedimiento administrativo, que sigue siendo el *modus operandi* de la elección autoritaria, a realizarse sobre la base de la legislación que confiere el poder y los fines atribuidos por la misma al organismo público titular del poder". Y añade, a la luz de la normativa supranacional, los tres principios que constituyen los elementos de garantía mínima para cada supuesto de uso de algoritmos en la toma de decisiones públicas: el principio de cognoscibilidad, que supone el derecho a conocer la existencia de decisiones que nos conciernen adoptadas por algoritmos y, correlativamente, como deber de

relación con la gravedad de su comportamiento social. Por lo tanto, dichos sistemas de IA deben prohibirse".

²³ Sentencia del Consejo de Estado núm. 8472, de 13 de diciembre de 2019. Disponible en: <https://www.giustizia-amministrativa.it>



quienes tratan los datos de forma automatizada, de ponerlo en conocimiento del interesado, acompañado de mecanismos capaces de descifrar la lógica utilizada por el algoritmo; el principio de no exclusividad de la decisión algorítmica, entendido como la necesaria existencia en el proceso de toma de decisiones de una contribución humana capaz de controlar, validar o denegar la decisión automática; y el principio fundamental de no discriminación algorítmica, según el cual procede que el responsable del tratamiento utilice procedimientos matemáticos o estadísticos apropiados para la elaboración de perfiles, poniendo en marcha medidas técnicas y organizativas adecuadas para asegurar, en particular, que se corrijan los factores que dan lugar a inexactitudes en los datos y se minimice el riesgo de error, y para garantizar la seguridad de los datos personales, de forma que se tengan en cuenta los riesgos potenciales existentes para la intereses y derechos de los interesados y se impidan, entre otros, los efectos discriminatorios contra las personas por motivos de raza u origen étnico u opiniones políticas. En estos casos, los datos de "entrada" deben corregirse para evitar efectos discriminatorios en la toma de decisiones; esta operación obviamente requiere la necesaria cooperación de quienes entrenan a las máquinas que producen tales decisiones. En su resolución, el Consejo de Estado concluye que la imposibilidad de comprender las formas en que a través del mencionado algoritmo se han asignado las plazas disponibles, constituye en sí mismo un defecto tal que invalida el procedimiento, por la falta de transparencia del mismo.

4.2.2.- Países Bajos

Sin embargo, la sentencia más conocida y comentada (Cotino Hueso. 2020) en esta materia ha sido la del Tribunal del Distrito de la Haya de 5 de febrero de 2020. Su peculiaridad reside en que impone un estándar muy elevado de garantías.

Esta sentencia fue dictada en relación con el programa "SyRI" utilizado por la Inspección de Trabajo y Seguridad Social holandesa para detectar y prevenir casos de fraude cometidos en la percepción de prestaciones a la seguridad social. El programa generaba informes de riesgo, es decir, señalaba si una persona física o jurídica debía ser investigada con respecto a un posible fraude, uso ilegal e incumplimiento de la legislación. El objeto del proceso judicial era la conformidad o no de la ley que regulaba este programa con el artículo 8 CEDH y no con la CDFUE o con el RGPD, dada la peculiaridad de la Constitución de los Países Bajos, donde no existe control jurisdiccional de constitucionalidad de las leyes (artículo 120) y este se ha articulado en relación con el CEDH, dada su apertura a los Tratados internacionales que son expresamente vinculantes (artículo 94).

El Tribunal partió de una concepción amplia del derecho al respeto a la vida privada para englobar también la protección de datos personales y aplicó un canon más estricto de control al legislador de SyRI, por un deber de responsabilidad especial cuando se aplican las nuevas tecnologías. En base a ello, aunque reconoció que la norma contemplaba con suficiente precisión y claridad las restricciones del derecho



