



La proyección del Reglamento de IA sobre la Administración Pública: el Principio de explicabilidad aplicado al procedimiento administrativo^(*)(**)

The Impact of the AI Regulation on Public Administration: The Principle of Explainability in Administrative Procedures

Patricia Mendilibar Navarro^(*)**

Universidad CEU Cardenal Herrera (Valencia, España)

Resumen: El Reglamento de IA ha supuesto un paso adelante en la regulación de la inteligencia artificial en la Unión Europea. En este trabajo, se analiza el impacto de esta norma en el uso de IA por la Administración Pública, particularmente, respecto de la explicabilidad y la motivación de las decisiones tomadas con sistemas de IA. El trabajo atiende a la insuficiencia del Reglamento para abordar los riesgos del uso de la IA en el procedimiento administrativo y se identifican las deficiencias del artículo 86 del Reglamento en relación con la necesidad de motivar y de explicar determinados elementos del procedimiento administrativo algorítmico para que sea comprensible para la ciudadanía.

Palabras clave: Inteligencia Artificial - Reglamento de IA - Procedimiento Administrativo - Explicabilidad – Motivación - Derecho Administrativo - España

Abstract: The AI Regulation has been a step forward in the regulation of artificial intelligence in the European Union. This paper analyzes the impact of this regulation on the use of AI by the public administration, particularly with regard to the explainability and motivation of decisions made with AI systems. The paper addresses the Regulation's failure to address the risks of using AI in administrative procedures and identifies the shortcomings of Article 86 of the Regulation in relation to the need to justify and explain certain elements of the algorithmic administrative procedure so that it is understandable to citizens.

Keywords: Artificial Intelligence - AI Regulation – Administrative Procedure – Explainability - Motivation - Administrative Law - Spain

(*) Nota del Equipo Editorial: Este artículo fue recibido el 24 de octubre de 2025 y su publicación fue aprobada el 19 de diciembre de 2025.

(**) Este artículo forma parte del Proyecto de I+D+i "Ética y Autorregulación de la Comunicación Social: Análisis de contenido de los Códigos Éticos de 2ª Generación y elaboración de Protocolos y Guías para su implementación", Ref. PID2021-124969NB-I00, financiado por MCIN/AEI/10.13039/501100011033/ y "FEDER Una manera de hacer Europa".

(***) Abogada por la Universidad CEU Cardenal Herrera (Valencia, España). Doctora por la Universidad CEU Cardenal Herrera. Profesora de Derecho Administrativo y Derecho de la Comunicación en el Departamento de Ciencias Jurídicas de la misma universidad. Entre sus líneas de investigación destaca el estudio de la relación entre el Derecho y las nuevas tecnologías, especialmente el impacto de la inteligencia artificial en el sector público. ORCID: <https://orcid.org/0009-0001-4672-8254>. Correo electrónico: patricia.mendilibar@uchceu.es.



1. Introducción

El uso de la inteligencia artificial (en adelante, “IA”) por parte de la Administración Pública española es una realidad. Aunque la Administración ha sido criticada históricamente por su obsolescencia, caracterizada por la tramitación en papel, la necesidad de citas presenciales y la burocratización de sus procesos, la Administración ha comenzado a adoptar una de las herramientas más avanzadas y poderosas del siglo XXI que le permitirá acabar con esta lacra: la IA.

A pesar de las asimetrías que presenta el uso de la IA en el sector público en función del área examinada⁽¹⁾, se advierte una progresiva delegación en el algoritmo para el dictado de actos administrativos porque, a pesar de que no existe un mapeo de usos de la IA en el sector público, se trata de una herramienta ampliamente utilizada por la Administración para múltiples finalidades.

En ese sentido, los usos de la IA para el dictado de actos administrativos abarcan dos tipos de decisiones (Huergo Lora, 2020). Por un lado, las actuaciones administrativas automatizadas, que son aquellas realizadas íntegramente a través de medios electrónicos por una Administración Pública en el marco de un procedimiento administrativo y en la que no interviene de forma directa un empleado público (artículo 41 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público). Por otro lado, las predicciones algorítmicas, término acuñado por Huergo Lora (2020), que son aquellas que predeterminan una decisión administrativa, como se desarrollará en este trabajo. Por tanto, la IA puede ser utilizada en diferentes fases del procedimiento administrativo y con múltiples finalidades.

La aprobación de una norma en esta materia resultaba necesaria, a nuestro juicio, por diversos motivos: en primer lugar, para impulsar la innovación y la digitalización dentro de un marco normativo que asegure el cumplimiento de las restantes normas que afectan a esta materia (como la protección de datos, la accesibilidad o las normas sectoriales materia de propiedad intelectual o derecho de la competencia, entre otros); en segundo lugar, para la estandarización de esta tecnología (Cerrillo I Martínez et al., 2020); y en tercer lugar, para dotar a la ciudadanía de mecanismos de acción eficaces frente al uso indebido de la IA, en particular, para el control de las decisiones discrecionales (Chaves, 2016).

Como consecuencia de todo ello, la norma incentiva a la Administración Pública a la utilización de estas tecnologías que, bien utilizadas, pueden redundar en la eficacia y eficiencia de las Administraciones, como Principios rectores regulados en la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público (en adelante, “Ley 40/2015”). Por ello, podemos afirmar que la vaguedad jurídica y la falta de una norma en esta materia ha jugado en detrimento de los derechos de la ciudadanía en este ámbito, tanto por las ineficiencias que puede encontrar el administrado frente a una Administración resistente al cambio como por el uso indiscriminado de IA por parte de las instancias públicas.

Por este motivo, la aprobación del Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo de 13 de junio de 2024 por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) 300/2008, (UE) 167/2013, (UE) 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (en adelante, el “Reglamento de IA”) merece, *a priori*, una valoración positiva para coadyuvar a la toma de decisiones basadas en algoritmos en el sector público y para dotar de mayor seguridad a los ciudadanos que sean objeto de decisiones algorítmicas.

A pesar de que se trata de un avance muy relevante y de una norma pionera en esta materia, el Reglamento de IA establece muy pocas previsiones en relación con el uso de la IA en el sector público y las exigencias que se deben imponer a las Administraciones Públicas cuando utilicen esta tecnología. De hecho, el Reglamento ha optado por un nuevo modelo de regulación propio de situaciones de incertidumbre en la que no se quiere dejar de regular una realidad, pero a la vez, no se

(1) Se observan asimetrías en función del ámbito en el que nos encontremos. Por ejemplo, en el ámbito tributario, se han implementado muchas herramientas de IA que permiten agilizar los procedimientos. Para profundizar en ello, véase, “*Big data e inteligencia artificial en la Administración tributaria*”, de Oliver Cuello (2021). O, también, en las investigaciones policiales, por ejemplo, para valorar el riesgo de reincidencia en el ámbito de la violencia de género. Sobre este tema, véase, “*Policía predictiva en España. Aplicación y retos de futuro*”, de (González- Álvarez, et al. 2020). Sin embargo, existen otros ámbitos, como la gestión que, por el momento, son impenetrables.



puede abordar con una regulación prematura que impida el desarrollo de una nueva tecnología emergente (De la Quadra-Salcedo, 2024). Se trata de una norma programática que intenta adaptarse a una situación de cambio constante.

Respecto al sector público, se advierte la poca profundización y exigencias que el Reglamento impone a la Administración cuando emplea esta tecnología, más allá de la evaluación de impacto en los derechos fundamentales cuando emplee sistemas de IA de alto riesgo. Esto se explica por el Principio de autonomía institucional, que supone la potestad que tienen los Estados para establecer parcelas de organización y procedimiento, “a fin de que sus legisladores y Administraciones puedan adoptar las soluciones formales que más les convengan, a partir de los Principios matrices de subsidiariedad y proporcionalidad” (Ballesteros, 2011).

Este principio reconoce que, aunque los Estados transfieren parte de su soberanía a la Unión Europea en áreas específicas, conservan un amplio margen para estructurar sus instituciones y procesos administrativos, siempre y cuando respeten los valores fundamentales de la UE, como la democracia, el Estado de Derecho y los derechos fundamentales (Díez-Picazo, 2004).

La vigencia del Principio de autonomía institucional explica que la Unión Europea haya apostado por no imponer un marco rígido y uniforme para el uso de la IA en la Administración, sino que permite que los Estados Miembros decidan en qué áreas y bajo qué condiciones implementen la IA, de manera flexible y adaptada a sus necesidades.

No obstante, esto no significa que la Administración Pública quede al margen de las obligaciones impuestas por el Reglamento. Por este motivo, es necesario realizar una reflexión sobre las disposiciones que resultan aplicables al sector público en esta materia, cómo se va a aplicar el Reglamento de IA a la Administración, con qué intensidad y qué papel juega, especialmente, el principio de explicabilidad de la IA (regulado en el artículo 86 del Reglamento) en el ámbito administrativo —teniendo en cuenta que la motivación de las decisiones administrativas es un derecho de la ciudadanía—.

