



Working Papers  
2025, 1(3):25-41, © DIGITAPIA  
Universitat Oberta de Catalunya  
<https://blogs.uoc.edu/digitapia/>

DIGITAPIA

## La transparencia algorítmica

### *The algorithmic transparency*

Agustí Cerrillo i Martínez<sup>1</sup>

Catedrático de Derecho Administrativo, Universitat Oberta de Catalunya  
[acerrillo@uoc.edu](mailto:acerrillo@uoc.edu)

Clara I. Velasco Rico<sup>2</sup>

Profesora Agregada Serra Húnter de Derecho Administrativo, Universitat Pompeu Fabra  
[clara.velasco@upf.edu](mailto:clara.velasco@upf.edu)

#### Resumen

El avance de las tecnologías de inteligencia artificial (IA) ha generado un profundo debate jurídico y ético en torno al control democrático de los sistemas automatizados de toma de decisiones. En este debate, los conceptos de transparencia y explicabilidad algorítmica han emergido como pilares fundamentales para garantizar la responsabilidad, la justicia y la protección de los derechos fundamentales. Aunque ambos términos suelen utilizarse de forma intercambiable en el lenguaje común e incluso en documentos técnicos, jurídicamente poseen alcances distintos y responden a exigencias diferenciadas.

**Palabras clave:** IA, Transparencia, Explicabilidad, Caja negra, Derechos fundamentales.

#### Abstract

The advancement of artificial intelligence (AI) technologies has sparked a profound legal and ethical debate regarding the democratic control of automated decision-making systems. Within this debate, the concepts of transparency and algorithmic explainability have emerged as key pillars to ensure accountability, fairness,

---

<sup>1</sup>Doctor en Derecho y licenciado en Ciencias Políticas y de la Administración. Catedrático de Derecho administrativo en la Universitat Oberta de Catalunya. Es miembro de la Comisión Jurídica Asesora de la Generalitat de Cataluña. Ha sido Director de sus Estudios de Derecho y Ciencia Política y Síndic de Greuges de la UOC. En la actualidad es co-investigador principal del proyecto Transición digital de las Administraciones públicas e inteligencia artificial: claves, riesgos y propuestas de política regulatoria (DIGITAPIA) (TED2021-132191B-I00).

<sup>2</sup>Doctora en Derecho, licenciada en Derecho y licenciada en Ciencias Políticas y de la Administración. Profesora agregada Serra Húnter de Derecho administrativo en la Universidad Pompeu Fabra. Vicepresidenta de la Comisión de Garantía del Derecho de Acceso a la Información Pública de la Generalitat de Catalunya. Ha sido secretaria del Consejo de Redacción de la RCDP. En la actualidad es co-investigadora principal del proyecto Transición digital de las Administraciones públicas e inteligencia artificial: claves, riesgos y propuestas de política regulatoria (DIGITAPIA) (TED2021-132191B-I00).

and the protection of fundamental rights. Although these terms are often used interchangeably in common language and even in technical documents, from a legal perspective they have distinct scopes and respond to different normative requirements.).

**Keywords:** AI, Tranparency, Explainability, Black box, Fundamental rights.

## 1. Introducción

La inteligencia artificial se caracteriza por la dificultad para conocer e interpretar su funcionamiento y explicar sus resultados. Esta opacidad, que convierte los sistemas de IA en cajas negras (black box) (Pasquale, 2015), impide poder conocer los procesos seguidos por los sistemas de IA para obtener sus resultados y limita la capacidad de auditar, supervisar y cuestionar el diseño de los algoritmos y de las decisiones que estos adoptan. Además, la falta de transparencia plantea cuestiones sobre la rendición de cuentas de las decisiones adoptadas con el uso de la inteligencia artificial dado que los sistemas de IA pueden tomar decisiones sin que se pueda identificar o conocer las razones.

La opacidad algorítmica es un fenómeno complejo y multifactorial de la opacidad algorítmica (Burrell, 2016); (Cerrillo i Martínez, 2021); (C. Velasco Rico, 2019).

Uno de los principales factores de opacidad es de carácter técnico y proviene de la naturaleza de los algoritmos que, al funcionar como cajas negras, ocultan su diseño y sus procesos internos. La opacidad también traducirse en el desconocimiento de los datos utilizados para entrenar o poner en funcionamiento el sistema, así como de las reglas que determinan cómo se procesan o descartan estos datos.

Otro conjunto de factores de opacidad de los sistemas de IA puede ser jurídico. En efecto, la legislación vigente prevé diferentes límites al principio de transparencia con la finalidad de proteger distintos intereses públicos o privados como puede ser la propiedad intelectual o industrial o la protección de los datos personales que pueden impedir el conocimiento del diseño de los algoritmos. Ciertamente, esta apertura informativa no está exenta de tensiones, especialmente en lo que respecta a los límites derivados de la protección de la propiedad intelectual e industrial sobre los algoritmos. En este sentido, Azuaje Pirela y Finol González (2020) advierten que el principio de transparencia puede entrar en conflicto con los derechos exclusivos del desarrollador del software, por lo que proponen mecanismos de equilibrio institucional que permitan la supervisión sin comprometer la innovación tecnológica.

Por último, un último conjunto de factores tiene carácter organizativo. Las Administraciones públicas con frecuencia no proporcionan información puntual ni actualizada sobre los sistemas de IA que utilizan para llevar a cabo sus funciones por no disponer de la información (Velasco Rico, 2020). En ocasiones, no facilitan la información para tratar de evitar que las personas afectadas por las decisiones automatizadas puedan tratar de eludir sus efectos (De Laat, 2018).

Ante la incorporación creciente de los sistemas de inteligencia artificial en las Administraciones públicas, se hace cada vez más difícil garantizar la transparencia en su funcionamiento y el control sobre las motivaciones de su actuación y, en última instancia, poder rendir cuentas y prevenir la arbitrariedad.

Frente a la opacidad algorítmica, las Administraciones públicas deben garantizar el acceso a la información necesaria para permitir que la ciudadanía pueda evaluar si un sistema de IA actúa de forma ajustada al ordenamiento jurídico. No obstante, la implementación de mecanismos de transparencia se enfrenta a desafíos técnicos, normativos y organizativos (Cerrillo i Martínez, 2021).

A nivel técnico, los proveedores de sistemas de IA se encuentran con el dilema de mantener la eficiencia y precisión de sus modelos mientras intentan hacerlos comprensibles para los usuarios y reguladores. En términos normativos, los reguladores deben equilibrar la transparencia con la protección de la propiedad intelectual y los secretos comerciales de las empresas que desarrollan estos sistemas, así como con la ciberseguridad. Por lo que respecta los aspectos organizativos, las Administraciones públicas deben desarrollar una gobernanza de la IA y de datos que permita facilitar el acceso a la información sobre los sistemas de IA.

En las próximas páginas, llevaremos a cabo una revisión de la literatura publicada en los últimos años en relación con la transparencia de la inteligencia artificial. Posteriormente, examinaremos los principales mecanismos a disposición de las Administraciones públicas para promover la transparencia de los sistemas de IA. Para concluir, haremos una propuesta normativa con la que abordar la regulación de los mecanismos de transparencia de los sistemas de IA.

## 2. Estado de la cuestión y/o revisión sistemática de la literatura

En la última década se han publicado decenas de artículos científicos y estudios que abordan desde diferentes perspectivas no sólo la transparencia, pero también en lo que se refiere a la transparencia y la explicabilidad algorítmica<sup>3</sup>.

Una búsqueda en Google Scholar muestra el gran interés entre expertos y académicos sobre la transparencia algorítmica que se concreta en más de seiscientas publicaciones en inglés y quinientas en español.

Además, algunas de las publicaciones poseen un alto volumen de citas.

Entre las publicaciones en inglés sobre transparencia algorítmica, se puede observar que 2020 fue el año en que más publicaciones se acumularon (96). El menor número de publicaciones está entre 1984 (1) y 2015 (7). A partir de 2015, se observa un aumento considerable de publicaciones sobre transparencia algorítmica en inglés en Google Scholar.

Las diez publicaciones en inglés sobre la transparencia de los algoritmos que han acumulado un mayor número de citas pueden ser consultadas en la Tabla 1.