y su finalidad era legítima, el órgano judicial declaró la violación del artículo 8 CEDH porque la legislación no respetó el principio de proporcionalidad. No había garantías suficientes debido a la gran cantidad de datos recopilados, de varios tipos y de fuentes diferentes, y no se conocían los indicadores ni el modelo de riesgo ni los criterios objetivos que subyacían a la validez de estos. Lo interesante del caso es que el sistema incluía garantías de anonimización, de división funcional, de borrado y de confidencialidad. Todo ello se valoró positivamente en la sentencia. No obstante, estas garantías no eran suficientes, porque la forma en que se llevó a cabo la selección definitiva del riesgo no era pública. La razón de la falta de información a los ciudadanos era que los defraudadores pudieran ajustar su conducta para eludir ser considerados perfiles de riesgo a inspeccionar, pero para el tribunal esto no era suficiente justificación. El órgano judicial echó en falta una revisión exhaustiva de antemano, o una revisión por un tercero independiente, con el fin de evaluar si era necesaria o no la restricción de derechos. La evaluación de impacto no se había llevado a cabo por excepción legal (artículo 30.10 RGPD). Sin embargo, la sentencia considera que hubiera servido para compensar la falta de garantías suficientes. Además del principio de transparencia (caja negra, opacidad, falta de información a los interesados), la sentencia tiene en cuenta los derechos de defensa y de no discriminación, aunque todo bajo la cobertura del art.8 CEDH. Los demandantes consideraban que tenía un efecto estigmatizador al investigar los barrios más marginales.

4.2.3.- España

En el caso de España, no hay apenas jurisprudencia sobre la materia. Conviene citar el caso del algoritmo (BOSCO) empleado en la concesión del bono social eléctrico. La solicitud de acceso al código fuente de este algoritmo fue rechazada por el Ministerio para la Transición Ecológica. Presentada reclamación ante el Consejo Estatal de Transparencia y Buen Gobierno, este la denegó amparándose en la protección de la propiedad intelectual del sistema²⁴, límite al derecho de acceso a la información pública. La sentencia 143/2021, de 30 de diciembre, del Juzgado Central Contencioso Administrativo número 8 de Madrid, desestimó igualmente la demanda de acceso al código fuente. Sobre el límite de la propiedad intelectual señala esta resolución que "hay que tener en cuenta que el código fuente de la mencionada aplicación informática no está dentro de las exclusiones de la propiedad intelectual, mencionadas en el artículo 13 del Real Decreto Legislativo 1/1996, de 12 de abril, de propiedad intelectual²⁵, precepto invocado por la entidad recurrente, pues dicho código no es una norma ni un acto administrativo".

²⁴ Sobre que el derecho de acceso a la información pública debe verse limitado por la protección de propiedad intelectual cuando se empleen algoritmos se alega el art.14.1 j) de la Ley 19/2013, de 9 de diciembre de transparencia, acceso a la información pública y buen gobierno.

²⁵ Artículo 13 del Real Decreto Legislativo 1/1996, de 12 de abril, de propiedad intelectual: "No son objeto de propiedad intelectual las disposiciones legales o reglamentarias y sus correspondientes proyectos,



5.- Los efectos discriminatorios de los algoritmos: el artículo 23 de la L15/2022

Si bien los algoritmos empleados por las Administraciones Públicas tienen el potencial de promover la eficacia, también pueden reforzar discriminaciones históricas que no siempre serán fáciles de apreciar, por ejemplo, contra las mujeres, ciertos grupos de edad, personas con discapacidad o personas de orígenes raciales o étnicos concretos o con una orientación sexual determinada.

Para que los supuestos de discriminación algorítmica puedan considerarse como de discriminación directa, será necesario que el sistema se entrene con datos que no deben tomarse en consideración por impedirlo el artículo 14 CE. Pero incluso si no se procesan esta clase de datos, también puede ocurrir que se tomen decisiones basándose en datos aparentemente inocuos. Serían los supuestos de discriminación directa por inferencia. Por ejemplo, si se utilizan datos no actualizados. En todos estos casos será especialmente difícil probar el nexo causal entre el tratamiento menos favorable y la pertenencia al grupo protegido si no se garantiza una plena transparencia de la instrucción del algoritmo. Es preciso que al menos sea accesible al tribunal²⁶. Los datos no representativos o los datos insuficientes pueden ser también ser causa de sesgos discriminatorios.