La explicabilidad se introdujo en el Reglamento a partir de las Directrices éticas para una IA fiable, en las que el Grupo de expertos de alto nivel sobre inteligencia artificial de la Comisión Europea alertó de que las personas solo podrían confiar en el desarrollo tecnológico si se contaba con una regulación que asegurase su fiabilidad. Así, el Grupo de Expertos señalaba acertadamente que:

La explicabilidad es crucial para conseguir que los usuarios confíen en los sistemas de IA y para mantener dicha confianza. Esto significa que los procesos han de ser transparentes, que es preciso comunicar abiertamente las capacidades y la finalidad de los sistemas de IA y que las decisiones deben poder explicarse —en la medida de lo posible— a las partes que se vean afectadas por ellas de manera directa o indirecta. Sin esta información,

no es posible impugnar adecuadamente una decisión. No siempre resulta posible explicar por qué un modelo ha generado un resultado o una decisión particular (ni qué combinación de factores contribuyeron a ello). Esos casos, que se denominan algoritmos de “caja negra”, requieren especial atención. En tales circunstancias, puede ser necesario adoptar otras medidas relacionadas con la explicabilidad (por ejemplo, la trazabilidad, la auditabilidad y la comunicación transparente sobre las prestaciones del sistema), siempre y cuando el sistema en su conjunto respete los derechos fundamentales. El grado de necesidad de explicabilidad depende en gran medida del contexto y la gravedad de las consecuencias derivadas de un resultado erróneo o inadecuado (2019).

La explicabilidad, tal y como estaba concebida por el Grupo de expertos, inspiró el marco normativo del Reglamento de IA para determinar en qué consiste el “derecho a una explicación”. Por ejemplo, el artículo 10 impone la obligación de implementar prácticas adecuadas de gobernanza y gestión de los datos; el artículo 11 requiere la existencia de documentación técnica; el artículo 12 la conservación de los registros la conservación de registros a lo largo de todo el ciclo de vida del sistema de IA para garantizar un nivel adecuado de trazabilidad de su funcionamiento; y el artículo 13, que exige que los sistemas de IA se diseñen de un modo que garanticen que funcionan con un nivel de transparencia suficiente para que los responsables del despliegue interpreten y usen correctamente su información de salida; finalmente, el artículo 14, que indica que el sistema de IA debe diseñarse de tal manera que la persona física a la que se encomiende su vigilancia pueda llevar a cabo adecuadamente su labor (López-Tarruella, 2024). En fin, el cumplimiento de los requisitos establecidos en los artículos 11 a 14 permite cumplir con el derecho a una explicación; explicación que, por otro lado, es especialmente relevante en el ámbito público por los derechos implicados.

Teniendo en cuenta todo lo anterior, en este trabajo se analiza, como punto de partida, la proyección del Reglamento de IA en la Administración Pública, ya que se parte de la hipótesis de que el uso de sistemas de IA en el ejercicio de potestades públicas plantea determinados problemas que no



pueden ser resueltos por una norma de alcance horizontal como el Reglamento.

En base a las conclusiones alcanzadas, el trabajo se centra en el principio de explicabilidad, porque constituye el punto de encuentro más relevante entre el régimen europeo de la IA y las garantías clásicas del Derecho administrativo. A nuestro juicio, la explicabilidad es un instrumento de transparencia técnica que resulta imprescindible para la motivación de los actos administrativos exigida en el artículo 35.1 de la Ley 39/2015, Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas; el control de la discrecionalidad administrativa previsto en el artículo 3.1 de la Ley 40/2015; y para el adecuado ejercicio del derecho de defensa y la tutela judicial efectiva, consagrados en el artículo 24.1 de la Constitución Española.

Así, los interrogantes que se pretenden resolver con el siguiente trabajo atienden a cómo se proyecta el Reglamento de IA sobre la Administración Pública española; si el principio de explicabilidad, tal y como está configurado en el artículo 86 del Reglamento, es suficiente para garantizar el control de las decisiones administrativas algorítmicas; y, finalmente, en qué medida el Derecho administrativo interno colma las lagunas del marco europeo analizado.

2. El ámbito de aplicación del Reglamento: especial referencia a la Administración Pública

2.1. Proveedor y responsable del despliegue en la Administración Pública

El Reglamento Europeo de IA establece un marco normativo que equipara el uso de IA en la Administración Pública con su uso en el sector privado. Esto implica que la Administración Pública española, en la medida en que actúe como proveedor o responsable del despliegue de un sistema de IA, queda sujeta a las mismas disposiciones que las empresas privadas. No se prevén distintos regímenes normativos en función de si el uso de la IA se realiza en el seno de una entidad pública o privada.

Esta falta de distinción no responde a la realidad en la que la norma debe aplicarse, ya que las implicaciones jurídicas y los derechos afectados por las decisiones públicas no son equiparables a las consecuencias de la aplicación de la IA en el ámbito privado. Las decisiones administrativas implican, por su propia naturaleza, a la dimensión pública de la persona (al pago de tributos, a la gestión sanitaria, o a la concesión de prestaciones sociales, entre otras). En consecuencia, la necesidad de que las decisiones administrativas sean transparentes es fundamental en el ámbito público (más, incluso, que en el ámbito privado). Como se analiza a continuación, la aplicación de un mismo régimen normativo puede resultar insuficiente para abordar las particularidades

del uso de la IA en la Administración Pública y, desde luego, el Reglamento de IA no regula esta cuestión en profundidad.

En el artículo 3 apartado 3 del Reglamento se incluye en la definición de proveedor a las autoridades públicas que desarrollen “un sistema de IA o un modelo de IA de uso general o para el que se desarrolle un sistema de IA o un modelo de IA de uso general y lo introduzca[n] en el mercado o ponga en servicio el sistema de IA con su propio nombre o marca, previo pago o gratuitamente”.

Es decir, sería el supuesto en el que la Administración Pública, *per se* o a través de otro, desarrolla un sistema de IA destinado a ser utilizado dentro de la estructura pública o por otros actores, públicos o privados. Esto podría incluir la creación de modelos de IA de uso general, es decir, modelos diseñados para ser aplicados en una amplia variedad de contextos y que pueden ser adaptados a las necesidades de otro tipo de entidades. Se trataría del desarrollo de un sistema de IA para la gestión de datos en la sanidad pública, que luego se pondría a disposición de hospitales de titularidad privada.

En el artículo 3 apartado 4 del Reglamento, se define como responsable del despliegue, entre otros, a aquella “autoridad pública, órgano u organismo” que “utilice un sistema de IA bajo su propia autoridad”. En este supuesto estaríamos hablando de un sistema de IA que ha sido desarrollado por un tercero, que puede ser una empresa privada, una institución académica o incluso otra Administración Pública y que es implementado por una Administración Pública.

Como se anticipaba, uno de los principales problemas de esta perspectiva regulatoria es que no distingue adecuadamente entre los riesgos inherentes a la utilización de IA en la Administración y en el sector privado. La normativa está claramente influenciada por el enfoque de las normas de seguridad de productos, que prioriza aspectos como la seguridad técnica, pero no aborda adecuadamente los riesgos específicos asociados al uso de IA en la toma de decisiones administrativas.

De hecho, la Administración queda sujeta a una serie de garantías que el Reglamento de



IA ni siquiera menciona (por ejemplo, la sujeción al Principio de legalidad, de responsabilidad, de transparencia o de eficacia y eficiencia). Y ello porque la aplicación de IA en la Administración Pública está vinculada al ejercicio de poder público y la prestación de servicios esenciales, lo que amplifica los posibles impactos en derechos de los administrados.

En definitiva, aunque podría haberse empleado la misma definición para referirse al ámbito subjetivo de aplicación del Reglamento, sí que se debería haber hecho una mención específica a la Administración Pública como responsable del despliegue o como proveedora de servicios de IA, especificando las cautelas a las que queda sujeta, entre otras, la necesidad de seguridad jurídica, el Principio de legalidad, la transparencia, la obligación de motivar sus actos y la rendición de cuentas, que se analizarán posteriormente.

2.2. Prestación de servicios públicos por entidades privadas

El Reglamento de IA también regula la prestación de servicios públicos por parte de entidades privadas, y destaca que estas entidades, cuando desempeñan funciones de interés público, están sujetas a los mismos requisitos que las Administraciones Públicas. En su artículo 27 se establece que, tanto los organismos de derecho público como las entidades privadas que prestan servicios públicos deberán elevar el estándar de seguridad y realizar una evaluación de impacto relativa a los derechos fundamentales antes de poner en funcionamiento los sistemas de IA.

Este artículo reconoce que los servicios públicos esenciales, como la educación, la asistencia sanitaria, los servicios sociales y la administración de justicia, a menudo pueden ser prestados por entidades privadas (artículo 27 del Reglamento). Estas entidades, al estar vinculadas al interés público, deben cumplir con las mismas exigencias regulatorias que las Administraciones Públicas en cuanto a la evaluación de los riesgos y el impacto sobre los derechos fundamentales. Esta evaluación será fundamental para identificar y mitigar los riesgos que la IA puede generar para la ciudadanía.