Tabla 1: Diez publicaciones sobre la transparencia de los algoritmos en inglés con mayor acúmulo de citas

<i>Autor</i>	<i>Título</i>	<i>Año</i>	<i>Citas</i>
B Goodman, S Flaxman	European Union regulations on algorithmic decision-making and a right to explanation”	2017	2571
M Ananny, K Crawford	Seeing without knowing	2018	1369
FK Došilović, M Brčić, N Hlupić	Explainable artificial intelligence	2018	1327
L Edwards, M Veale	Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for	2017	1099
MK Lee	Understanding perception of algorithmic decisions	2018	985
A Datta, S Sen, Y Zick	Algorithmic transparency via quantitative input influence	2016	981
N Diakopoulos	Algorithmic accountability	2015	872
ID Raji, A Smart, RN White, M Mitchell	Closing the AI accountability gap	2020	770
B Lepri, N Oliver, E Letouzé, A Pentland	Fair, transparent, and accountable algorithmic decision-making processes	2018	663
RF Kizilcec	How much information? Effects of transparency on trust in an algorithmic interface	2016	517

Tabla elaborada por los autores.

Por lo que se refiere a las publicaciones en castellano, hasta 2018 (5), las publicaciones fueron moderadas (6, en 2017) pero aumentaron gradualmente hasta llegar a 18 en 2019 y 39 en 2020. Destaca

especialmente el salto a 140 publicaciones en 2023. Respecto a las publicaciones sobre explicabilidad de los sistemas de IA, el resultado de la investigación en Google Scholar muestra también la publicación de

<sup>3</sup>Esta recopilación es un resultado del proyecto EAPC. A continuación, se recogen los resultados de la revisión sistemática de la literatura sobre transparencia y explicabilidad algorítmica en la base de datos Google Scholar. A estos efectos se han realizado cuatro búsquedas en Google Scholar en cuatro ocasiones diferentes utilizando las palabras clave .<sup>a</sup>algorithmic transparency”(transparencia algorítmica) y .<sup>a</sup>algorithmic explainability”(explicabilidad algorítmica) en inglés. Después, el mismo criterio con las mismas palabras claves en castellano, es decir, transparencia algorítmica y explicabilidad algorítmica.

seiscientos trabajos, teniendo la publicación más citada 7517 menciones y las de menor impacto con sólo una mención de 21 textos. En 2022 fue el año con más publicaciones (153). Por otro lado, los datos muestran que la explicabilidad algorítmica ha generado un creciente interés en lengua castellana, especialmente en los últimos años. En 2019 se publicaron 10 trabajos,

en 2020 24 y, 47, en 2021.

Las diez publicaciones en inglés sobre la explicabilidad de los algoritmos que han acumulado un mayor número de citas pueden ser consultadas en la Tabla 2.

Tabla 2: Diez publicaciones sobre la explicabilidad de los algoritmos en inglés con mayor acúmulo de citas

<i>Autor</i>	<i>Título</i>	<i>Año</i>	<i>Citas</i>
AB Arrieta, N Díaz Rodríguez, J Del Ser, A Bennetot	Explainable Artificial Intelligence (XAI)	2020	7517
A Adadi, M Berrada	Peeking inside the black-box	2018	5400
J Burrell	How the machine ‘thinks’	2016	2767
LH Gilpin, D Bau, BZ Yuan, A Bajwa	Explaining explanations	2018	2718
BD Mittelstadt, P Allo, M Taddeo	The ethics of algorithms	2016	2354
E Tjoa, C Guan	A survey on explainable artificial intelligence (xai)	2020	1696
DV Carvalho, EM Pereira, JS Cardoso	Machine learning interpretability	2019	1626
A Holzinger, G Langs, H Denk	Causability and explainability of artificial intelligence in medicine	2019	1361
FK Došilović, M Brčić, N Hlupić	Explainable artificial intelligence	2018	1327
I Zhang, X Chen	Explainable recommendation	2020	1003

Tabla elaborada por los autores.

### 3. Experiencias prácticas relacionadas con el objeto

Las Administraciones públicas han ido implementando diferentes mecanismos para promover la transparencia de los sistemas de IA que utilizan.

Tal vez el que ha tenido mayor reconocimiento son los registros de sistemas de IA con el fin de canalizar la transparencia algorítmica en el ámbito público. Otro mecanismo utilizado por algunas Administraciones públicas es la publicación del código fuente de los sistemas de IA. Por último, las Administraciones públicas están utilizando en relación con los sistemas de IA los mecanismos de transparencia previstos en la legislación de administración digital y de transparencia (difusión proactiva en sedes electrónicas y portales de transparencia, acceso a la información pública, etc.).

#### A) *Los registros de algoritmos*

Una de las principales herramientas que se están

impulsando para canalizar la transparencia algorítmica son los registros de algoritmos. Los registros de algoritmos son bases de datos que recogen información sobre los sistemas algorítmicos utilizados por las Administraciones públicas.

De acuerdo con (Boet & Pérez-Batlle, 2023), la puesta en marcha de estas herramientas se enfrenta a distintos retos como la estandarización de un esquema de datos que defina el tipo de información a inventariar en los registros de algoritmos o la selección de los sistemas de IA que deben incorporarse en el registro y que deben ser públicos.

Distintas ciudades, como Ámsterdam o Helsinki, han sido pioneras en la implementación de estas herramientas. Estas y otras iniciativas se basan en el modelo impulsado desde la Coalición de Ciudades por los Derechos Digitales que parte de la *Internet Rights and Principles Coalition (IRPC)*, inspiradas a su vez por la *United Nations Internet Governance Forum*

(UNIGF)<sup>4</sup>. La Declaración de la Coalición de Ciudades por los Derechos Digitales adoptada por la Coalición de Ciudades por los Derechos Digitales (*Cities Coalition for Digital Rights*) identifica los principios que deben regir la actividad de las ciudades con el fin de garantizar los derechos digitales entre los que destacan los de transparencia y de rendición de cuentas. De acuerdo con la declaración, el principio de transparencia implica que “Todo el mundo debería tener acceso a una información comprensible y precisa sobre los sistemas tecnológicos. afectan a su vida, así como a la capacidad de poner en duda y cambiar los sistemas injustos, parciales o discriminatorios.”

Por otra parte, el *Digital Forum Lab de Eurocities* ha promovido el Estándar de transparencia algorítmica que es un conjunto de categorías comunes de información útiles para informar sobre el funcionamiento y la finalidad de los algoritmos utilizados en las administraciones locales. El uso de estas categorías compartidas permite la comparación de distintos algoritmos tanto dentro de una misma ciudad como entre distintos entes locales. Asimismo, el uso de la norma permite obtener información útil en los procesos de toma de decisiones en la ciudad<sup>5</sup>. El *Digital Forum Lab* también ha fomentado abrir los algoritmos y detallar su funcionamiento en la toma de decisiones. El Estándar de transparencia algorítmica de Eurocities es un modelo clave para la creación de registros de algoritmos, ofreciendo un marco robusto para promover la transparencia, seguridad y confianza en la implementación de sistemas de inteligencia artificial (IA). Según el Estándar de transparencia algorítmica, la creación de un registro de transparencia algorítmica se desarrolla en siete etapas. La versión 0.3.1 (beta) del esquema de registro de transparencia algorítmica incluye las categorías de información que deben completarse para que los algoritmos funcionen de forma transparente, no discriminatoria y velando por los derechos de la ciudadanía. Estas categorías cubren distintas fases del ciclo de vida de un algoritmo (por ejemplo, las de prueba, implementación, funcionamiento y revisión periódica). La información que integra el registro de algoritmos se clasifica en 6 categorías que se utilizan para estructurar y difundir la información. En total, el registro público de algoritmos incluye hasta 49 campos de información sobre algoritmos.

Entre las iniciativas que se han impulsado, pode-

mos traer a colación las siguientes:

a) Dutch Algorithmic Transparency Standard<sup>6</sup>. El Estándar de Metadatos para el Registro de Algoritmos de los Países Bajos tiene como objetivo apoyar la implementación del Reglamento de Inteligencia Artificial y, en particular, la previsión sobre la creación de una base de datos de sistemas de IA de alto riesgo. Este estándar también persigue dar respuesta a la Moción del parlamento neerlandés (35510-16) aprobada por la Cámara de Representantes de los Países Bajos que insta al gobierno a crear un registro de algoritmos que describa los algoritmos utiliza el gobierno, con qué objetivos y en qué bases de datos se apoyan con la finalidad de permitir al público realizar el público concordancia con los principios del estado de derecho. Sin embargo, esta moción todavía no se ha convertido en ley.

b) UK’s Algorithmic Transparency Recording Standard Hub<sup>7</sup>. El Estándar de Registro de Transparencia Algorítmica del Reino Unido pretende ayudar a las organizaciones del sector público a ofrecer información clara sobre las herramientas algorítmicas que utilizan y las razones de su uso. En su concepción, la transparencia algorítmica implica ser abiertos respecto a cómo estas herramientas se utilizan para apoyar las decisiones de los poderes públicos, proporcionando información completa, accesible y comprensible. El estándar incluye un modelo de registro para inventariar el uso de herramientas algorítmicas, las guías para cumplimentar estos registros y un repositorio de registros públicos sobre transparencia algorítmica. Este estándar forma parte de la Estrategia Nacional de Datos del gobierno británico y se alinea con la Estrategia Nacional de Inteligencia Artificial. Desde 2024, se exige a todos los departamentos del gobierno central que utilicen este estándar, con el propósito de extenderlo a otros organismos públicos.

c) The Scottish AI Register<sup>8</sup>. El registro

<sup>4</sup>Desde otras latitudes también se ha promovido la creación de registros de algoritmos, entre otros, a partir del impulso del Centro Latinoamericano de Administración para el Desarrollo -CLAD a través de la *Carta Iberoamericana de Inteligencia Artificial en la Administración Pública*.