Al propio tiempo, cabe la posibilidad de situaciones de discriminación algorítmica indirecta, esto es, decisiones aparentemente neutras que dan lugar a resultados más perjudiciales para los miembros de un grupo especialmente protegido que para los ajenos al mismo. Para que una justificación se considere suficiente debe acreditarse que la medida aparentemente neutra persigue un objetivo legítimo y supera el test de proporcionalidad: adecuación, necesidad y proporcionalidad en sentido estricto. La tendencia de los seres humanos a confiar en la neutralidad de los sistemas automatizados incrementa el riesgo de que se tienda a aceptar los sesgos generados por estos sistemas como necesarios, concluyendo que superan el test de proporcionalidad de las medidas indirectamente discriminatorias. Este riesgo se encuentra especialmente presente en aquellos casos en los que se demuestre que el sistema tiene un elevado nivel de precisión. Dicha precisión en muchas ocasiones será el resultado de la incorporación y reflejo por parte del algoritmo de la estructura social discriminatoria y, sin embargo, por las características propias de los sistemas automatizados y la percepción de neutralidad que se tiene de estos sistemas, es probable que los tribunales terminen por aceptar decisiones automatizadas que perpetúan la situación de desventaja de algunos grupos (Soriano Aranz. 2022:162).

las resoluciones de los órganos jurisdiccionales y los actos, acuerdos, deliberaciones y dictámenes de los organismos públicos, así como las traducciones oficiales de todos los textos anteriores".

²⁶ AGENCIA DE LOS DERECHOS FUNDAMENTALES DE LA UNIÓN EUROPEA Y CONSEJO DE EUROPA, 2018. *Manual de legislación europea contra la discriminación*, pp. 43-71. Disponible en: <https://www.echr.coe.int/>



Asimismo, hay muchas posibilidades de que se produzcan errores o se introduzcan sesgos durante el desarrollo del sistema algorítmico y durante su funcionamiento. Por eso debe ponerse en tela de juicio la aparente mayor objetividad de los algoritmos, porque también las máquinas tienen fallos. Las conclusiones alcanzadas por estos sistemas se basan en correlaciones que se consideran suficientemente fiables, aunque no exista una prueba de causalidad. Estamos acostumbrados a que las decisiones se tomen en función de criterios razonados, explicables, lo que nos remite a juicios de causalidad o de tipo normativo. En cambio, los algoritmos predictivos suponen una nueva perspectiva. Se analiza una gran cantidad de datos referidos a experiencias anteriores de las que sabemos qué resultados produjeron y se extraen correlaciones, es decir, se identifica automáticamente qué características o grupos de características han llevado a los mejores resultados y se toman esos criterios como base para las decisiones (Huergo Lora. 2020:35). Con el riesgo de que se produzcan errores en los resultados finales.

Los algoritmos son diseñados por seres humanos que de manera voluntaria o involuntaria pueden verse influidos por los estereotipos sobre determinados grupos sociales. En la composición de los equipos de diseñadores de estos sistemas se ha encontrado el origen de los sesgos que incorporan esos sistemas algorítmicos (Soriano Arnanz.2021:92)²⁷. Al mismo tiempo, los algoritmos se desarrollan en una realidad social en la que existen situaciones discriminatorias, que pueden ayudar a reproducir y perpetuar.

Adicionalmente, puede suceder que los algoritmos no contengan sesgos perjudiciales contra personas pertenecientes a grupos desaventajados, pero que de las decisiones tomadas en relación a los mismos puedan derivarse situaciones discriminatorias. Esto es lo que ha ocurrido en España con el sistema público de concesión del bono social energético, que ha llevado a que se denieguen estas ayudas a personas que reunían los requisitos legales.