2.3 Disposiciones aplicables a la Administración Pública

El Reglamento establece obligaciones específicas para los sistemas de IA utilizados en diversos ámbitos del sector público, principalmente, aquellos clasificados como de alto riesgo. Estos sistemas pueden tener implicaciones en áreas que tienen una especial afección para la ciudadanía, lo que exige una regulación más estricta y, en particular, la realización de una evaluación sobre el impacto en los derechos fundamentales. No se va a profundizar sobre las implicaciones en cada ámbito por exceder del objeto del presente trabajo, pero se señalan a continuación las características del uso de la IA por ámbitos.

El Reglamento de IA clasifica como **sistemas de alto riesgo** a aquellos utilizados en áreas críticas, donde su implementación

podría afectar directamente los **derechos fundamentales** de la ciudadanía.

En primer lugar, los sistemas de IA utilizados en la Administración de Justicia son considerados de alto riesgo debido a su capacidad para influir en la interpretación y aplicación de la ley. Estos sistemas no pueden reemplazar la decisión adoptada por la persona humana, sino que solo deben asistir a jueces y funcionarios judiciales. El artículo 6 del Reglamento establece que: “Los sistemas de IA utilizados para asistir en la investigación e interpretación judicial serán considerados de alto riesgo. Su implementación deberá ser precedida de una evaluación detallada de impacto en derechos fundamentales [...]”.

En este sentido, el artículo 6 reconoce que todos los sistemas de IA empleados en la justicia deben pasar por una evaluación exhaustiva de impacto antes de ser implementados. Este proceso es esencial para garantizar que el uso de IA no comprometa la independencia judicial ni la tutela judicial efectiva del artículo 24 de la Constitución Española.

En segundo lugar, el uso de IA en procesos de gestión de fronteras, migración y solicitudes de asilo también está sujeto a una regulación específica debido al impacto potencial sobre los derechos fundamentales, como el derecho al asilo o el Principio de no devolución. El Reglamento establece que los sistemas de IA utilizados en estos contextos deben ser evaluados para garantizar que no violen las obligaciones internacionales de la Unión Europea. El artículo 7 dispone literalmente que “los sistemas de IA empleados en el control de fronteras y procesos de asilo deberán ser considerados de alto riesgo. Las autoridades deberán garantizar que estas herramientas no vulneren los derechos de los solicitantes”.

En tercer lugar, los sistemas de IA utilizados en la prestación de servicios públicos esenciales, como sanidad, educación y vivienda se consideran de alto riesgo, ya que pueden influir directamente en el acceso a estos derechos fundamentales, en el Principio de igualdad y en la distribución de recursos. El Reglamento obliga a las autoridades públicas que utilicen estos sistemas a realizar



una evaluación de impacto en derechos fundamentales antes de su implementación. El artículo 8 regula este ámbito con el siguiente tenor literal: “Los sistemas de IA que intervienen en la asignación de servicios sanitarios o educativos deben considerarse de alto riesgo. La evaluación de impacto garantizará que estos sistemas respeten los Principios de igualdad de oportunidades y no discriminación”.

Finalmente, la identificación biométrica en espacios públicos, especialmente cuando se realiza en tiempo real, está fuertemente regulada debido a su potencial invasivo. Aunque su uso está restringido, puede permitirse en situaciones excepcionales, como la prevención de delitos graves o el terrorismo.

Los sistemas de identificación biométrica en tiempo real, especialmente en espacios públicos, son considerados de alto riesgo en el Reglamento de IA, ya que su uso puede implicar una vigilancia masiva y una invasión significativa de la privacidad de los ciudadanos. Estos sistemas incluyen tecnologías de reconocimiento facial, escaneo de iris o huellas dactilares, que pueden ser utilizados por autoridades públicas para la seguridad y la prevención de delitos.

El Reglamento impone restricciones muy estrictas sobre el uso de estos sistemas, permitiendo que los Estados miembros introduzcan excepciones por ley que permitan el uso de la identificación biométrica en casos como actividades policiales relacionadas con determinados delitos, búsqueda selectiva de víctimas específicas, secuestros, trata y explotación sexual de seres humanos y personas desaparecidas o prevención de amenazas para la vida o la seguridad física de las personas o respuesta a la amenaza actual o previsible de un ataque terrorista.

En definitiva, el Reglamento de IA clasifica el uso de identificación biométrica en tiempo real en espacios públicos como una práctica de alto riesgo y sujeta a una estricta regulación y supervisión. Las autoridades públicas solo pueden implementar estos sistemas en situaciones excepcionales y en unas condiciones que garanticen el respeto a los derechos fundamentales.

3. La evaluación de impacto en los derechos fundamentales

Como se ha analizado en el apartado anterior, los organismos públicos responsables del despliegue de sistemas de IA de alto riesgo, así como entidades privadas que presten servicios públicos en sectores como la educación, la salud y la justicia, deben realizar una evaluación de impacto antes de poner en marcha dichos sistemas (artículo 27 del Reglamento de IA). Como hemos visto, cuando la IA se emplee en determinadas áreas críticas, deberá ser sometida a una evaluación de impacto en los derechos fundamentales.

La Administración Pública no es ajena a la evaluación de impacto cuando adopta decisiones, por ejemplo, en materia ambiental o en materia de género. Por ejemplo, la Ley 21/2013, de 9 de diciembre, de evaluación ambiental, exige que los proyectos públicos o privados con potencial impacto ambiental sean evaluados antes de su aprobación, garantizando la protección del medio ambiente. Asimismo, la Ley Orgánica 3/2007 establece la obligación de realizar informes de impacto de género para evaluar la igualdad entre mujeres y hombres cuando se adopten determinadas políticas públicas. Sin embargo, la evaluación de impacto en los derechos fundamentales, como la exigida por el Reglamento de IA, introduce un enfoque novedoso, ya que obliga a analizar cómo el uso de la IA afecta directamente derechos fundamentales como el derecho a la intimidad, a la protección de datos o al derecho a la igualdad y no discriminación.

Como ha señalado la doctrina (Laukyte, 2024), la evaluación no solo tiene que considerar los efectos potenciales de la IA sobre los derechos individuales, sino también el impacto en grupos vulnerables y colectivos que podrían verse afectados de manera desproporcionada.

Sin embargo, esta evaluación resulta deficiente por diversos motivos (Laukyte, 2024):

- En primer lugar, el Reglamento de IA indica que la evaluación debe ser revisada si hay cambios en los factores de riesgo, pero la decisión de realizar una nueva evaluación depende del responsable del despliegue. Es decir, se realiza una evaluación sobre la primera versión del sistema, pero, posteriormente, la necesidad de evaluar las actualizaciones será ponderada por el responsable del despliegue. El problema es que los responsables del sistema podrían no detectar o ignorar cambios significativos.
- El artículo 27 apartado 2 establece que, “en casos similares, el responsable del despliegue podrá basarse en evaluaciones de impacto relativas a los derechos fundamentales realizadas previamente o a evaluaciones de impacto existentes realizadas por los proveedores”. Sin



embargo, el Reglamento de IA no define qué constituye un “caso similar”, lo que da lugar a inseguridad jurídica y a la posibilidad de que no se evalúen correctamente todos los casos que deben ser evaluados.

- También existe el riesgo de evaluaciones superficiales, ya que la evaluación inicial se realiza por los responsables del despliegue o los proveedores, que tienen un interés en que sus sistemas sean aprobados. La sociedad civil ha criticado que no existe una supervisión efectiva e independiente de estas evaluaciones, lo que aumenta el riesgo de evaluaciones superficiales o sesgadas.
- Finalmente, a pesar de la obligación de realizar una evaluación de impacto, la Ley no proporciona mecanismos de reparación claros para los ciudadanos cuyos derechos fundamentales puedan verse vulnerados por un sistema de IA mal evaluado o que cause daños una vez implementado. Esto crea una laguna en la protección efectiva de los derechos, ya que los afectados podrían encontrarse sin vías claras para reclamar o exigir cambios en el uso de los sistemas de IA.

El Reglamento también menciona la necesidad de involucrar a expertos, organizaciones de la sociedad civil y grupos afectados en la realización de estas evaluaciones. Este enfoque busca asegurar que se consideren todas las perspectivas posibles, se introduce la participación pública en la evaluación y se mitigan los riesgos respecto a grupos vulnerables.

Sin embargo, este punto ha sido criticado por la falta de mecanismos claros para facilitar la participación efectiva de la sociedad civil y garantizar que sus preocupaciones sean tomadas en cuenta en el diseño y despliegue de sistemas de IA (Laukyte, 2024).

En cualquier caso, además, debe tenerse en cuenta que debe realizarse una evaluación de impacto en los derechos fundamentales real. Esta obligación no quedaría satisfecha, en nuestra opinión, mediante la mera cumplimentación de un formulario en la que se indique si afecta o no afecta. En este sentido, debe tenerse en cuenta la doctrina sentada por la Sentencia del Tribunal Supremo que resuelve el recurso de casación interpuesto por la Generalitat Valenciana contra la anulación del Plan de Acción Territorial de la Infraestructura Verde del Litoral (en adelante, “PATIVEL”) (Sentencia de la Sala Tercera del Tribunal Supremo de 27 de abril de 2022 (n.º rec. 4049/2021)).