<sup>5</sup>Accesible en: [algorithmregister.org](https://algorithmregister.org) (última consulta: marzo de 2025).

<sup>6</sup>Accesible en: [standaard.algoritmeregister.org](https://standaard.algoritmeregister.org) (última consulta: marzo de 2025).

<sup>7</sup>Accesible en: [gov.uk](https://gov.uk) (última consulta: marzo de 2025).

<sup>8</sup>Accesible en: [The Scottish AI Register](https://www.scottishairegister.org) (última consulta: marzo de 2025).

de IA escocés es una herramienta para conocer los sistemas de IA que está usando o desarrollando el gobierno escocés y su sector público con el fin de favorecer una mejor comprensión de cómo funciona la IA. El registro se deriva de lo dispuesto en la Estrategia de IA de Escocia. El registro facilita una descripción general de los sistemas así como una información detallada sobre los procesos empleados y cómo los datos contribuyen a mejorar la toma de decisiones y los servicios públicos. El sistema está diseñado para mostrar la evolución de los sistemas de IA concretos, desde su concepción y desarrollo hasta su puesta en marcha, y en su caso, su desmantelamiento<sup>9</sup>.

d) El Registro de sistemas de inteligencia artificial y otros algoritmos de la Generalidad de Cataluña. El Gobierno de la Generalidad ha aprobado la creación de un registro de algoritmos de la Generalidad de Cataluña, un proyecto que actualmente está en fase de desarrollo<sup>10</sup>. Asimismo, la Administración de la Generalitat ha impulsado el registro de algoritmos del Observatorio de la IA en la Salud nace en el marco del Programa Salud/IA que pretende establecer un entorno propicio para la innovación en el ámbito de la salud, impulsando el desarrollo y la implementación la ciudadanía<sup>11</sup>. En el Observatorio de IA en Salud se identifican, clasifican y analizan las herramientas de Inteligencia Artificial que se emplean en el ámbito de la salud y quedan registradas. Hasta la fecha se han incluido en el registro 170 herramientas. El acceso al registro no es accesible público, pero se ofrece información agregada sobre su contenido y una panorámica, a través de indicadores, sobre los tipos de herramientas que se han hecho constar en el mismo. Por último, también resulta de interés la iniciativa impulsada por el Consorcio de la Administración Abierta de Cataluña que publica unas fichas de algoritmos de IA con un contenido comprensible y su publicación en la sede electrónica<sup>12</sup>.

e) Los registros locales de algoritmos. Amsterdam, Barcelona, Bruselas, Eindhoven, Mannheim, Rotterdam y Sofía son las primeras municipalidades que se han adherido al modelo de gobernanza de transparencia algorítmica del Algorithmic Transparency Standard desde la red Eurocities y que han impulsado diferentes iniciativas que se encuentran en la actualidad en diferentes estadios de aplicación. En particular, en el caso de Barcelona, la Medida de Gobierno de la Estrategia municipal de algoritmos y datos para el impulso ético de la inteligencia artificial adoptada por el Ayuntamiento de Barcelona en abril de 2021 previó la creación de un registro municipal de algoritmos con el fin de que toda la ciudadanía pueda consultar los datos que se han utilizado para la base de IA<sup>13</sup>. No obstante, aún no se ha puesto en funcionamiento.

Mención aparte puede hacerse a la Ordenanza de administración digital del Ayuntamiento de Tarragona que regula el uso de la inteligencia artificial y, en particular, prevé la creación del Registro de sistemas de inteligencia artificial del Ayuntamiento de Tarragona que será impulsado por la Comisión de Administración Digital para dar a conocer los sistemas utilizados por el municipio<sup>14</sup>. Asimismo, la Ordenanza-tipo de Transparencia y Administración Digital elaborada por el Consorcio Localret establece la creación de un registro municipal de sistemas de inteligencia artificial<sup>15</sup>.

f) Estándares de transparencia algorítmica impulsados por la sociedad civil. La organización AI, Algorithmic, Automation Incidents and Controversies (AIAAIC) es una iniciativa independiente que busca promover la transparencia y la apertura en los sistemas de IA, algoritmos y automatización cuyo objetivo es hacer que la IA sea más comprensibles y accesibles mediante la recopilación y difusión de datos útiles para la academia y la sociedad civil, la educación de ciudadanos y políticos,

<sup>9</sup>Actualmente, el registro escocés tiene contabilizados 10 casos de uso de sistemas de IA.

<sup>10</sup>Acuerdo de Gobierno GOV/45/2024

<sup>11</sup>Resolución SLT/954/2023 de 19 de marzo de 2023.

<sup>12</sup>Acceso a las fichas aquí: [aoc.cat](http://aoc.cat) (última consulta, 13.11.2024).

<sup>13</sup>[Ajuntament de Barcelona](http://Ajuntament de Barcelona)

<sup>14</sup>Boletín Oficial de la Provincia de Tarragona de 30 de septiembre de 2024.

<sup>15</sup>Accesible en: [localret.cat](http://localret.cat) (última consulta: marzo de 2025).

considerando el impacto del uso de la IA y arquitecturas algorítmicas también para gobiernos y empresas.

El Canadian Tracking Automated Government (TAG) Register de Canadá es una iniciativa diseñada para monitorizar el uso de sistemas automatizados por parte del gobierno canadiense. Este registro tiene como objetivo documentar y analizar cómo se utilizan estas tecnologías en la toma de decisiones gubernamentales, centrándose en la transparencia y posibles impactos desiguales. Se basa en información obtenida a través de solicitudes de acceso a datos públicos, preguntas en el Parlamento y búsquedas en medios de comunicación y contratos gubernamentales.

El Repositorio de Algoritmos Públicos recopila y difunde información sobre los sistemas de soporte o toma de decisiones automatizadas desarrollados y/o implementados por instituciones públicas en Chile<sup>16</sup>. Este repositorio está impulsado por GobLab una iniciativa de la Escuela de Gobierno de la Universidad Adolfo Ibáñez.

En el ámbito de Cataluña, la Mesa de entidades del Tercer Sector Social lleva tiempo impulsando el proyecto *Radar de algoritmos de IA y procesos de decisión automatizada para el acceso a los derechos sociales en Cataluña*. En este estudio se recogen experiencias internacionales e identifican 12 sistemas de IA actualmente implementados o en proceso de desarrollo por las Administraciones públicas catalanas.

#### B) *La difusión del código fuente de los sistemas de IA*

Otro mecanismo para promover la transparencia algorítmica es la difusión del código fuente de los algoritmos<sup>17</sup>. En esta dirección, algunas Administraciones públicas están publicando el código de las aplicaciones que utilizan a través de repertorios como Github<sup>18</sup>. De todos modos, no se puede desconocer que en general la difusión no afecta a sistemas de inteligencia artificial, sino a otros programas informáticos.

En cualquier caso, el acceso a los algoritmos es necesario, pero no suficiente, para garantizar su transparencia. En efecto, el acceso al código fuente del algoritmo no lo hace necesariamente más transparente,

particularmente, cuando estamos ante complejos modelos de aprendizaje automático o si los destinatarios de la información son personas sin un conocimiento adecuado<sup>19</sup>.

Por ello, más allá de las iniciativas que puedan impulsarse para difundir el código fuente de los sistemas de IA será necesario garantizar la comprensibilidad de su funcionamiento, por ejemplo, facilitando una descripción del diseño del algoritmo en lenguaje natural. También puede ser útil utilizar lenguajes y procedimientos de programación que faciliten la transparencia y la auditoria de los sistemas de IA.