El artículo 14 de la Constitución de 1978 proclama el derecho a la igualdad y a la no discriminación, citando como motivos especialmente rechazables el nacimiento, la raza, el sexo, la religión u opinión, y prohibiendo la discriminación por cualquier otra circunstancia personal o social. En el ámbito europeo, los artículos 14 CEDH y 1 del Protocolo 12 del CEDH y el artículo 21 CDFFUE son los instrumentos que deben aplicarse en los casos de discriminación algorítmica. Estos prohíben la discriminación por razón de sexo, raza, color, idioma, religión, opiniones políticas u otras, origen nacional o social, pertenencia a una minoría nacional, fortuna, nacimiento o cualquier otra situación. La CDFFUE añade origen étnico, discapacidad, edad y orientación sexual.

²⁷ EL PERIÓDICO.COM, 2022, "Algoritmos sin perspectiva de género: solo el 0,5% de los técnicos en inteligencia artificial son mujeres", disponible en: <https://www.elperiodico.com>



La normativa de protección de datos sirve, sin lugar a dudas, como instrumento de protección de la igualdad y la prohibición de discriminación. Ello tiene su reflejo en los artículos 9 y 22 RGPD. En concreto, el párrafo cuarto del artículo 22, en conexión con el primero, prohíbe expresamente que las decisiones individuales automatizadas -autorizadas conforme al Derecho de la Unión o de los Estados miembros- se basen en las categorías especiales de datos personales contempladas en el artículo 9, apartado 1, es decir de datos personales que revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, y el tratamiento de datos genéticos, datos biométricos dirigidos a identificar de manera unívoca a una persona física, datos relativos a la salud o datos relativos a la vida sexual o la orientación sexual de una persona física. Exceptuando razones de interés público esencial y que se hayan tomado medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado²⁸.

Con el objeto de dar una respuesta normativa más determinada a estos supuestos y prevenir otros que se puedan presentar, se ha aprobado en España la Ley 15/2022, de 12 de julio, que, según su exposición de motivos, tiene como objetivo "crear un instrumento eficaz contra toda discriminación que pueda sufrir cualquier persona y que aborde todos los ámbitos desde los que esta se pueda producir, acogiendo la concepción más moderna de los derechos humanos". Se trata de un derecho antidiscriminatorio específico, que viene a dar cobertura a las discriminaciones que existen y a las que pueden surgir, ya que los desafíos de la igualdad cambian con la sociedad y, en consecuencia, también deberán hacerlo en el futuro las respuestas debidas. En este sentido, su ámbito objetivo de aplicación se extiende a la Inteligencia Artificial y gestión masiva de datos, así como otras esferas de análoga significación. Se trata de la primera norma que señala cómo deberán las Administraciones Públicas diseñar los algoritmos utilizados en la toma de decisiones. Así, dispone en su artículo 23:

"Inteligencia Artificial y mecanismos de toma de decisión automatizados.

1. En el marco de la Estrategia Nacional de Inteligencia Artificial, de la Carta de Derechos Digitales y de las iniciativas europeas en torno a la Inteligencia Artificial, las administraciones públicas favorecerán la puesta en marcha de mecanismos para que los algoritmos involucrados en la toma de decisiones que se utilicen en las administraciones públicas tengan en cuenta criterios de minimización de sesgos, transparencia y rendición de cuentas, siempre que sea factible técnicamente. En estos mecanismos se incluirán su diseño y datos de entrenamiento, y abordarán su potencial impacto discriminatorio. Para lograr este fin, se promoverá la realización de evaluaciones de impacto que determinen el posible sesgo discriminatorio.

²⁸ Artículo 9.2 g) RGPD.



2. Las administraciones públicas, en el marco de sus competencias en el ámbito de los algoritmos involucrados en procesos de toma de decisiones, priorizarán la transparencia en el diseño y la implementación y la capacidad de interpretación de las decisiones adoptadas por los mismos.
3. Las administraciones públicas y las empresas promoverán el uso de una Inteligencia Artificial ética, confiable y respetuosa con los derechos fundamentales, siguiendo especialmente las recomendaciones de la Unión Europea en este sentido.
4. Se promoverá un sello de calidad de los algoritmos".