Justamente, uno de los motivos por los que se anuló el PATIVEL fueron las deficiencias en los informes de impacto de género, familia e infancia. La doctrina sentada por esta Sentencia resulta aplicable analógicamente al supuesto que nos ocupa:

Si en el caso enjuiciado cabe deducir que realmente se llevó a cabo el análisis del impacto que las determinaciones del plan de ordenación territorial podrían tener sobre el género, familia, infancia y adolescencia, los informes “neutros” no serán

equivalentes a inexistentes; por el contrario, cuando pueda inferirse que se ha utilizado una fórmula rituarial para afirmar la neutralidad de la norma a esos efectos, pero que no ha existido realmente un análisis sobre el particular, esos informes deben reputarse inexistentes (2021).

Por tanto, el Tribunal establece que estos informes deben incluir un análisis real del impacto del plan en estas materias, y no una simple afirmación de que no hay afectación.

Además, a pesar de que la norma habla de “derechos fundamentales”, se considera que el más implicado va a ser el derecho a la igualdad en conexión con otros derechos, pero el principio de no discriminación debería ser, a nuestro juicio, la piedra angular de este informe. Por ejemplo, se podrá infringir el derecho a la igualdad en relación con el derecho a la educación —si, por ejemplo, el algoritmo otorga una puntuación distinta a una persona para acceder a un colegio y esta puntuación se debe a una característica como el origen o la ideología—.

En cualquier caso, la evaluación de impacto se erige como una herramienta fundamental para garantizar que la implementación de IA respete los derechos de los ciudadanos. El Reglamento extiende esta obligación a las entidades privadas que prestan servicios públicos, asegurando que cualquier sistema de IA de alto riesgo que se utilice en este contexto esté sujeto a un análisis detallado antes de su implementación. No obstante, la deficiente redacción del Reglamento da lugar a determinados interrogantes y lagunas que han quedado expuestos y que deberán ser resueltos durante su aplicación y/o en la norma interna que se apruebe al efecto.

4. Otras obligaciones de la Administración Pública cuando actúe como responsable del despliegue

En cuanto a las obligaciones impuestas a la Administración como responsable del despliegue, el artículo 26 del Reglamento (2024) obliga a lo siguiente:

1. Los responsables del despliegue de sistemas de IA de alto riesgo adoptarán medidas técnicas y organizativas adecuadas para garantizar que utilizan



dichos sistemas con arreglo a las instrucciones de uso que los acompañen (apartado 1).

2. Los responsables del despliegue encomendarán la supervisión humana a personas físicas que tengan la competencia, la formación y la autoridad necesarias (apartado 2).
3. El responsable del despliegue se asegurará de que los datos de entrada sean pertinentes y suficientemente representativos en vista de la finalidad prevista del sistema de IA de alto riesgo, en la medida en que ejerza el control sobre dichos datos (apartado 4).
4. Los responsables del despliegue vigilarán el funcionamiento del sistema de IA de alto riesgo basándose en las instrucciones de uso y, cuando proceda, informarán a los proveedores. Cuando los responsables del despliegue tengan motivos para considerar que utilizar el sistema de IA de alto riesgo conforme a sus instrucciones puede dar lugar a que ese sistema de IA presente un riesgo, informarán, sin demora indebida, al proveedor o distribuidor y a la autoridad de vigilancia del mercado pertinente y suspenderán el uso de ese sistema. Cuando los responsables del despliegue detecten un incidente grave, informarán asimismo inmediatamente de dicho incidente, en primer lugar, al proveedor y, a continuación, al importador o distribuidor y a la autoridad de vigilancia del mercado pertinente (apartado 5). Estas obligaciones de información son especialmente relevantes en el caso del uso de la IA, para que la Administración Pública se pueda exonerar de responsabilidad en el caso de que el error no sea imputable a la misma.
5. Antes de poner en servicio o utilizar un sistema de IA de alto riesgo en el lugar de trabajo, los responsables del despliegue que sean empleadores informarán a los representantes de los trabajadores y a los trabajadores afectados de que estarán expuestos a la utilización del sistema de IA de alto riesgo (apartado 7).
6. Los responsables del despliegue de sistemas de IA de alto riesgo que sean autoridades públicas o instituciones, órganos y organismos de la Unión cumplirán las obligaciones de registro del Reglamento (apartado 8).

En definitiva, el Reglamento de IA no impone unas obligaciones reforzadas a la Administración Pública por el mero hecho de serlo, sino que queda sujeta a las genéricas obligaciones que el artículo 26 reserva a los responsables del despliegue.

Lo analizado hasta ahora ya pone de manifiesto las limitadas exigencias que el Reglamento impone a la Administración Pública. Esta insuficiencia es todavía más clara cuando se analiza el Reglamento desde la perspectiva de la explicabilidad, ya que el texto europeo no determina cómo deben hacerse comprensibles las decisiones administrativas adoptadas con la intervención de sistemas de IA.

5. El Principio de explicabilidad en el Reglamento de IA

5.1. Definición y carencias del Reglamento respecto a la explicabilidad

El Reglamento de IA parte de la premisa de que la explicabilidad es una consecuencia de la transparencia de los sistemas de IA. En parte, así es, ya que, si los algoritmos son tan complejos que no se comprende su funcionamiento interno, resulta imposible descifrar el resultado y, por tanto, explicar cómo se ha llegado hasta él en términos accesibles para los diferentes actores involucrados.

En el Considerando 27, el Reglamento de IA reconoce que, “por transparencia, se entiende que los sistemas de IA se desarrollan y utilizan de un modo que permita una trazabilidad y explicabilidad adecuadas [...]”. Aunque en el ámbito administrativo la explicabilidad —motivación, en términos administrativistas— y la transparencia son dos obligaciones independientes, interesa destacar que, en el ámbito de la IA, la transparencia permite conocer que una persona ha sido objeto de una decisión algorítmica y su funcionamiento, pero no siempre garantiza que su razonamiento pueda ser comprendido por los afectados. Sin embargo, la explicabilidad se refiere a que las personas puedan entender, de forma clara, el proceso lógico llevado a cabo por la IA para adoptar una decisión.

Así, el derecho a una “explicación”, se encuentra regulado en el artículo 86 del Reglamento de IA en los siguientes términos:

Toda persona que se vea afectada por una decisión que el responsable del despliegue adopte basándose en los resultados de salida de un sistema de IA de alto riesgo que figure en el Anexo III [...] tendrá derecho a obtener del responsable del despliegue explicaciones claras y significativas acerca del papel que el sistema de IA ha tenido en el proceso de toma de decisiones y los principales elementos de la decisión adoptada (2024).

Según este artículo, si se toma una decisión sobre una persona utilizando un sistema de IA de alto riesgo, y esa decisión tiene un impacto significativo en la salud, la



seguridad o los derechos básicos de la persona, ésta tiene derecho a una explicación clara de cómo participó el sistema de IA en el proceso de toma de decisiones. Sin embargo, esto no se aplica si existen excepciones o restricciones en la legislación de la Unión o nacional.

Este precepto ha sido criticado porque el Reglamento no establece qué elementos debe incluir esta explicación ni cómo se conjuga el Reglamento de IA con otras normas concurrentes, como la normativa en materia de protección de datos (Barrio Andrés, 2024). Particularmente, los motivos de crítica se articulan en torno a diversos fundamentos: en primer lugar, la explicabilidad no se predica respecto de todos los sistemas de IA, sino solo respecto a las decisiones adoptadas por los responsables de despliegue de sistemas de IA de alto riesgo que figuren en el Anexo III, con excepción de los enumerados en el punto 2. Los sistemas excluidos son, precisamente, los referidos a infraestructuras críticas. La importancia que tienen estas infraestructuras para los poderes públicos da lugar a que la gestión de su explicabilidad “quede enteramente en manos del Estado” (López-Turruella, 2024)⁽²⁾. En segundo lugar, no existe el derecho a una explicación en relación con los sistemas de IA que no son considerados de alto riesgo. Finalmente, este derecho solo lo ostentan aquellas personas en relación con una decisión que “produzca efectos jurídicos o le afecte considerablemente, del mismo modo, de manera que considere que tiene un efecto perjudicial para su salud, su seguridad o sus derechos fundamentales” (artículo 86.1 del Reglamento de IA), lo que exigirá un juicio casuístico en cada caso.