Además, no podemos desconocer los obstáculos que pueden poner las Administraciones públicas para facilitar el acceso al código fuente de las aplicaciones que utilizan. Para ello, podemos traer a colación el conocido asunto Bosco promovido por la Fundación Civio que hace unos años solicitó poder acceder al código fuente y otra información relevante sobre el funcionamiento de BOSCO, una aplicación desarrollada por el gobierno español que determina quién puede acceder al bono social eléctrico, una ayuda destinada a los hogares vulnerables para reducir la factura de la luz. En respuesta a la solicitud formulada por Civio, el gobierno únicamente proporcionó parte de la información, alegando que el código fuente estaba protegido por derechos de propiedad intelectual. Ante esta negativa, presentó una reclamación ante el Consejo de Transparencia y Buen Gobierno y posteriormente, Civio presentó un recurso contencioso-administrativo contra la resolución del Consejo que le desestimó la reclamación, para exigir la publicación del código que fue desestimado en 2022 por el Juzgado Central de lo Contencioso Administrativo número 8 argumentando que la divulgación del código fuente de BOSCO podría suponer un riesgo para la seguridad pública y la defensa nacional. Posteriormente, Civio presentó en mayo de 2024 un recurso ante la Audiencia Nacional que dictó una nueva sentencia desestimando el recurso y manteniendo la negativa a facilitar el código fuente de BOSCO. Ante esta situación, en junio del mismo año, Civio presentó un recurso de casación ante el Tribunal Supremo<sup>20</sup>.

#### C) *Los mecanismos de transparencia de la legislación de administración digital y de transparencia*

La legislación vigente prevé diversos mecanismos

<sup>16</sup>[algoritmospublicos.cl](https://algoritmospublicos.cl)

<sup>17</sup>Alamillo observa que “la liberación del código fuente de las aplicaciones coadyuva de forma extraordinariamente potente a generar confianza en el correcto funcionamiento de dichas aplicaciones por parte de los ciudadanos” (Alamillo Domingo, 2017).

<sup>18</sup>Por ejemplo, [Github](https://github.com) (última consulta: abril de 2025).

<sup>19</sup>Véase Brauneis & Goodman (2017).

<sup>20</sup>Véase al respecto [civio.es](https://civio.es) (última consulta: abril de 2025). Asimismo, (Cotino Hueso, 2024; Ponce Solé, 2024).

para canalizar la difusión de información pública y, en particular, información sobre las aplicaciones que utilizan. Por un lado, en la normativa se prevé la difusión proactiva de información sobre sistemas de IA<sup>21</sup>.

En esta dirección, el Real Decreto 203/2021, de 30 de marzo, por el que se aprueba el Reglamento de actuación y funcionamiento del sector público por medios electrónicos prevé que la sede electrónica debe publicar la relación actualizada de las actuaciones administrativas automatizadas vinculadas a los servicios, procedimientos y trámites y, en particular, la descripción de su diseño y funcionamiento, los mecanismos de rendición de cuentas y transparencia, así como los datos utilizados en su configuración y aprendizaje<sup>22</sup>.

También, la Ley 1/2022, de 13 de abril, de la Generalitat, de Transparencia y Buen Gobierno de la Comunitat Valenciana. En particular, prevé que las Administraciones públicas deben 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 forma comprensible de su diseño y funcionamiento, el nivel de riesgo que implican y el punto de contacto al que poder dirigirse a cada caso, de acuerdo con los principios de transparencia y explicabilidad.

En cambio, la Ley 19/2013, de 9 de diciembre, de Transparencia, Acceso a la Información Pública y Buen Gobierno, no incluye disposiciones específicas para abordar la cuestión de la transparencia de los

algoritmos o sistemas de IA<sup>23</sup>.

Por otro lado, la información sobre los sistemas de IA puede ser conocida en ejercicio del derecho de acceso a la información pública<sup>24</sup>.

Al respecto, debemos tener en cuenta que la información relativa a los sistemas de IA que utilizan las Administraciones públicas es información pública. En esta dirección, podemos traer a colación las pioneras resoluciones de la Comissió de Garantia del Dret d'Accés a la Informació Pública en relación una solicitud de acceso al algoritmo matemático que condiciona el orden de posiciones en los tribunales correctores de las PAU. En particular, la Resolución de 21 de septiembre de 2016, de estimación de las Reclamaciones 123/2016 y 124/2016 considera que el algoritmo puede ser expresado en palabras (lenguaje natural) o, para mayor precisión (para evitar las ambigüedades del lenguaje natural), en lenguaje matemático o, donde resulta necesario utilizar la potencia de cálculo de una computadora, en lenguaje informático y que constituye información pública<sup>25</sup>. En la misma dirección, la resolución del Consejo de Transparencia y Buen Gobierno de 20 de mayo de 2021, 058/2021 dictada en relación con la solicitud de acceso al algoritmo utilizado para el cálculo de las pensiones por la Tesorería General de la Seguridad Social en la que se afirma que “un algoritmo como el que es objeto de la presente reclamación, que el solicitante describe como aquél ‘con el que el programa de cálculo de la pensión realiza los porcentajes de cálculo correspondientes a los años completos de servicio efectivo...’, entra sin lugar a dudas dentro del ámbito del concepto de información

---

<sup>21</sup>En esta dirección, los artículos 311-3-1 y 311-3-2 Código de relaciones entre la ciudadanía y la Administración de Francia prevé que las decisiones adoptadas a través de algoritmos informen de manera explícita a la persona interesada, indicando la finalidad perseguida y el derecho a conocer las reglas que definen el tratamiento y las principales características de su aplicación y cómo se ha de ejercer el recurso ante la *Commission d'Accès aux Documents Administratifs*.

<sup>22</sup>Art.11.1.i) RAFME.

<sup>23</sup>Véase Martín Delgado (2023, p. 164).

<sup>24</sup>La resolución 20144578 de 8 de enero de 2015 de la *Commission d'Accès aux Documents Administratifs* francesa (CADA) sobre la solicitud de acceso al código fuente del software que simula el cálculo del IRPF desarrollado por la Dirección General de Tributos, consideró que “un código fuente es un programa informático que contiene las instrucciones que debe ejecutar un micro-procesador, mientras que los archivos informáticos que constituyen el programa buscado en este caso, producidos por la administración como parte de su misión de servicio público, eran documentos administrativos”. Posteriormente, en la resolución 20161990, de 23 de junio de 2016, sobre el algoritmo desarrollado por el Ministerio de Educación Nacional conocido como admisión post-bac (APB) también se afirma que un algoritmo es un documento administrativo a los efectos del régimen jurídico del acceso a la información

Asimismo, el Tribunal administrativo regional de Lazio (Italia) (sección tercera) en cuya sentencia de 14 de febrero de 2017 reconocía el derecho de acceso al algoritmo del programa de gestión de un procedimiento administrativo, en particular en el caso en cuestión para la transferencia interprovincial de los docentes, en la medida en que se puede considerar como un acto administrativo informático..

<sup>25</sup>La Resolución 200/2017, de 21 de junio, la GAIP refuerza estos argumentos afirmando que el código fuente de un programa informático utilizado por la Administración en la designación de los miembros de tribunales evaluadores constituye información pública.

<sup>26</sup>La resolución también indica que “Mientras no se instauren otros mecanismos que permitan alcanzar los fines señalados con garantías equivalentes –como podrían ser, por ejemplo, auditorías independientes u órganos de supervisión-, el único recurso eficaz a tales efectos es el acceso al algoritmo propiamente dicho, a su código, para su fiscalización tanto por quienes se puedan sentir perjudicados por sus resultados como por la ciudadanía en general en aras de la observancia de principios éticos y de justicia”.

pública consagrado en la LTAIBG”<sup>26</sup>.

#### D) *La apertura de datos*

Por último, también debemos referirnos a la transparencia de los datos utilizados para el entrenamiento o la evaluación de los sistemas de IA<sup>27</sup>.

En efecto, el acceso a los datos es necesario para conocer el diseño y el funcionamiento de los algoritmos y para poder supervisar cómo operan y, en última instancia, para poder controlar las decisiones adoptadas automáticamente a través de algoritmos. En esta dirección, en los últimos años se han impulsado distintas iniciativas tanto para facilitar el acceso y su apertura, es decir, su puesta a disposición de los datos sin límites técnicos y/o jurídicos que impidan o dificulten su reutilización.

