



Qué concreta transparencia e información de algoritmos e inteligencia artificial es la debida

Lorenzo Cotino Hueso¹

*Catedrático de Derecho Constitucional. Universitat de Valencia. España
Investigador de la Universidad Católica de Colombia*

ORCID: [0000-0003-2661-0010](https://orcid.org/0000-0003-2661-0010)

RECIBIDO: 23 de diciembre de 2022

ACEPTADO: 24 de enero de 2023

El presente estudio es resultado de investigación del proyecto ELPIS: Análisis de datos para la toma de decisiones espaciales e inteligencia artificial para la administración, Grupo de Investigación en Derecho Público y TIC como investigador de la Universidad Católica de Colombia; proyecto MICINN Retos "Derechos y garantías frente a las decisiones automatizadas... (RTI2018-097172-B-C21); también "La regulación de la transformación digital ..." grupo de investigación de excelencia Generalitat Valenciana "Algoritmia law" (Prometeo/2021/009, 2021-24), proyecto "Derecho, Cambio Climático y Big Data", proyecto "Transición digital de las Administraciones públicas... (TED2021-132191B-I00); Inteligencia artificial y Derecho Administrativo... (UPO-1381574). Miembro de Valgrai.

¹ Lorenzo Cotino Hueso, www.cotino.net, cotino@uv.es. Catedrático Derecho Constitucional U. Valencia, Magistrado TSJ C. Valenciana hasta 2019, Vocal del Consejo de Transparencia C. Valenciana desde 2015. Director Regulación y Derechos de OdiselA, Coordinador Red www.derechotics.com, cofundador de la Red DAIA (Derecho Administrativo de la Inteligencia artificial). Doctor y licenciado en Derecho (UVEG), Máster en la especialidad de derechos fundamentales en Barcelona (ESADE), Licenciado y Diplomado de Estudios Avanzados de Ciencias políticas (UNED). Ha recibido premios de investigación nacionales, así como Premio Extraordinario de Doctorado. Profesor invitado en Konstanz (Alemania) desde 2004 y de diversas universidades colombianas. Autor de 16 libros, 160 artículos o capítulos y más 450 ponencias sobre gobierno abierto, inteligencia artificial, protección de datos, libertad en internet, e-Administración, temas sobre los que ha dirigido cinco proyectos nacionales de investigación.



RESUMEN: El estudio detalla qué datos, información o conocimiento debe facilitarse sobre los sistemas de IA. Se menciona si son obligaciones normativas presentes o futuras o de autoridades regulatorias. O si son las mejores prácticas que se detectan en Reino Unido, Países Bajos o Francia. Así, hay que facilitar información sobre la existencia del sistema, obligaciones de difusión activa y futuros registros de algoritmo públicos; información sobre quiénes son los proveedores de sistema de IA y quiénes los usuarios, los responsables de protección de datos, por qué existe el sistema y sus concretas finalidades. También información sobre los datos de entrenamiento o datos de entrada, así como la tipología y especificaciones técnicas del sistema IA, incluso su código. Se concreta la información en razón del artículo 22 RGPD. Igualmente, información sobre las garantías del sistema de IA, como estudios de impacto, mitigaciones de riesgos o supervisión humana.

PALABRAS CLAVE: Inteligencia artificial, transparencia de los algoritmos, explicabilidad, protección de datos, derechos.

CONTENIDOS: 1.- Transparencia y explicabilidad algorítmicas: el qué depende de para qué o para quién. 2.- La premisa: la "comunicación" de la existencia del sistema IA y notificaciones de su uso. Registros y bases de datos. 3.- Hacia la transparencia activa sobre la existencia de sistemas de IA públicos y su grado de incidencia en las decisiones públicas. 4.- Transparencia sobre quiénes son los proveedores de sistemas de IA y quiénes los utilizan y un contacto. 5.- Transparencia sobre por qué existe el sistema IA y explicación de que no había mejores alternativas. 6.- Transparencia sobre la finalidad prevista del sistema IA. 7.- Transparencia sobre los datos del sistema de IA (de entrenamiento, de entrada y los inferidos por el sistema). 8.- Transparencia sobre la tipología y especificaciones técnicas del sistema IA. 8.1.- *Transparencia del tipo de sistema y algoritmos, sus módulos, grado de madurez o sistema de entrenamiento.* 8.2.- *El acceso al "código fuente".* 8.3.- *Transparencia técnica sobre el sistema IA que impone el AIA para el público en general, usuarios o entidades y autoridades supervisoras.* 9.- Transparencia sobre el fundamento y lógica del sistema IA y su incidencia en la toma de decisiones y la transparencia en razón del artículo 22 RGPD. 10.- Transparencia sobre las garantías del sistema, estudios de impacto, medidas de mitigación de riesgos, garantías de supervisión. 10.1.- *Transparencia de las evaluaciones de impacto y su confidencialidad.* 10.2.- *Transparencia de las medidas y mitigaciones frente a riesgos.* 10.3.- *Transparencia sobre supervisión humana y auditorías.* 10.4.- *Transparencia "interna" sobre garantías del sistema en razón del AIA.* 11.- Transparencia de los algoritmos públicos y la motivación comprensible de las decisiones públicas con IA. 12.- Transparencia en la contratación pública de IA. 13.- ¿Y cómo ha de ser la información? 14.- Para concluir: Recapitulación del detalle de la información algorítmica, ¿quién accederá a ella? - Bibliografía.