Asimismo, interesa destacar que, a pesar de la importancia dada a la explicabilidad en los documentos anteriores a la aprobación del Reglamento de IA (desde el Libro Blanco de 19 de febrero de 2020, pasando por la Propuesta de Reglamento (Artificial Intelligence Act) de la Comisión: COM(2021) 206 final, de 21 de abril de 2021, la redacción del Reglamento no facilita que la explicabilidad se cumpla, especialmente porque no quedan claros qué elementos de la decisión adoptada mediante IA deben ser motivados. Tampoco se regula el plazo en el que el responsable del despliegue debe ofrecer las explicaciones solicitadas, lo que resulta relevante para el ejercicio posterior de derechos al que se refiere

el Considerando 171 del Reglamento⁽³⁾. Tampoco se establece qué acciones legales se pueden emprender frente a la falta de explicaciones. El artículo 85 no regula un derecho a presentar una reclamación ante la autoridad de vigilancia del mercado, sino que, en la práctica la persona afectada deberá iniciar acciones ante la jurisdicción civil (si el responsable es una entidad privada) o ante la jurisdicción contencioso-administrativa (si el responsable es una Administración Pública, como el supuesto que nos ocupa) (López-Tarruella, 2024).

Otra cuestión que limita el ejercicio del derecho son los límites a la explicabilidad, que son los derechos de propiedad intelectual y la información confidencial, ya que numerosos preceptos del Reglamento de IA disponen que la información sobre el sistema de IA se establece “sin perjuicio de la necesidad de observar y proteger los derechos de propiedad intelectual e industrial, la información empresarial confidencial y los secretos comerciales, de conformidad con el Derecho nacional y de la Unión” (artículo 25.5; artículo 53.1 b). En esta tensión entre los intereses del responsable del despliegue —que van a ofrecer la menor cantidad de información para preservar los derechos de propiedad intelectual— y los derechos de la persona afectada, se debe tener en cuenta que la obtención de una explicación interesa al sujeto afectado pero también “a la sociedad en general por cuanto, gracias a ello, se favorece la detección de errores en los sistemas de IA que, al fin y al cabo, conllevan un beneficio general” (López-Turruella, 2024).

Junto a lo anterior, hay que tener en cuenta que muchos de los sistemas de IA son opacos, y es complicado explicar cuál ha

(2) También se excluyen los sistemas de IA que constituyen componentes de seguridad de los productos regulados por los actos legislativos enumerados en el Anexo I, que tienen menor relevancia en el ámbito de la Administración Pública; y tampoco se aplica el derecho a una explicación en relación con los sistemas de IA que no son considerados de alto riesgo.

(3) “Las personas afectadas deben tener derecho a obtener una explicación cuando la decisión de un responsable del despliegue se base principalmente en los resultados de salida de determinados sistemas de IA de alto riesgo que entran dentro del ámbito de aplicación del presente Reglamento y cuando dicha decisión produzca efectos jurídicos o afecte significativamente de modo similar a dichas personas, de manera que consideren que tiene un efecto negativo en su salud, su seguridad o sus derechos fundamentales. Dicha explicación debe ser clara y significativa y servir de base para que las personas afectadas puedan ejercer sus derechos. El derecho a obtener una explicación no debe aplicarse a la utilización de sistemas de IA para los que se deriven excepciones o restricciones con arreglo al Derecho nacional o de la Unión, y solo debe aplicarse en la medida en que este derecho no esté ya previsto en el Derecho de la Unión”.



sido el razonamiento seguido para adoptar una decisión. Por otro lado, Agustino y Guilayn (2024) considera que una mera explicación de los procesos técnicos podría ser inadecuada porque podría dar lugar a una sobrecarga de información o, en el sentido contrario, dar una explicación incompleta.

En definitiva, el derecho a una explicación, tal y como está configurado en el Reglamento de IA, ha sido limitado por diferentes preceptos legales. Solo se disfruta en relación con determinadas categorías de sistemas de IA de alto riesgo (los enumerados en el Anexo III con excepción del punto 2).

5.2. Definición y carencias del Reglamento respecto a la explicabilidad

La explicabilidad en el Reglamento cobra su pleno sentido si se pone en relación con otros preceptos y obligaciones. La explicabilidad garantiza, en último término, la transparencia y el control humano de las decisiones tomadas mediante IA. Por ello, debemos hacer una breve referencia a las obligaciones de transparencia, supervisión humana y trazabilidad y registro, por su íntima conexión con la explicabilidad.

Por un lado, la transparencia está prevista en los artículos 13, 52 y en los Considerandos 44, 47 y 70, que disponen que los sistemas de IA deben ser desarrollados y utilizados de manera que se facilite una comprensión clara de su funcionamiento. Esto implica proveer documentación e información accesible sobre el modelo algorítmico, sus capacidades y limitaciones, y advertir cuando los ciudadanos están interactuando con una IA. En concreto, el RIA exige informar a las personas afectadas de sus derechos cuando son objeto de una decisión automatizada y asegurar que los usuarios del sistema conocen, en líneas generales, cómo funciona y qué datos trata. Si el algoritmo es explicable, en los términos definidos en el apartado anterior, se garantiza la transparencia.

Por otro lado, la trazabilidad exige llevar registros del funcionamiento de los algoritmos de alto riesgo. Por ello, se deben conservar datos de entrenamiento, parámetros y resultados de las decisiones para posibilitar auditorías posteriores. La trazabilidad es condición de la explicabilidad porque, si no se registra el proceso interno de la IA, no se puede reconstruir *a posteriori* por qué el algoritmo llegó a una determinada conclusión.

Finalmente, el Reglamento impone mecanismos de supervisión humana, que se traduce en obligaciones de diseño y de capacitación de los usuarios. El propio Reglamento dispone que “los sistemas de IA de alto riesgo se diseñarán y desarrollarán de modo que puedan ser vigilados de manera efectiva por personas físicas durante el período que estén en uso, lo que incluye dotarlos de herramientas de interfaz humano-máquina adecuadas” (artículo 14). En otras palabras, la vigilancia humana y la interpretabilidad técnica son condiciones para la fiabilidad del sistema y para que

las decisiones automatizadas puedan ser entendidas y, si procede, corregidas por funcionarios. De hecho, lo que se verifica mediante la supervisión humana en el ámbito administrativo es que la decisión tomada mediante IA se ajusta al marco jurídico. De ahí que “el control supone una especie de revisión de la motivación (una motivación a la inversa), o una forma de remontar la motivación para comprobar si la decisión se ajusta al marco jurídico” (Chang, 2025).

6. El deber de motivación de los actos administrativos como concreción del Principio de explicabilidad

6.1. Los actos administrativos dictados mediante IA

Antes de profundizar sobre la proyección del Principio de explicabilidad en el procedimiento administrativo, se deben distinguir los distintos tipos de actos administrativos que pueden dictarse en un procedimiento administrativo en el que se ha utilizado IA.

Respecto al tipo de funciones que puede realizar la IA en el sector público, se debe distinguir cada una de ellas y qué objetivos se persiguen en cada caso. Asimismo, se deben distinguir dos categorías para determinar las consecuencias jurídicas que se van a desencadenar en el caso de que se empleen algoritmos para tomar decisiones en el seno de la Administración (Huerco, 2020). En relación con esto, resulta especialmente pertinente aludir al debate sobre la idoneidad del uso de sistemas de IA en decisiones discrecionales, en las que la Administración no se limita a aplicar las reglas preestablecidas en la norma. En ese sentido, diversas propuestas normativas recientes han tratado de regularlo. Por ejemplo, la Ley 2/2025 de 2 de abril, para el desarrollo e impulso de la inteligencia artificial en Galicia, establece que la IA solo se podrá utilizar cuando las Administraciones Públicas “ejercen potestades regladas y, excepcionalmente, potestades discrecionales, en caso de que el margen de discrecionalidad se hubiera agotado previamente en el diseño del sistema de inteligencia artificial” (artículo 50.1). Hay que tener en cuenta que las



potestades discrecionales incluyen las apreciaciones subjetivas del decisor que, en muchos casos, están basadas en criterios extrajurídicos, bien porque exigen que se interpreten conceptos jurídicos indeterminados, o bien porque exijan la aplicación de la diligencia de un buen padre de familia o la empatía. En cuanto a la empatía, debemos tener en cuenta que, aunque existen máquinas programadas para captar emociones, ellas mismas carecen de la “consciencia, emoción y humanidad precisas para no degenerar en un psicópata de sílice” (Ponce Solé, 2019). En este punto, existen dos corrientes doctrinales, la que niega la utilización de algoritmos para el ejercicio de potestades discrecionales y la que considera que sí que es posible.

En la primera corriente, según Ponce (2019) se aboga por una regulación que prohíbe la utilización de algoritmos para ejercer potestades discrecionales, en aras a respetar el principio de precaución, que establece las obligaciones de debido cuidado o debida diligencia asociadas al derecho a una buena administración exigen un determinado comportamiento administrativo en el procedimiento de adopción de decisiones administrativas. En la segunda corriente doctrinal, de acuerdo a Martín Delgado (2009), incluso en artículos anteriores a la irrupción de la IA, se permitía el uso de algoritmos para ejercer potestades que permiten un cierto margen de discrecionalidad técnica. Boix (2020) tiene una visión todavía más avanzada y considera que la completa potencialidad de la IA se logrará en el ejercicio de las potestades discrecionales. Huergo (2020) también aboga por esta interpretación.