En primer lugar, debemos traer a colación la Estrategia Europea de Datos en febrero de 2020<sup>28</sup>. En el marco de esta estrategia, desde la Unión Europea se han adoptado distintas normas que promueven la reutilización de datos en relación con los sistemas de IA y, en general, fomentar el uso de datos abiertos y estimular la innovación de los productos y servicio como la Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativa a los datos abiertos y la reutilización de la información del sector público (Directiva de datos abiertos); el Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos; y el Reglamento (UE) 2023/2854 del Parlamento Europeo y del Consejo, de 13 de diciembre de 2023, sobre normas armonizadas para un acceso justo a los datos y su utilización, y por el que se modifican el Reglamento (UE) 2017/2394 y la Directiva (UE) 2020/1828 (Reglamento de Datos)<sup>29</sup>.

Asimismo, debemos referirnos a la Ley 37/2007, de 16 de noviembre, de reutilización de la información del sector público (LRISP) que transpone al ordenamiento español la normativa europea y en consonancia con ella fomenta la difusión de datos en formatos abiertos. En particular, esta norma dispone que “las Administraciones y organismos del sector público facilitarán sus documentos en cualquier formato o lengua pre-

existente, pero también procurarán, siempre que ello sea posible y apropiado, proporcionarlos en formato abierto y legible por máquina conforme a lo previsto en el apartado anterior y conjuntamente con sus metadatos, con los niveles más elevados de precisión y desagregación. Tanto el formato como los metadatos, en la medida de lo posible, deben cumplir estándares y normas formales abiertas”<sup>30</sup>.

La reutilización y la apertura de los datos también se recoge en la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno que prevé entre los principios generales que guían la publicidad activa, la reutilización<sup>31</sup>.

La apertura de datos, y en general, el uso de los datos por las Administraciones públicas está estrechamente vinculado a la existencia de una gobernanza de datos que asegure su calidad y facilite la toma de decisiones sobre los datos y su reutilización<sup>32</sup>.

Asimismo, también a los principios y garantías previstas en el Reglamento General de Protección de datos en la medida en que con frecuencia los datos utilizados por los algoritmos de las Administraciones públicas tratan datos personales lo que exigirá tener en cuenta lo previsto en esta disposición.

A pesar de lo anterior y de los esfuerzos de muchas Administraciones públicas en difundir datos abiertos y reutilizables, no es habitual aún que los portales de datos abiertos difundan los datos utilizados por los algoritmos.

#### D) *Transparencia vs. explicabilidad algorítmica*

Como ya hemos apuntado, el despliegue progresivo de sistemas de inteligencia artificial (IA) en sectores tan diversos como la sanidad, la administración pública, el ámbito financiero o la justicia penal ha exigido al Derecho una respuesta normativa que equilibre el fomento de la innovación con la tutela efectiva de los derechos fundamentales. En este contexto, los conceptos de transparencia y explicabilidad algorítmica han adquirido un protagonismo creciente como principios estructurantes del nuevo marco regulatorio europeo.

<sup>27</sup> Como recuerda el Libro blanco sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza “Sin datos, el desarrollo de la IA y otras aplicaciones digitales resulta imposible” [COM (2020) 65 final].

<sup>28</sup> COM(2020) 66 final.

<sup>29</sup> Véase Cerrillo i Martínez (2023).

<sup>30</sup> Art.5.3 Ley 37/2007.

<sup>31</sup> Art. 5.4 Ley 19/2013. Este principio se incluye entre los principios técnicos que deben guiar el diseño y funcionamiento del portal de transparencia de la Administración General del Estado (art.11.c).

<sup>32</sup> Véase Cerrillo i Martínez & Casadesús de Mingo (2021).

Ambos términos, aunque emparentados, responden a lógicas jurídicas distintas. La transparencia se refiere a la disponibilidad de información sobre los componentes, procesos y objetivos de un sistema algorítmico, con la finalidad de permitir su supervisión por parte de entidades reguladoras, auditores o incluso por la ciudadanía, en función del contexto<sup>33</sup>. Por su parte, la explicabilidad algorítmica alude a la capacidad del sistema para ofrecer razones inteligibles y comprensibles sobre decisiones concretas, especialmente desde la perspectiva del sujeto afectado. Esta distinción no es meramente técnica, sino que tiene hondas implicaciones normativas y procedimentales en la garantía de los derechos individuales.

Desde el plano normativo, la transparencia encuentra anclaje en diversos instrumentos del Derecho europeo. El Reglamento General de Protección de Datos (RGPD) configura este principio como una exigencia fundamental del tratamiento de datos personales (artículo 5.1.a) y desarrolla su alcance en los artículos 13 y 14, relativos a los deberes de información previa al tratamiento. Además, el artículo 22 RGPD incorpora el derecho del interesado a no ser objeto de decisiones exclusivamente automatizadas, reconociendo en consecuencia su derecho a obtener información significativa sobre la lógica aplicada. A ello se suma ahora el nuevo Reglamento de Inteligencia Artificial (RIA) que refuerza las obligaciones de transparencia para los sistemas de alto riesgo (artículo 13), imponiendo a los proveedores la obligación de diseñar sistemas que permitan comprender su funcionamiento, sus limitaciones, los datos utilizados y los riesgos potenciales. La norma también prevé obligaciones de transparencia en modelos fundacionales, como los sistemas de IA generativa, exigiendo documentación técnica accesible y garantías de no manipulación informativa (artículo 52 RIA).

En cuanto a su contenido, la transparencia jurídica exige que los sistemas algorítmicos sean documentados adecuadamente, que sus finalidades sean explícitas, que los datos de entrenamiento sean trazables y que los responsables del sistema sean claramente identificables<sup>34</sup>. Esta transparencia está orientada fundamentalmente al control institucional, por lo que su eficacia depende, en gran medida, de que existan mecanismos reales de auditoría y de supervisión algorítmica —como las pruebas de impacto algorítmico, contempladas también en el RIA— que garanticen que los sistemas cumplen con los principios de legalidad,

proporcionalidad y no discriminación<sup>35</sup>.

Ahora bien, la transparencia no asegura por sí sola la comprensión efectiva de las decisiones automatizadas. En este punto irrumpe el concepto de explicabilidad algorítmica, cuyo foco es la persona afectada. Se trata de dotar al sujeto de capacidad para entender por qué una decisión concreta fue adoptada por el sistema, qué factores fueron determinantes y de qué manera podría impugnarla o revisarla. Esta dimensión conecta directamente con los principios de justicia procedimental, con el derecho a la tutela judicial efectiva (art. 47 Carta de los Derechos Fundamentales de la UE) y con el derecho a ser informado (art. 8 Carta de los Derechos Fundamentales de la UE) conforme ha argumentado Ortiz de Zárate (2022), quien destaca que la explicabilidad debe traducirse en una herramienta procesal para garantizar la equidad decisional y no solo en un principio técnico o ético.

Desde el punto de vista técnico-jurídico, la explicabilidad puede alcanzarse mediante modelos intrínsecamente comprensibles (como árboles de decisión), o a través de herramientas de explicación post hoc que permitan interpretar modelos complejos, como las redes neuronales<sup>36</sup>. Métodos como LIME o SHAP se han desarrollado precisamente para ofrecer explicaciones localizadas y adaptadas al usuario<sup>37</sup>. La finalidad no es solo permitir la comprensión técnica, sino garantizar que el sujeto pueda ejercer sus derechos de forma efectiva, en especial cuando los sistemas se utilizan en contextos donde la decisión automatizada puede implicar consecuencias jurídicas relevantes (rechazo de un crédito, negativa en una solicitud administrativa, inclusión en un sistema de puntuación de riesgo, etc.).

Los retos que plantea la explicabilidad algorítmica en los sistemas de inteligencia artificial más complejos —especialmente aquellos basados en modelos de aprendizaje automático opacos o de tipo caja negra, como redes neuronales profundas, bosques aleatorios o máquinas de vectores soporte— han impulsado el desarrollo de herramientas metodológicas capaces de proporcionar explicaciones post hoc. Estas se generan después de que el modelo haya tomado una decisión, sin requerir acceso directo al código fuente ni a la arquitectura interna del sistema. Entre las técnicas más influyentes y extendidas actualmente destacan LIME (*Local Interpretable Model-Agnostic Explanations*) y SHAP (*SHapley Additive exPlanations*), ambos diseñados para generar explicaciones localizadas,

<sup>33</sup>Véase Edwards & Veale (2018)

<sup>34</sup>Véase Cotino Hueso (2023)

<sup>35</sup>Véase Gobierno de Cataluña (2022).

<sup>36</sup>Véase Wachter, Mittelstadt & Russell (2017).