De las tres propuestas que menciona este artículo, merece la pena destacar la Carta de Derechos Digitales, por su contenido mucho más completo que el del artículo 23. Como sabemos, este documento de julio de 2021 no tiene valor normativo, sino que su objetivo es, según sus consideraciones previas, reconocer los novísimos retos de aplicación e interpretación que la adaptación de los derechos al entorno digital plantea, así como sugerir principios y políticas referidas a ellos en el citado contexto. En concreto, en su apartado XXV, la Carta establece:

"Derechos ante la inteligencia artificial

1. La inteligencia artificial deberá asegurar un enfoque centrado en la persona y su inalienable dignidad, perseguirá el bien común y asegurará cumplir con el principio de no maleficencia.
2. En el desarrollo y ciclo de vida de los sistemas de inteligencia artificial:
 - a) Se deberá garantizar el derecho a la no discriminación cualquiera que fuera su origen, causa o naturaleza, en relación con las decisiones, uso de datos y procesos basados en inteligencia artificial.
 - b) Se establecerán condiciones de transparencia, auditabilidad, explicabilidad, trazabilidad, supervisión humana y gobernanza. En todo caso, la información facilitada deberá ser accesible y comprensible.
 - c) Deberán garantizarse la accesibilidad, usabilidad y fiabilidad.
3. Las personas tienen derecho a solicitar una supervisión e intervención humana y a impugnar las decisiones automatizadas tomadas por sistemas de inteligencia artificial que produzcan efectos en su esfera personal y patrimonial".

Frente a lo dispuesto en la Carta, el contenido del artículo 23 de la Ley 15/2022 no menciona ni la supervisión humana de las decisiones adoptadas por el sistema, ni el derecho a impugnar la decisión adoptada. Según este nuevo precepto, cuando las Administraciones Públicas utilicen algoritmos para la toma de decisiones, deberán favorecer que los mismos estén dotados de unos mecanismos sobre los que no se dan más detalles, y, *siempre que ello sea factible técnicamente*, que tengan en cuenta la minimización de sesgos, la transparencia y la rendición de cuentas. Lo cual deja



en el aire la puesta en marcha de estos mecanismos, que dependerá de que se dote a las Administraciones Públicas de medios técnicos adecuados.

Con respecto a la minimización de sesgos, la propuesta europea de ley de IA (AIA) expone que si el sistema de IA no está entrenado con datos de buena calidad, no cumple los requisitos oportunos en términos de precisión o solidez, o no se diseña y prueba debidamente antes de introducirlo en el mercado o ponerlo en servicio, puede señalar a personas de manera discriminatoria, incorrecta o injusta²⁹. Por eso regula en su artículo 10 los datos y la gobernanza de los datos de los sistemas de IA considerados de alto riesgo. Y se precisa que los conjuntos de datos utilizados para el entrenamiento, validación y prueba de estos sistemas deberán cumplir los criterios de calidad que se detallan en sus distintos apartados. En particular, según el apartado tercero "serán pertinentes y representativos, carecerán de errores y estarán completos".

En cuanto a la transparencia, resulta muy interesante el documento sobre el Estándar de Transparencia Algorítmica elaborado por la Oficina Central de lo Digital y el Dato británica³⁰.

En relación con la rendición de cuentas, el artículo 17 de la propuesta europea de ley de IA (AIA) señala que los sistemas de IA de alto riesgo establecerán un sistema de gestión de la calidad que garantice el cumplimiento del presente Reglamento, que incluirá "m) un marco de rendición de cuentas que defina las responsabilidades del personal directivo y de otra índole en relación con todos los aspectos enumerados en este apartado".

Por último, resaltar que se crea, conforme al art. 40 de la Ley integral de igualdad de trato, la Autoridad Independiente para la Igualdad de Trato y la No Discriminación, en el ámbito de la Administración del Estado, como encargada de proteger y promover la igualdad de trato y no discriminación de las personas por razón de las causas y en los ámbitos competencia del Estado previstos en esta ley, tanto en el sector público como en el privado.