En primer lugar, es posible que la IA sirva para dictar una actuación administrativa automatizada. La Ley 40/2015 sí que regula el régimen de la actuación administrativa automatizada, que es aquella actuación realizada íntegramente por medios electrónicos en el marco de un procedimiento y en la que no hay intervención de empleado público. Véase que el precepto es muy amplio porque se refiere a “medios electrónicos” y no a IA o a algoritmos. En este sentido, conviene precisar que no toda actuación administrativa automatizada implica el uso de IA; del mismo modo que pueden emplearse sistemas de IA sin que ello dé lugar a una actuación administrativa automatizada según el tenor literal del artículo. La actuación administrativa automatizada regulada en la Ley 40/2015 suele operar en el ámbito de las potestades regladas, que son aquellas en las que la decisión administrativa se obtiene mediante la aplicación de los criterios previamente establecidos en la norma. Es decir, las potestades regladas se caracterizan por la predeterminación de los presupuestos de hecho y la aplicación automática de la consecuencia jurídica. Por este motivo, en estos supuestos, el principio de explicabilidad resulta más fácilmente aplicable, en la medida en que se conocen los presupuestos de hecho que han dado lugar a la consecuencia jurídica aplicada por la IA, y también, la revisión de su adecuación o no al ordenamiento jurídico.

En segundo lugar, existen las predicciones algorítmicas. Son aquellas que permiten a la Administración tomar una decisión más informada —por ejemplo, al basarse en un mapa de delincuencia basado en un algoritmo—.

Según Huergo (2020), su función sería similar a la de un informe en el seno de un procedimiento administrativo. Las predicciones algorítmicas tienen lugar en el ejercicio de las potestades discrecionales, en las que la norma no predetermina todas las condiciones y consecuencias de su ejercicio, sino solo algunas, por lo que la Administración goza de mayor margen de discrecionalidad.

Las principales diferencias serían, en primer lugar, que en la actuación administrativa automatizada puede no haber intervención humana; (al menos, no de forma directa), en las predicciones algorítmicas, sí; la actuación administrativa automatizada está prevista en la Ley 40/2015, aunque sea deficientemente, las predicciones algorítmicas, no; la actuación administrativa automatizada sirve para el ejercicio de potestades regladas, las predicciones algorítmicas, para el ejercicio de las potestades discrecionales.

6.2. La motivación de los actos administrativos dictados mediante IA

El hecho de que el Reglamento de IA no regule pormenorizadamente las exigencias de la IA en el seno de la Administración Pública en esta materia no implica que la Administración no tenga esas obligaciones reforzadas por imposición de la normativa interna, en particular, por la Constitución Española, la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas y la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público.

El Tribunal Constitucional considera que la motivación es un riguroso requisito del acto restrictivo de derechos, de los que se separen del criterio seguido en actuaciones precedentes, los que se dicten en ejercicio de potestades discrecionales, etc. (Sentencias de la Sala Primera del Tribunal Constitucional 26/1982, de 17 de julio; 8/1992, de 2 de enero; 52/1995, de 23 de febrero, entre otras).



La motivación, además, está directamente con los principios del Estado de Derecho (artículo 1.1 de la Constitución) y con el ejercicio de las potestades por parte de la Administración (artículo 103 de la Constitución y 3.1 de la Ley 40/2015) y ha sido definida por el Tribunal Supremo como un “derecho subjetivo del interesado no sólo en el ámbito sancionador sino en todos los sectores de la actuación administrativa” (Sentencia de la Sala de lo Contencioso-Administrativo del Tribunal Supremo 8073/2002, de 3 de diciembre).

Esto es, la motivación de los actos es una garantía formal de los mismos que resulta plenamente aplicable a los procedimientos en los que se emplea IA, aunque con dos matices: por un lado, las máquinas no pueden emitir declaraciones de voluntad, propiamente dichas, ni formular declaraciones de voluntad, juicio, conocimiento o deseo (Alamillo Domingo; Urios Aparisi, 2011); por otro lado, hay que tener en cuenta que la Ley 39/2015 ni la Ley 40/2015 reconocen qué elementos del procedimiento administrativo tramitado mediante IA deben motivarse.

La motivación en el caso de las actuaciones administrativas automatizadas o las predicciones algorítmicas supone explicar el proceso lógico que conduce a la adopción de la decisión. La motivación es la explicación transparente —y sucinta— del proceso lógico que conduce a la adopción del acto. (artículo 35 de la Ley 39/2015).

Urios y Alamillo han considerado que la exigencia de motivación supone en este caso:

La necesidad de codificar y poder reconstruir, para cada caso singular, las reglas lógicas – ya hemos visto que en una aproximación híbrida, con la aplicación de un método integrado por la lógica de predicados de primer orden, por las lógicas modales, deóntica y refutable aplicables, y por la lógica descriptiva en cuanto a la representación del dominio de conocimiento jurídico – que han sido aplicadas en el acto administrativo automático singular (2011).

El Consejo de Estado italiano, equiparable a la Sala Tercera del Tribunal Supremo español, fue pionero al definir también la exigencia de motivación de los actos administrativos automatizados. En la Sentencia de 13 de diciembre de 2019 considera que la aplicación de los algoritmos requiere “el pleno conocimiento del módulo utilizado y de los criterios aplicados”. En esta resolución, indica los extremos a los que se extiende la motivación: los autores del acto; 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.

Todo ello permitirá verificar que los criterios del procedimiento cumplen con las prescripciones de la Ley y el administrado pueda cuestionar el resultado alcanzado.

En la Sentencia de 8 de abril de 2019 (sentencia n.º 2270), el Consejo de Estado italiano también se enfrentó a la denuncia de la infracción del principio de motivación con ocasión de

un concurso de provisión de puestos de trabajo de personal docente. Los recurrentes consideraron que el algoritmo no había tenido en cuenta las preferencias incorporadas en la solicitud, ni en cuanto al nivel educativo (educación primaria o educación secundaria) ni tampoco la zona geográfica de destino. El recurso fue desestimado en primera instancia, pero el recurso de apelación presentado por los recurrentes ha sido estimado por el Consejo de Estado italiano. Además, este órgano ha establecido las reglas de aplicación de las decisiones automatizadas. En primer lugar, “la regla técnica que gobierna cualquier algoritmo no deja de ser en cualquier caso una regla administrativa general”.

Las consecuencias de esta regla son, por un lado, que las decisiones automatizadas se rigen por los principios generales de la actividad administrativa, que incluyen, además de la motivación, la publicidad, la transparencia, la razonabilidad y la proporcionalidad. Por otro lado, la Administración ha de velar por los intereses en presencia, y debe probar, actualizar y perfeccionar el algoritmo, especialmente en el caso del aprendizaje profundo. En este último caso, el algoritmo puede empezar a tomar decisiones que no estaban previstas por el programador inicialmente, por lo que la Administración debe ser especialmente cautelosa al revisar los resultados que arroja la máquina. Finalmente, desde el punto de vista del control judicial, el algoritmo se considera como un acto administrativo informático y el órgano judicial ha de poder evaluar la corrección del proceso automatizado (De la Sierra Morón, 2021).

Consideramos que la motivación debe incluir también la acreditación de que el algoritmo empleado ha superado las correspondientes auditorías que verifican su adecuado funcionamiento y, especialmente, la evaluación de impacto sobre los derechos fundamentales contenida en el Reglamento de IA.

Además, no podemos obviar la reciente Sentencia del Tribunal Supremo (n.º rec. 7878/2024), en la que condena a la Administración a proporcionar a la Fundación Ciudadana Civio el acceso al código fuente



de la aplicación informática BOSCO, desarrollada por el Ministerio para la Transición Ecológica para que las empresas comercializadoras de energía eléctrica puedan comprobar si los solicitantes del bono social cumplen con los requisitos para poder ser considerados consumidores vulnerables. Sin ánimo de ser exhaustivos, ya que el análisis de la Sentencia excedería del objeto de este artículo, sí que exige que el código fuente forme parte de la motivación del acto: “en el caso de actuaciones administrativas automatizadas, el acceso al código fuente posibilita la comprobación de la conformidad del sistema algorítmico con las previsiones normativas que debe aplicar, pues la motivación —parametrizada— de la decisión administrativa que tiene lugar reside en el diseño de los parámetros que determinan ese código fuente. Observación que adquiere singular importancia ante el hecho de que las exigencias de motivación de los actos administrativos sean predicables también de las actuaciones administrativas automatizadas, en aplicación del artículo 35 de la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas”.

Es decir, esta Sentencia, que es pionera en el ámbito de la transparencia algorítmica, debe ser puesta en relación con el deber de explicabilidad y, de hecho, el propio Tribunal Supremo afirma que, “en este nuevo contexto digital democrático se impone a los Poderes públicos la obligación, entre otras, de explicar de forma comprensible el funcionamiento de los algoritmos que se emplean en la toma de decisiones que afectan a los ciudadanos para permitirles conocer, fiscalizar y participar en la gestión pública”. Es decir, debe traducir las instrucciones de funcionamiento interno del sistema en una justificación comprensible. En este fallo, se integra la explicabilidad en el deber de motivar, ya que la motivación debe integrar un razonamiento accesible y comprensible de la decisión algorítmica.