<sup>37</sup>Véase Kirat et al. (2022).

comprensibles por humanos y adaptadas a contextos de alta relevancia jurídica o social. El método LIME, desarrollado por Ribeiro, Singh y Guestrin (2016), parte de la premisa de que el comportamiento global del modelo puede ser opaco o incomprensible, pero que es posible generar una aproximación interpretable en torno a una predicción concreta. Para ello, LIME realiza pequeñas perturbaciones sobre los datos de entrada —modificando valores de ciertas variables— y observa cómo varían las predicciones resultantes. Con base en dichas variaciones, construye un modelo local y lineal, fácilmente interpretable, que permite al usuario conocer qué características específicas del input han influido más en una determinada decisión algorítmica. Esta aproximación es especialmente útil en entornos en los que se exige transparencia funcional de la decisión, como en procesos administrativos o judiciales con intervención algorítmica<sup>38</sup>.

Por su parte, SHAP, desarrollado por Lundberg y Lee (2017), se basa en una formulación matemática inspirada en los valores de Shapley de la teoría de juegos cooperativos. SHAP asigna a cada variable de entrada una contribución cuantificada a la predicción final, calculando el efecto marginal de cada característica en todas las posibles combinaciones de variables. Esta metodología garantiza una descomposición precisa y equitativa de la predicción, cumpliendo propiedades deseables como la simetría, la aditividad y la consistencia, lo que la convierte en una técnica especialmente robusta desde la perspectiva de la responsabilidad algorítmica. Estas propiedades tienen implicaciones jurídicas, pues aseguran que la explicación no esté sesgada ni arbitrariamente construida, aportando seguridad jurídica y confiabilidad en escenarios de toma de decisiones automatizadas con relevancia legal<sup>39</sup>.

La utilidad de estas técnicas no se limita al plano técnico o computacional. Como ha destacado Ortiz de Zárate Alcarazo (2022), la explicabilidad debe entenderse como una garantía procesal de los derechos fundamentales, no como una simple herramienta de auditoría. En efecto, la posibilidad de entender el porqué de una decisión automatizada es un prerequisite indispensable para que el sujeto afectado pueda ejercer adecuadamente su derecho a la impugnación, al recurso y, en última instancia, a una tutela judicial efectiva (art. 47 de la Carta de Derechos Fundamentales de la Unión Europea). Por ello, los métodos como LIME y SHAP no solo cumplen una función de tra-

ducción técnica, sino también de instrumentalización normativa del derecho a la explicación, consagrado implícitamente en el artículo 22.3 del RGPD.

En este sentido, Kirat, Tambou, Do y Tsoukiàs (2022) subrayan que el uso de herramientas como SHAP permite avanzar hacia sistemas de decisión automatizada “explicablemente justos”, especialmente si se integran en procesos de gobernanza algorítmica basados en derechos. La trazabilidad de las decisiones, la posibilidad de reconstruir los factores causales y la descomposición clara de las contribuciones individuales de cada variable son elementos que resultan claves para dotar de efectividad al principio de no discriminación, a la transparencia sustantiva y al control democrático de la inteligencia artificial.

No obstante, el uso de estas metodologías no está exento de críticas y limitaciones. Wachter, Mittelstadt y Russell (2017) advierten que, aunque estas técnicas son prometedoras, sus explicaciones son aproximaciones heurísticas que no equivalen a una justificación jurídica completa. Además, en contextos de alta no linealidad o cuando se emplean datos con alto grado de correlación, los resultados de LIME o SHAP pueden ser inestables o difíciles de interpretar correctamente. Estas observaciones deben tenerse en cuenta si se pretende utilizar dichas explicaciones como prueba o como motivación formal en procesos contenciosos o administrativos.

En suma, LIME y SHAP representan instrumentos fundamentales en la operacionalización práctica del principio de explicabilidad, particularmente en entornos donde el uso de sistemas de IA comporta riesgos concretos para los derechos fundamentales. Su integración en el diseño de sistemas de inteligencia artificial, especialmente aquellos regulados como de alto riesgo por el Reglamento de Inteligencia Artificial de la Unión Europea (RIA) —arts. 13 y 52—, constituye un paso necesario para cumplir las exigencias de transparencia, rendición de cuentas y acceso efectivo al razonamiento algorítmico.

Ambas dimensiones —transparencia y explicabilidad— son necesarias para articular un verdadero régimen de rendición de cuentas algorítmica (algorithmic accountability). No obstante, su naturaleza y destinatarios difieren, como se sintetiza en la Tabla 3:

Tabla elaborada por los autores.

---

<sup>38</sup>Véase Ribeiro et al. (2016).

<sup>39</sup>Véase Lundberg & Lee (2017).

Tabla 3: Cuadro comparativo sobre transparencia y explicabilidad

<i>Elemento comparado</i>	<i>Transparencia</i>	<i>Explicabilidad</i>
Finalidad principal	Permitir auditoría y control institucional	Permitir comprensión e impugnación por parte del afectado
Destinatarios	Reguladores, desarrolladores, peritos	Ciudadanos, usuarios, personas afectadas
Contenido	Información técnica y estructural	Justificación clara, inteligible y adaptada a humanos
Naturaleza	Técnica, documental, institucional	Cognitiva, comunicativa, procedimental
Mecanismos típicos	Códigos abiertos, documentación, informes de impacto, registros	Herramientas de interpretación, visualización, lenguaje llano

En definitiva, la transparencia puede considerarse un requisito estructural *ex ante*, mientras que la explicabilidad opera como una garantía funcional *ex post*, orientada a la defensa efectiva del derecho a comprender y a ser tratado con justicia en contextos de automatización.

En definitiva, tanto la transparencia como la explicabilidad deben entenderse como instrumentos complementarios en la arquitectura del nuevo Derecho de la inteligencia artificial. El RIA, en su redacción definitiva, da un paso relevante al consolidar ambas exigencias como elementos normativos vinculantes, lo que representa un avance significativo hacia un modelo de IA compatible con los valores democráticos, la responsabilidad institucional y la protección sustantiva de los derechos fundamentales en la era digital<sup>40</sup>.

## 4. Propuesta regulatoria

La propuesta normativa que se formula persigue garantizar la transparencia de los sistemas de inteligencia artificial que utilizan las Administraciones públicas. En particular, la transparencia debe garantizarse en relación con el procedimiento de decisión sobre el uso de la inteligencia artificial; sobre el diseño de los sistemas de IA y sobre las decisiones automatizadas dictadas utilizando sistemas de IA.

En primer lugar, en relación con la transparencia de la decisión sobre el uso de la inteligencia artificial, cabe recordar que la Ley 30/1992, de 26 de noviembre, sobre el régimen jurídico de las administraciones públicas y del procedimiento administrativo común incluía la obligación de divulgar públicamente las características de los programas y aplicaciones informáticas que utilizaran, siempre que hubieran sido aprobadas previamente. Actualmente, el régimen jurídico vigente del procedimiento administrativo común no contempla esta obligación con igual claridad y alcance<sup>41</sup>.

En segundo lugar, por lo que se refiere a la información sobre el diseño de los sistemas de IA, sería necesario consolidar los principios básicos de la difusión de la información tal y como recoge reglamentariamente el artículo 11.1.i) RAFME y garantizar no solo la publicación de la relación actualizada de las actuaciones administrativas automatizadas vinculadas a los servicios, procedimientos y trámites sino que la información que se difunda sea suficiente tanto para poder supervisar, auditar y controlar como para que la ciudadanía pueda conocer el alcance de la automatización de la Administración pública. En esta dirección, debe garantizarse que la información sea suficiente y comprensible.

En tercer lugar, sobre las decisiones automatiza-

<sup>40</sup>Véase Cotino Hueso (2023).

<sup>41</sup>La doctrina no solo ha identificado esta circunstancia, sino que también ha propuesto recuperar esta obligación. Véase al respecto Cerrillo i Martínez (2019); Valero Torrijos (2019). Ver asimismo, Martín Delgado (2023) quien considera que parte de esta información —al menos el algoritmo en sí— debería formar parte del expediente administrativo. Esta obligación podría canalizarse a través de la tramitación de un procedimiento de automatización Vaquer Caballería (2025).

da dictadas utilizando sistemas de IA, la legislación básica debería garantizar que estas incorporen una referencia a la automatización tal y como, por ejemplo, está previsto en la legislación francesa que exige informar de manera explícita a la persona interesada de la finalidad perseguida y sobre las reglas que definen el tratamiento y las principales características de su aplicación y cómo se ha de ejercer y el recurso específico que se pueda interponer respecto a la decisión automatizada.