6.- Conclusiones

No cabe duda el extraordinario potencial que la revolución digital 4.0 representa para las Administraciones Públicas. Es evidente la ventaja que entraña la IA para la calidad de los servicios prestados a los ciudadanos, en términos de eficiencia y precisión. Estas herramientas están llamadas a sustentar la toma de decisiones administrativas o a ser fuente de decisiones automatizadas. Su funcionamiento depende en gran medida de instrucciones precisas y suficientes para empezar a operar y de los datos de referencia utilizados para entrenar el sistema. De ahí la

²⁹ Considerando 38 del AIA.

³⁰ CENTRAL DIGITAL AND DATA OFFICE, "Algorithmic Transparency Recording Standard", Disponible en: <https://www.gob.uk>



importancia de que tanto las decisiones automatizadas, como las tomadas con apoyo en sistemas de IA por las Administraciones Públicas respeten el principio de transparencia. En nuestro Estado de Derecho, todos, ciudadanos y poderes públicos, estamos sujetos al principio de legalidad (la Constitución y el resto del ordenamiento jurídico, artículo 9.1 CE), que implica el deber inexcusable de respetar los derechos fundamentales. Asimismo, como Estado democrático, nuestro sistema fomenta el control eficaz de la actuación de los poderes públicos por parte de la ciudadanía.

Sin embargo, en la práctica, como hemos podido apreciar, el principio de transparencia se encuentra con demasiadas limitaciones. En primer lugar, la opacidad que rodea a los sistemas algorítmicos, por la escasa información que se tiene sobre cuales son y cómo funcionan. A esto cabe añadir la dificultad para comprender cómo actúan, dada su enorme complejidad. A pesar de estos inconvenientes, el uso de la IA no debe estigmatizarse, sino más bien fomentarse, profundizando en la necesidad de unos mecanismos de seguridad y de garantía para el respeto a los derechos fundamentales. De hecho, estos procedimientos informáticos conllevan numerosas ventajas, como la reducción del tiempo para operaciones meramente repetitivas en las decisiones automatizadas, o la mayor percepción de neutralidad en las decisiones basadas en predicciones algorítmicas.

En cuanto a los mecanismos de seguridad, creemos que es necesario incentivar la publicidad activa de los sistemas algorítmicos empleados por las Administraciones Públicas de impacto social, con la descripción de manera comprensible de su diseño y funcionamiento. Igualmente, potenciar la obligación de realizar evaluaciones de impacto como medida de prevención que sirva a proveedores y usuarios de la IA para identificar los riesgos, tomar medidas técnicas y organizativas desde el diseño, para mitigar los daños a los derechos fundamentales y supervisar su evolución. A tal efecto, es urgente requerir una correcta formación del personal al servicio del sector público que va a gestionar y controlar el uso de las herramientas de IA, para que tome conciencia de los efectos discriminatorios que estas decisiones pueden conllevar y comprenda mejor los riesgos que puede provocar su uso sobre los derechos fundamentales de los ciudadanos. Y, en último término, desarrollar metodologías que permitan auditar los algoritmos que tomen decisiones directa o indirectamente sobre las personas, para comprobar que funcionan con corrección y detectar su posible vulnerabilidad.

Desde el punto de vista jurídico, es importantísima la labor que está desarrollando la UE en este campo. Al propio tiempo, España no se ha quedado atrás con la aprobación de la Ley 15/2022 de 12 de julio, integral para la igualdad de trato y la no discriminación. Sin embargo, aún queda mucho por hacer, puesto que, en esta materia, como en tantas otras, el derecho va siempre por detrás de la tecnología. Las normas de protección de datos han servido hasta ahora como respuesta a los retos que ha planteado el fenómeno de la IA, pero no son suficientes. Hay que tomar nota de algunas de sus garantías para aplicarlas a los sistemas de inteligencia artificial,



como la obligación de realizar evaluaciones de impacto o el derecho a la supervisión humana. El desarrollo del principio de transparencia y el derecho a una buena administración pueden servir en el sector público para ampliar las garantías frente a las decisiones automatizadas tomadas por las Administraciones. Esto se ha puesto de manifiesto por la jurisprudencia comparada, junto con el principio de no discriminación. La importancia de la transparencia es un hecho porque hay que tener en cuenta las dificultades de construir una mínima prueba de discriminación algorítmica debido a la opacidad de estos sistemas, que incluso pueden conllevar a que la persona afectada ni siquiera tenga constancia del trato discriminatorio. De todo ello resulta, la importancia de que las personas que gestionen estos sistemas, que no serán juristas en su mayoría, tengan una formación específica en la protección de los derechos fundamentales y en ética algorítmica, que tenga en cuenta los riesgos que el diseño de estos sistemas conlleva para los derechos a la igualdad y no discriminación y los valores y principios democráticos.