Por otro lado, la motivación puede insertarse en el propio acto administrativo o en los informes incorporados en el procedimiento administrativo. El artículo 88.6 de la Ley 40/2015 establece que “la aceptación de informes o dictámenes servirá de motivación a la resolución cuando se incorporen al texto de la misma”. Esta motivación por remisión ha sido aceptada por la jurisprudencia, entre otras, en la Sentencia de la Sala de lo Contencioso-Administrativo del Tribunal Supremo 1446/2012, de 5 de marzo, en la que el Alto Tribunal admite la motivación *in aliunde*.

Es decir, el acto administrativo de que se trate no tiene que incorporar obligatoriamente todos los detalles sobre la programación del sistema, pero sí que debería proporcionar la información necesaria para que el administrado pueda localizar estos datos, o bien deberían constar en el expediente o en otras fuentes de información que se encuentren al alcance del interesado (Sentencia del Tribunal Supremo 1806/2017, de 23 de noviembre de 2017).

La falta de motivación del acto determinaría su invalidez. De hecho, la jurisprudencia ha establecido que los actos administrativos carentes de motivación generan indefensión (Sentencia de la Sala de lo Contencioso-Administrativo del Tribunal Supremo 2872/2019, de 19 de septiembre). Además, cuando el acto afecta a un derecho fundamental, como el derecho de reunión o los procedimientos sancionadores, la motivación tiene alcance constitucional y determinaría un vicio de nulidad radical del acto (Sentencia del Tribunal Constitucional 26/1982, de 17 de julio).

No obstante, lo anterior, esta cuestión debe ser objeto de un tratamiento normativo específico que explicita qué elementos deben incorporarse a la motivación, de qué forma pueden motivarse los actos administrativos dictados con IA y qué consecuencias tiene la falta de motivación en este caso, entre otros.

A la motivación del acto administrativo debe añadirse una nueva variable: ¿qué pasaría si el decisor público ignorase la predicción algorítmica? Un ejemplo sería el supuesto en el que el algoritmo valorase todas las ofertas en un procedimiento de contratación pública y determinase que una empresa tiene un elevado riesgo de quiebra, obteniendo el peor resultado a juicio de la máquina. Si, a pesar de este resultado, se decidiera adjudicar el contrato a esta mercantil, ¿debería indicarse que este acto es contrario al criterio del algoritmo y motivarse con especial intensidad? A estos efectos, el artículo 35 de la Ley 39/2015 no exige la motivación de los actos que se separen del criterio del algoritmo, pero sí cuando el acto se separa de actuaciones precedentes o del dictamen de consejos consultivos (artículo 35.1 c) de la Ley 39/2015) (Capdeferro Villagrasa, 2020).

A nuestro juicio, en este supuesto, si se trata de una decisión discrecional donde el algoritmo es un mero instrumento de apoyo, la motivación siempre resultaría exigible (artículo 35.1 i) de la Ley 39/2015). Si se trata de una actuación administrativa automatizada por encontrarnos ante una potestad reglada, en ese caso, habría que realizar una ponderación del interés público



concurrente y motivar especialmente las razones por las que se aparta del criterio del algoritmo.

En definitiva, la motivación supone que la Administración debe justificar de manera clara y comprensible las decisiones adoptadas mediante IA, de manera que se asegure de que los ciudadanos y ciudadanas puedan entender los criterios utilizados y, en su caso, las resoluciones que resulten lesivas para sus derechos. En este sentido, la explicabilidad es un imperativo del Reglamento de IA pero también del ordenamiento jurídico interno. Además, la motivación debe incluir información suficiente sobre el funcionamiento del sistema de IA utilizado, las reglas lógicas aplicadas, los datos empleados y los factores que influyeron en la decisión. No se trata únicamente de una formalidad, sino de una verdadera exigencia para alcanzar la transparencia en el uso de tecnologías automatizadas dentro de la Administración.

7. Conclusiones

Primero, el Reglamento de IA de la Unión Europea representa un avance significativo en la regulación del uso de sistemas de IA en todos los sectores, incluido el sector público. Esta norma introduce un enfoque basado en el riesgo, ya que clasifica los sistemas de IA de alto riesgo o bajo riesgo en función de su potencial impacto sobre los derechos fundamentales, la seguridad y la privacidad de los ciudadanos. La aprobación del Reglamento marca un paso importante hacia la armonización de las normativas en Europa, puesto que establece estándares comunes que fomentan la innovación tecnológica, mientras se protegen los derechos fundamentales de los ciudadanos.

Segundo, el Reglamento de IA incluye a la Administración Pública y autoridades públicas dentro de su ámbito de aplicación, equiparando en gran medida el uso de la IA en el sector público al de los usos en el sector privado. De hecho, introduce muy pocas obligaciones en materia de IA en la Administración, por lo que se considera que el Reglamento no aborda de manera específica los riesgos que el uso de la IA en el sector público puede representar. La explicación de esta deficiencia es la vigencia del principio de autonomía institucional, que supone que los Estados Miembros tienen libertad para configurar sus procedimientos y órganos dentro de los márgenes que marca la legislación europea.

Tercero, entre los usos de la IA en el sector público, destacan algunos sectores clave como la justicia, el asilo y la identificación biométrica, que han sido clasificados como de alto riesgo. En estos casos, los sistemas deben someterse a una evaluación de impacto en los derechos fundamentales para evitar comprometer derechos como la privacidad, la igualdad y la no discriminación. A pesar de su importancia, la regulación de esta evaluación presenta deficiencias, ya que la flexibilidad en

su aplicación y la falta de supervisión efectiva pueden dar lugar a evaluaciones superficiales o incompletas.

Cuarto, a pesar de la escasa regulación del Reglamento de IA sobre las obligaciones específicas de la Administración Pública, la normativa interna española impone una serie de garantías que se aplican a la Administración por el mero hecho de serlo, en particular, el deber de motivación de los actos administrativos.

Quinto, la explicabilidad en los sistemas de IA constituye un principio fundamental para garantizar la transparencia y la rendición de cuentas en la toma de decisiones automatizadas. Si bien la transparencia permite conocer que un individuo ha sido objeto de una decisión algorítmica, la explicabilidad va un paso más allá al exigir que esa decisión sea comprensible y significativa para los afectados. Sin embargo, el Reglamento de IA presenta ciertas lagunas en cuanto a los elementos específicos que debe incluir la explicación y su relación con otras normativas, como la protección de datos. Por ello, es fundamental desarrollar directrices más detalladas que permitan implementar este principio de manera efectiva y uniforme en todos los ámbitos de aplicación de la IA.

Sexto, la aplicación del Principio de explicabilidad en el ámbito administrativo está directamente vinculado con la motivación de los actos, ya que la automatización de las decisiones no exime a la Administración de su deber de justificar sus actos administrativos. En este sentido, la doctrina y la jurisprudencia han enfatizado que la motivación debe incluir una reconstrucción lógica del proceso algorítmico, con indicación de los datos utilizados, las reglas aplicadas y los criterios que han llevado a la decisión final. Recientemente, también se ha incluido dentro de la motivación de las actuaciones administrativas automatizadas, la necesidad de aportar el código fuente. Esta exigencia cobra especial relevancia en el caso de decisiones discrecionales, en las que el algoritmo actúa como un instrumento de apoyo y la Administración debe justificar si se aparta de sus recomendaciones y en qué términos.



Séptimo, el marco normativo actual aún presenta lagunas en lo que respecta a la regulación específica de la motivación en decisiones dictadas mediante sistemas de IA. La Ley 39/2015 y la Ley 40/2015 no establecen con precisión los elementos que deben incluirse en la motivación de actos dictados con IA, lo que genera incertidumbre sobre la suficiencia de la argumentación en estos casos. De ahí la necesidad de una regulación específica que delimite con claridad qué información debe proporcionarse al ciudadano, de qué manera puede integrarse en el procedimiento administrativo y cómo debe evaluarse en sede de recursos, si los hubiere.

Referencias bibliográficas

Alamillo Domingo, I. & Urios Aparisi, F. (2011). *L'actuació administrativa automatitzada en l'àmbit de les administracions públiques*. Escola d'Administració Pública de Catalunya.

<https://eapc.bibliotecadigital.gencat.cat/bitstream/handle/20.500.14227/23/Alamillo-actuacio-administrativa-cat.pdf?sequence=1&isAllowed=y>

Alcolea Azcárraga, C. (2022). La responsabilidad patrimonial de la Administración y el uso de algoritmos. *Revista General de Derecho Administrativo*, (59), p.135-188.