Adición el apartado j) en el artículo 53 Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas

1. Además del resto de derechos previstos en esta Ley, los interesados en un procedimiento administrativo, tienen los siguientes derechos:

[...]

j) Derecho a recibir información comprensible y en lenguaje plano sobre los sistemas de IA utilizados en la actuación administrativa automatizada, su alcance y funcionamiento.

Artículo 88. Contenido

4. En el caso de las actuaciones administrativas automatizadas, la resolución debe informar de que la actuación ha sido realizada de forma automatizada, de las consecuencias que tiene asociadas y del espacio de la sede electrónica a que se refiere el artículo 40.3 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público.

Adición de los siguientes artículos en el Capítulo V del Título Preliminar de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público

Artículo 40. Sede electrónica

3. La sede electrónica ha de difundir la relación actualizada de actuaciones administrativas automatizadas con una descripción de los parámetros, reglas e instrucciones que definen su diseño y funcionamiento, los mecanismos de rendición de cuentas y de transparencia, los datos

utilizados en su configuración y aprendizaje, los recursos que procedan contra la actuación y los informes de evaluación de impacto previstos en el RIA.

Artículo 41.bis. Principio de transparencia de los sistemas de IA

1. El uso de sistemas de IA debe basarse en el principio de transparencia.
2. Las Administraciones públicas deben garantizar la disponibilidad, difusión, acceso y reutilización de la información relativa a los sistemas de IA que utilizan.
3. La información sobre los sistemas de IA debe ser completa, accesible, comprensible, actualizada y de calidad.
4. Las Administraciones públicas deben garantizar que el diseño de los sistemas de IA permita la trazabilidad de su funcionamiento y la explicabilidad de sus actuaciones automatizadas.

Artículo 41.tri. Procedimiento para la aprobación de la actuación administrativa automatizada

1. Las Administraciones públicas deben determinar las especificaciones técnicas, las características básicas del modelo, los datos utilizados para el entrenamiento y la supervisión, los mecanismos de supervisión, auditoría y control.
2. La incorporación de sistemas de inteligencia artificial debe tener en cuenta la información disponible sobre el diseño y funcionamiento de los algoritmos y, en su caso, los datos empleados en la configuración y aprendizaje de los sistemas de inteligencia artificial y la capacidad de explicar de forma comprensible la forma en que se ha tomado la decisión.

Artículo quater. El registro público de sistemas de inteligencia artificial

1. Los sistemas de inteligencia artificial que utilicen las Administraciones públicas deben estar inscritos en un registro público de sistemas de inteligencia artificial
2. El registro público de sistemas de inteligencia artificial debe contener como mínimo la descripción de los parámetros, reglas e instrucciones que definen el diseño del sistema y funcionamiento, los mecanismos

de rendición de cuentas y de transparencia, los datos utilizados en su configuración y aprendizaje, en su caso, y los informes de evaluación de impacto previstos en el RIA.

3. El registro es de acceso público a través de la sede electrónica.

4. Las Administraciones públicas pueden crear su propio registro o registrar los sistemas de IA en el registro de otra Administración pública previo el oportuno convenio.

## 5. Justificación

Tal y como se ha ido apuntando al examinar los distintos mecanismos a través de los que canalizar la transparencia de los sistemas de IA y de los datos utilizados en su diseño, entrenamiento y evaluación, tal y como sostuvimos en una ocasión anterior “(...) más allá del acceso al algoritmo, que ya ha sido concedido en algunos casos en nuestro país, al ser considerado ‘información pública’, y que, para muchos, no representa ninguna solución completa, es necesario establecer otros mecanismos para garantizar la transparencia de los sistemas algoritmos”<sup>42</sup>.

Con carácter previo a la propuesta normativa para regular los mecanismos a través de los que garantizar la transparencia de los sistemas de IA, a continuación, se examinan los fundamentos normativos en los que deba basarse.

### A) *Los fundamentos normativos de los mecanismos de transparencia algorítmica en la legislación europea*<sup>43</sup>

En primer lugar, la regulación de los mecanismos de transparencia de los sistemas de IA debe basarse en lo que dispone la legislación europea y, en particular, el Reglamento (UE) 2016/679 del Parlamento

Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos, RGPD) y el RIA.

El RGPD reconoce que el responsable del tratamiento debe facilitar a la persona interesada, entre otra información, la relativa a la existencia de decisiones automatizadas y, en particular, información significativa sobre la lógica aplicada, así como la importancia y las consecuencias previstas de dicho tratamiento para el interesado<sup>44</sup>. Para ello, deberá tomar las medidas oportunas para facilitar al interesado toda esta información<sup>45</sup>. El derecho a obtener esta información está estrechamente relacionado al derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado que produzca efectos jurídicos en la persona interesada o le afecte significativamente de modo similar<sup>46</sup>. A este respecto, Karminsky, en el contexto de los debates sobre la interpretación del artículo 22 RGPD, defiende que “los individuos tienen ‘derecho a ser informados’ sobre la toma de decisiones algorítmicas”<sup>47</sup>.

A su vez, el RIA establece como uno de sus pilares el principio de transparencia lo que se traduce en distintas obligaciones que se imponen a proveedores y responsables del despliegue en relación con los sistemas de IA que desarrollen y comercialicen<sup>48</sup>. Por lo general, cuanto más riesgo suponga el sistema para los bienes jurídicos protegidos, más estrictas serán las obligaciones de transparencia que deberán cumplir sus proveedores, comercializadores o responsables del despliegue cuando sea necesario. Las previsiones del RIA en materia de transparencia difieren y tienen distinto alcance respecto al Reglamento General de Protección de Datos, ya que implican a actores y destinatarios diferentes. En este sentido, el Reglamento de Inteligencia Artificial, que se centra específicamente en los sistemas de IA, define la transparencia como la información que los proveedores de sistemas de IA deben proporcionar a los usuarios, entendidos como las entidades que utilizan estos sistemas. En el considerando

<sup>42</sup>Véase C. I. Velasco Rico (2019). En la misma dirección, Edwards & Veale (2017) quienes proponen la implementación de sistemas de evaluación de algoritmos y certificación, tanto de los resultados obtenidos como de los profesionales que los diseñan o utilizan.

<sup>43</sup>Antes de entrar en las disposiciones específicas del Reglamento de IA de la UE sobre transparencia, cabe recordar que el Grupo Independiente de Expertos de Alto Nivel en IA, creado por la Comisión Europea, ya estableció en 2019, en las Directrices Éticas para una IA fiable, el principio de transparencia como uno de los siete principios éticos esenciales.

<sup>44</sup>Art.13.2.f) RGPD.

<sup>45</sup>Art.12. RGPD

<sup>46</sup>Art.22. RGPD.

<sup>47</sup>Véase Kaminski (2021).

<sup>48</sup>Véase Velasco Rico (2024). Tal y como se destaca en este trabajo en la versión en español del RIA, la transparencia se menciona hasta 41 veces, en diversos contextos, con alcances y objetivos diferenciados.

27, el Reglamento especifica que la “transparencia” implica que los sistemas de IA se desarrollen y utilicen de forma que permitan una trazabilidad y explicabilidad adecuadas, asegurando, además, que las personas sean conscientes de cuándo se comunican con un sistema de IA. También exige informar adecuadamente tanto a los responsables del despliegue como a las personas afectadas sobre las capacidades y limitaciones del sistema de IA, así como sobre los derechos que les asisten. Asimismo, el considerando 67 destaca la necesidad de transparencia en la gestión y gobernanza de datos para cumplir con la normativa de protección de datos de la UE, como el RGPD. El Reglamento también exige que esta información esté siempre actualizada, como se indica en el considerando 101, a fin de que cualquier cambio o mejora en el sistema de IA quede reflejado en la documentación. Así, el Reglamento de IA no sólo promueve la transparencia como valor, sino que la convierte en una obligación activa y constante para los proveedores, exigiéndoles un compromiso continuo en la gestión y actualización de la información disponible para las autoridades, y, de forma indirecta, también para la ciudadanía.

En particular, el principio de transparencia algorítmica se regula en el artículo 13 RIA y se materializa, entre otros instrumentos, en la creación de un Registro de sistemas algorítmicos. Según el artículo 49 RIA, cualquiera que quiera comercializar o implementar un sistema de IA de alto riesgo deberá registrarlo previamente en una base de datos de la UE, con toda la información. También deberá registrarse en otra base de datos cualquier sistema de IA considerado no de alto riesgo. Este registro será público, salvo en determinados casos sensibles.

Esta estructura de la transparencia algorítmica refleja la voluntad del legislador europeo de garantizar una supervisión rigurosa de los sistemas de IA en el ámbito público, combinando la protección de la seguridad en áreas sensibles con el fomento de la confianza ciudadana en el uso de la inteligencia artificial en la gestión pública.