Bibliografía

- BAZ LOMBA, C. 2021. "Los algoritmos y la toma de decisiones administrativas. Especial referencia a la transparencia", *Revista CEFlegal*, núm.243: 119-160.
- CERRILLO I MARTINEZ, A. 2020. "¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?", *European Review of Digital Administration & Law* . Vol.1: 17-36.
- (2019) "El impacto de la inteligencia artificial en el Derecho administrativo ¿Nuevos conceptos para nuevas realidades técnicas?", *Revista General de Derecho Administrativo*, núm.50:1-38.
- COTINO HUESO, L .2020. "Holanda "SyRI, ¿a quién sanciono? Garantías frente al uso de la Inteligencia Artificial y decisiones automatizadas en el sector público y la sentencia holandesa de 2020". *La ley privacidad*, núm.4.
- DE LA SIERRA MORÓN, S. 2020. "Inteligencia artificial y justicia administrativa: una aproximación desde la teoría del control de la Administración Pública", *Revista General de Derecho Administrativo*, núm.53: 1-19.
- GAMERO CASADO, E. 2021. "El enfoque europeo de inteligencia artificial", *Revista de Derecho Administrativo- CDA*, núm.20:268-289.
- HUERGO LORA, A. 2020. "Una aproximación a los algoritmos desde el Derecho Administrativo", en HUERGO LORA, A (dir.) *La regulación de los algoritmos*, Aranzadi, pp.23-87.
- (2021) "Administraciones Públicas e Inteligencia Artificial: ¿más o menos discrecionalidad?", *El Cronista del Estado Social y Democrático de Derecho*, núms.96-97:78-95.



- MARTÍN DELGADO, I. 2009. "Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada", *Revista de Administración Pública*, núm.180:353-386.
- MARTÍNEZ GUTIÉRREZ, R. 2021. "Inteligencia artificial, algoritmos y automatización en la Justicia. Propuestas para su efectiva implantación", *Práctica de Tribunales: Revista de Derecho Procesal Civil y Mercantil*, núm.149:1-20.
- MEDINA GUERRERO, M. 2022. "El derecho a conocer los algoritmos utilizados en la toma de decisiones". *Teoría y realidad constitucional*, núm. 49:141-171.
- PALMA ORTIGOSA, A. 2019. "Decisiones automatizadas en el RGPD. El uso de algoritmos en el contexto de la protección de datos", *Revista General de Derecho Administrativo*. núm. 50:1-35.
- PONCE SOLÉ, J. 2019. "Inteligencia artificial, Derecho Administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico", *Revista General de Derecho Administrativo*, núm.50:1-52.
- ROIG, A. 2020. *Las garantías frente a las decisiones automatizadas: del Reglamento General de Protección de Datos a la gobernanza algorítmica*. Barcelona, Bosch.
- SORIANO ARNAZ, A. 2021. "Decisiones automatizadas: problemas y soluciones jurídicas. Más allá de la protección de datos", *Revista de Derecho Público: Teoría y Método*, Vol. 3: 85-127.
- 2022. "Discriminación algorítmica: garantías y protección jurídica" en COTINO HUESO, L. *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*. Cizur Menor. Aranzadi, pp.134-169.
- VELASCO RICO, C.I. 2022. "Las tecnologías disruptivas en la administración pública: Inteligencia Artificial y Block Chain", en coord. CASTILLO RAMOS-BOSSINI, S.E y CERRILLO I MARTÍNEZ, A (Dirs.): *La Administración digital*. Dykinson. pp.227-256.