Ballesteros Moffa, L. (2011). *Las dimensiones aplicativas del procedimiento administrativo europeo*. Parlamento Europeo. [https://www.europarl.europa.eu/RegData/etudes/note/join/2011/432764/IPOL-JURI_NT\(2011\)432764_ES.pdf](https://www.europarl.europa.eu/RegData/etudes/note/join/2011/432764/IPOL-JURI_NT(2011)432764_ES.pdf)

Barrio Andrés, M. (Dir.). (2024). *Comentarios al Reglamento Europeo de Inteligencia Artificial*. LA LEY.

Capdeferro Villagrasa, O. (2020). La inteligencia artificial del sector público: desarrollo y regulación de la actuación administrativa en la cuarta revolución industrial. *Revista de Internet, Derecho y Política*, (30), 1-14. <http://dx.doi.org/10.7238/idp.v0i30.3219>

Cerrillo I Martínez, A., Galindo Caldés, R. & Velasco Rico, C. (2020). *La personalización de los servicios públicos. La contribución de la inteligencia artificial y los datos masivos*, XXXIII Concurso del CLAD sobre Reforma del Estado y Modernización de la Administración Pública.

Chang Chuyes, G. (2025). Motivación e inteligencia artificial. En P. Valcárcel Fernández & L. Hernández González (Coords.), *El Derecho Administrativo en la era de la inteligencia artificial* (pp. 351–360). Instituto Nacional de Administración Pública (INAP).

Chaves, J. (2016). *Principio de buena administración: nuevo paradigma de control de la discrecionalidad*. De La Justicia.

<https://delajusticia.com/2016/06/07/principio-de-buena-administracion-nuevo-paradigma-de-control-de-la-discrecionalidad/>

Cotino Hueso, L. (2021). Hacia la transparencia 4.0: el uso de la inteligencia artificial y *big data* para la lucha contra el fraude y la corrupción y las (muchas) exigencias constitucionales. En C. Ramio (Coord.) *Repensando la Administración Pública: Administración digital e innovación pública* (169-197). Instituto Nacional de Administración Pública. <https://www.uv.es/cotino/publicaciones/INAPCOTINOPUBLICADO.pdf>

De la Quadra-Salcedo, T. (2024). Estudio preliminar. El Reglamento Europeo de Inteligencia Artificial y su regulación en condiciones de permanente adaptación a los riesgos y a la evolución de los modelos y sistemas de IA de uso general. En Barrio Andrés, M. (Dir.), *Comentarios al Reglamento Europeo de Inteligencia Artificial* (pp. 23-48). Tirant Lo Blanch.

<https://www.aranzadilaley.es/MK/PDF/Comentarios-al-reglamento-europeo-de-inteligencia-artificial/publication.pdf>

De la Sierra Morón, S. (2021). Control judicial de los algoritmos: robots, administración y estado de derecho. *ElDerecho*. <https://elderecho.com/control-judicial-de-los-algoritmos-robots-administracion-y-estado-de-derecho>

Díez Picazo, L. (2024). Il principio di autonomia istituzionale degli Stati membri dell'Unione europea. *Quaderni costituzionali*, 24(4), 865-867. DOI: 10.1439/19041

Díaz Senés, J. (2024). Panorámica internacional de la regulación de la inteligencia artificial: Reglamento Europeo de Inteligencia Artificial (AI ACT). *Revista CESCO de Derecho de Consumo*, (49), 117-130. <https://dialnet.unirioja.es/servlet/articulo?codigo=9462903>

Eguiluz Castañeira, J. (2023). Cinco aclaraciones necesarias sobre el Reglamento de IA (con un prólogo para tecnólogos y un epílogo para optimistas. *Diario La Ley*. https://diariolaley.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAEAFWNsQ6CMBRF_6ZzX1EThk7AHZC4kUJvzYulj7TVyN8rgyae8SbnHvbeDlf9gVoYjVZP5MKSrNGmIUNnlcRj7Dv7SB6BE7yKbkbsjhmvam9lyLwMOUs-bA57L8u4b7DBxQKFWeT-F5l-55VXIOrW7duTEOmTad-npxWQlwAAAA=WKE

González-Álvarez, J., Hermoso Santos, J. & Camacho Collados, M. (2020). *Policía predictiva en España. Aplicación y retos de futuro*. *Behavior & Law Journal*, 6(1), 26-41. <https://doi.org/10.47442/blj.v6.i1.75>

Grupo independiente de Expertos de Alto Nivel sobre Inteligencia Artificial de la Comisión Europea. (2019). *Directrices éticas para una inteligencia artificial fiable*. Comisión Europea. <https://digital-strategy.ec.europa.eu/es/library/ethics-guidelines-trustworthy-ai>

Guilayn Agustinoy, A. (2024). Artículo 86. Derecho a explicación de decisiones tomadas individualmente. En Barrio Andrés, M. (Dir.), *Comentarios al Reglamento Europeo de Inteligencia Artificial* (pp. 774-449). Tirant Lo Blanch.

Huergo Lora, A. (2020). Una aproximación a los algoritmos desde el Derecho Administrativo. En J. Huergo Lora & G. Díaz González (Eds.), *La regulación de los algoritmos* (pp. 23-87). Aranzadi.



Huici Sancho, L. (2010). Los gobiernos locales en la Unión Europea: del principio de autonomía institucional al principio de autonomía local. *Anuario del Gobierno Local*, 2010(1), 521-547. <https://repositorio.gobiernolocal.es/xmlui/handle/10873/746>

Laukyte, M. (2024). Reflexión sobre los derechos fundamentales en la nueva Ley de la Inteligencia Artificial. *Derechos y Libertades*, (51), 151-175. <https://doi.org/10.20318/dyl.2024.8586>

Ley Orgánica 3/2007 [Cortes Generales]. Ley para la igualdad efectiva de mujeres y hombres. 22 de marzo de 2007.

Ley 21/2013 [Cortes Generales]. Ley de Evaluación Ambiental. 9 de diciembre de 2013.

Ley 39/2015 [Cortes Generales]. Ley del Procedimiento Administrativo Común. 1 de octubre de 2015.

Ley 40/2015 [Cortes Generales]. Régimen Jurídico del Sector Público. 1 de octubre de 2015.

Ley 2/2025 [Cortes Generales]. Ley para el desarrollo e impulso de la inteligencia artificial en Galicia. 2 de abril de 2025.

López-Tarruella Martínez, A. (2024). Derecho a presentar una reclamación y derecho a una explicación: Vías de recurso para los particulares en el reglamento de inteligencia artificial. En L. Cotino Hueso & P. Simón Castellano (Dir.), *Tratado sobre el reglamento de inteligencia artificial de la Unión Europea* (pp. 893–918). Aranzadi.

Martín Delgado, I. (2009). Naturaleza, concepto y régimen jurídico de la actuación administrativa automatizada. *Revista de Administración Pública*, (180), 353-386. <https://www.cepc.gob.es/sites/default/files/2021-12/27614isaacmartindelgadordap180.pdf>

Oliver Cuello, R. (2021). Big data e inteligencia artificial en la Administración tributaria. *Revista de Internet, Derecho y Política*, (33), 1-13. <http://dx.doi.org/10.7238/idp.v0i33.381275>

Ponce Solé, J. (2019). Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico. *Revista General de Derecho Administrativo*, 50, 1-35. <https://laadministracionaldia.inap.es/noticia.asp?id=1509505>

Reglamento de inteligencia artificial de la Unión Europea 2024/1689 [Parlamento Europeo y Consejo de la Unión Europea]. Reglamento (UE) 2024/1689 Del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial). 13 de junio de 2024.

Sentencia del Tribunal de Justicia de la Unión Europea, *Leandro Tontodonati contra Comisión de las Comunidades Europeas* (15 de diciembre de 1971). Tribunal de Justicia de la Unión Europea.

Sentencia de la Sala Primera del Tribunal Constitucional 26/1981, de 17 de julio, (1981, 17 de julio). Tribunal Constitucional.

Sentencia de la Sala Primera del Tribunal Constitucional 8/1992 (1992, 2 de enero). Tribunal Constitucional.

Sentencia de la Sala Tercera del Tribunal Supremo 4049/2021 (2022, 27 de abril). Tribunal Supremo.

Sentencia del Tribunal Constitucional 26/1982 (1981, 17 de julio). Tribunal Constitucional.

Sentencia del Tribunal Constitucional 8/1992 (1992, 16 de enero). Tribunal Constitucional.

Sentencia del Tribunal Constitucional 52/1995 (1995, 23 de febrero). Tribunal Constitucional.

Sentencia del Tribunal Supremo 8073/2002 (2002, 3 de diciembre). Tribunal Supremo.

Sentencia de la Sala de lo Contencioso-Administrativo del Tribunal Supremo 1446/2012 (2012, 5 de marzo). Tribunal Supremo.

Sentencia de la Sala de lo Contencioso-Administrativo del Tribunal Supremo 1119/2025 (2025, 11 de septiembre). Tribunal Supremo.

Valero Julián, J. (2019). Les garanties jurídiques de la intel·ligència artificial en l'activitat administrativa des de la perspectiva de la bona administració. *Revista Catalana de Dret Públic*, (58), 82-96. DOI: <https://doi.org/10.2436/rcdp.i58.2019.3307> 