B) *Los fundamentos normativos de los mecanismos de transparencia algorítmica en la legislación española*

Desde el punto de vista jurídico, el fundamento de los mecanismos para la transparencia de los sistemas de IA que utilizan las Administraciones públicas se encuentra en distintas normas que regulan el régimen jurídico de las Administraciones públicas y, en particular, la transparencia pública tal y como se ha expuesto anteriormente.

En particular, la transparencia se configura como un principio fundamental en la actuación de la administración pública en el artículo 3 de la Ley de Régimen Jurídico del Sector Público (LRJSP). Este principio tiene su engarce en el artículo 105 de la Constitución Española y ha sido desarrollado a través de la Ley 19/2013, de Transparencia, Acceso a la Información Pública y Buen Gobierno, junto con las correspondientes normativas autonómicas.

Junto a estas normas, debe tenerse en cuenta que el artículo 23 de la Ley 15/2022, de 12 de julio, de igualdad de trato y no discriminación ha venido a reforzar las obligaciones de transparencia en el uso de algoritmos, aunque de forma poco precisa. Concretamente, esta ley incorpora tres mecanismos clave: primero, exige la inclusión de mecanismos que integren los principios de transparencia y rendición de cuentas; segundo, refuerza la necesidad de interpretación de las decisiones tomadas con algoritmos; y, finalmente, introduce sellos de calidad y evaluaciones de impacto.

Asimismo, también resulta de interés traer a colación, si bien con un alcance diverso al ser un documento de *soft law*, la Carta de Derechos Digitales, que en su apartado XVIII establece en su punto quinto que el ente público que desarrolle cualquier actividad en el entorno digital deberá identificar los órganos responsables de la misma. Esta disposición debe ponerse en relación con el artículo 41.2 de la LRJSP, que exige indicar “el órgano que debe ser considerado responsable a efectos de impugnación” en las actuaciones automatizadas. Por último, el sexto punto del epígrafe XVIII reconoce explícitamente el derecho a la transparencia, incluyendo la obligación de informar sobre el “carácter decisivo o no decisivo” del uso de inteligencia artificial en el ámbito.

## Referencias

- Alamillo Domingo, I. (2017). La regulación de la tecnología: La superación del modelo papel como elemento de transformación digital innovadora. En I. Martín Delgado (Ed.), *La reforma de la administración electrónica: Una oportunidad para la innovación desde el derecho* (pp. 79–130). Instituto Nacional de Administración Pública.
- Azuaje Pirela, M., y Finol González, D. (2020). Transparencia algorítmica y la propiedad intelectual e industrial: Tensiones y soluciones. *Revista La Propiedad Inmaterial*(30), 111–146.
- Boet, P., y Pérez-Batlle, M. (2023). *A propósito de los registros de algoritmos en las administraciones públicas locales*. Descargado de <https://cidai.eu/es/a-proposit-dels-registres-dalgoritmes-en-les-administracions-publicues-locales/>
- Brauneis, R., y Goodman, E. P. (2017). Algorithmic transparency for the smart city. *Yale Journal of Law & Technology*, 20, 104–176.
- Burrell, J. (2016). How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1). doi: 10.1177/2053951715622512
- Cerrillo i Martínez, A. (2019). Com obrir les caixes negres de les administracions públiques? transparència i rendició de comptes en l’ús dels algoritmes. *Revista Catalana de Dret Públic*, 58, 13–28. doi: 10.2436/rcdp.i58.2019.3277
- Cerrillo i Martínez, A. (2021). La transparencia de los algoritmos que utilizan las administraciones públicas. En R. Camp Batalla (Ed.), *Anuario de transparencia local* (Vol. 3, pp. 41–78). Fundación Democracia y Gobierno Local.
- Cerrillo i Martínez, A. (2023). Reutilización de la información del sector público e inteligencia artificial. En E. Gamero Casado (Ed.), *Inteligencia artificial y sector público: Retos, límites y medios* (pp. 466–501). Tirant lo Blanch.
- Cerrillo i Martínez, A., y Casadesús de Mingo, A. (2021). Data governance for public transparency. *El Profesional de la Información*, 30(4). doi: 10.3145/epi.2021.jul.02
- Cotino Hueso, L. (2024a). Caso bosco, a la tercera tampoco va la vencida. mal camino en el acceso a los algoritmos públicos. *Diario La Ley*, 84.
- Cotino Hueso, L. (2024b). *Tratado sobre el reglamento de inteligencia artificial de la unión europea*. Descargado de <https://www.uv.es/cotino/publicaciones>
- De Laat, P. B. (2018). Algorithmic decision-making based on machine learning from big data: Can transparency restore accountability? *Philosophy & Technology*, 31(4), 525–541.
- Edwards, L., y Veale, M. (2017). Slave to the algorithm: Why a right to an explanation is probably not the remedy you are looking for. *Duke Law & Technology Review*, 16, 18–84.
- Edwards, L., y Veale, M. (2018). *Enslaving the algorithm: From a “right to an explanation” to a “right to better decisions”?* <https://arxiv.org/abs/1803.07540>.
- Gobierno de Cataluña. (2022). *Transparencia algorítmica en el sector público*. Generalitat de Catalunya.
- Kaminski, M. E. (2021). The right to explanation, explained. En *Research handbook on information law and governance* (pp. 278–299). Edward Elgar Publishing.
- Kirat, T., Tambou, O., Do, V., y Tsoukiàs, A. (2022). Fairness and explainability in automatic decision-making systems: A challenge for computer science and law. *arXiv preprint*. Descargado de <https://arxiv.org/abs/2206.03226>
- Lundberg, S. M., y Lee, S.-I. (2017). A unified approach to interpreting model predictions. En *Advances in neural information processing systems* (Vol. 30).
- Martín Delgado, I. (2023). La aplicación del principio de transparencia a la actividad administrativa algorítmica. En E. Gamero Casado y F. L. Pérez Guerrero (Eds.), *Inteligencia artificial y sector público: Retos, límites y medios* (pp. 131–194). Tirant lo Blanch.
- Ortiz de Zárate Alcarazo, L. (2022). Explicabilidad (de la inteligencia artificial). *Eunomía. Revista en Cultura de la Legalidad*, 22, 328–344.
- Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.
- Ponce Solé, J. (2024). La transparencia algorítmica y los entes locales: Lecciones de las sentencias recaídas en el caso bosco, con especial atención a la de la audiencia nacional de 30 de abril de 2024. *Consultor de los Ayuntamientos y de los Juzgados: Revista técnica especializada en administración local y justicia municipal*, 3.
- Ribeiro, M. T., Singh, S., y Guestrin, C. (2016). “why should i trust you?”: Explaining the predictions of any classifier. En *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and*

*data mining.*

- Valero Torrijos, J. (2019). Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración. *Revista Catalana de Dret Públic*, 58, 82–96. doi: 10.2436/rcdp.i58.2019.3307
- Vaquer Caballería, M. (2025). El procedimiento de automatización y los efectos de la automatización sobre los procedimientos administrativos. En M. Vaquer Caballería y J. Pedraza Córdoba (Eds.), *La actuación administrativa automatizada: Sus claves jurídicas*. Tirant lo Blanch.
- Velasco Rico, C. (2019). La ciudad inteligente: Entre la transparencia y el control. *Revista General de Derecho Administrativo*(50).
- Velasco Rico, C. I. (2019). Vigilando al algoritmo: Propuestas organizativas para garantizar la transparencia. En B. Puentes Cociña y A. Quintiá Pastrana (Eds.), *El derecho ante la transformación digital: Oportunidades, riesgos y garantías* (pp. 73–90). Atelier.
- Velasco Rico, C. I. (2020). Personalización, proactividad e inteligencia artificial: ¿un nuevo paradigma para la prestación electrónica de servicios públicos? *Revista de Internet, Derecho y Política*(30).
- Velasco Rico, C. I. (2024). Implementación de sistemas algorítmicos en el sector público, preferencias ciudadanas y obligaciones de transparencia a la luz del reglamento de ia de la unión europea. En A. Cerrillo i Martínez, F. Di Lascio, I. Martín Delgado, y C. I. Velasco Rico (Eds.), *Inteligencia artificial y administraciones públicas: Una triple visión en clave comparada*. Iustel.
- Wachter, S., Mittelstadt, B., y Russell, C. (2017). *Counterfactual explanations without opening the black box: Automated decisions and the gdpr*. <https://arxiv.org/abs/1711.00399>.