What concrete transparency and information from algorithms and artificial intelligence is due

ABSTRACT: The study details what data, information or knowledge should be provided on AI systems. It is mentioned whether these are present or future regulatory authority obligations, or if they are best practices found in the UK, the Netherlands or France. Thus, information must be provided on the existence of the system, active dissemination obligations and future public algorithm registers; information on who are the algorithm registries; information on who the AI system providers are and who the users are, the data protection officers, why the system protection officers, why the system exists and its specific purposes. Also information about the training data or input data, as well as the type and technical specifications of the AI system, including its code. The information is specified in accordance with Article 22 RGPD. Likewise, information on the guarantees of the IA system, such as impact studies, risk mitigations or human supervision.

KEYWORDS: Artificial intelligence, transparency of algorithms, explainability, data protection, rights.



1.- Transparencia y explicabilidad algorítmicas: el qué depende de para qué o para quién

El principio de transparencia algorítmica está presente en 73 de 84 de los documentos de referencia sobre inteligencia artificial (IA)². Otro estudio esencial sobre los principios éticos de la IA señala que el 94% de los documentos exhaustivamente analizados incluyen la transparencia como uno de los principios consensuados en los documentos básicos (la referencia, Fjeld *et al*, 2020, por mi parte, Cotino, 2019 a). El principio de transparencia y explicabilidad es la "pieza crucial que falta en el rompecabezas" pues "complementa los otros cuatro principios" (HLEG, 2019: 10) básicos de la ética de la IA: beneficencia ("hacer el bien"), no maleficencia ("no hacer daño"), autonomía o acción humana ("human agency") ("respeto por la autodeterminación y elección de los individuos") y justicia ("Trato justo y equitativo para todos", ídem, también, Floridi *et al*, 2018: 699-700). Así, sin transparencia no se entiende el bien o el daño de la IA, sin transparencia no se sabe cómo actuaría la IA en lugar de nosotros (autonomía). Tampoco se puede determinar la justicia y la responsabilidad. La transparencia -y explicabilidad- son el principio más consensuado, eso sí, con muchos contenidos y significados bajo dicho *paraguas*.

Y precisamente la transparencia es una herramienta, un instrumento y, por ello hay que tener en cuenta en los muy variados motivos, finalidades, intereses o derechos a los que la transparencia sirve. No en vano, el contenido o alcance de la transparencia se hace depender en muy buena medida de tales finalidades. Así, esencialmente la transparencia algorítmica es un elemento instrumental que favorece la necesitada confiabilidad en la IA (Cotino, 2022 c). Por un lado, la transparencia algorítmica es la herramienta básica de la garantía de toda una serie de principios democráticos y derechos fundamentales. Así, se exige en razón de la *accountability* y el principio democrático especialmente en los supuestos en los que se "traduce" la ley en un código. También, se dan obligaciones de publicidad activa como de derecho de acceso a la información pública respecto de los programas, algoritmos y sistemas de IA usados por el poder público, los datos que utilizan y los registros que deja la actuación automatizada. Asimismo, el régimen de protección de datos incluye fuertes obligaciones de transparencia y de derecho de acceso, incrementadas en el caso de decisiones adoptadas por sistemas algorítmicos y de IA. La transparencia también es esencial para poder controlar errores, discriminación

² Mantelero (2022: 98), siguiendo a Jobin *et al* (2019: 98) señala que la transparencia aparece en 73/84; no maleficencia 60/84; responsabilidad 60/84; privacidad 47/84; beneficencia 41/84; libertad y autonomía 34/84; confianza 28/84; sostenibilidad 14/84; dignidad 13/84, y solidaridad 6/84. Menciona datos de Hagendorff en donde la transparencia también era uno de los principios mencionados en más del 50% de documentos.



y sesgo algorítmicos públicos y privados. Y se refuerzan las garantías de transparencia y explicabilidad cuando hay un uso de la IA en el ámbito sancionador, policial y judicial, en razón de las garantías de la actuación administrativa, buena Administración, no arbitrariedad, debido proceso o no discriminación. Todos estos caminos conducen a la *Roma de la transparencia*. Y en buena medida se traducen en fuertes garantías de transparencia "externa", es decir, de facilitación de datos, información o conocimiento de cara a toda la ciudadanía, a la comunidad científica o cuanto menos a todos los interesados o afectados por la IA³, que también se denomina "explicabilidad pública"⁴.

Pero, por otro lado, más allá de estos derechos, principios e intereses públicos (aunque relacionada con ellos), la transparencia es instrumento esencial para el conocimiento y comprobación del buen funcionamiento del sistema de IA por los sujetos de la cadena de valor (usuarios del sistema, importadores, distribuidores, etc.), así como todos aquellos que tienen que verificar, comprobar el mismo (autoridades, evaluadores, etc.). Ésta sería la "transparencia interna", es decir, la transparencia e información para los usuarios afectado por el sistema de IA, así como en beneficio de los auditores internos y externos, las autoridades supervisoras en su caso judiciales. Para qué y para quién es la transparencia algorítmica son elementos esenciales para su determinación sobre los que he tenido ocasión de dedicar otros estudios. El grado de detalle de la información obviamente ha de variar en unos y otros casos y también la aplicabilidad de conocidas técnicas de información por capas (información más superficial o más profunda). De igual modo, según el destinatario de la transparencia se pueden articular técnicas para preservar la reserva, confidencialidad o el secreto que hay que garantizar en algunos supuestos. Así sucede especialmente cuando se trata de la transparencia interna y la posibilidad de facilitar información más sensible a quienes contratan el sistema de IA o quienes los evalúan o controlan.

Como especialmente he analizado en reciente estudio sobre transparencia y explicabilidad (Cotino 2022 c), la "transparencia" algorítmica es un concepto general incluso con una miríada de nociones que orbitan a su alrededor, a saber: *trazabilidad, explicabilidad, interpretabilidad, comprensibilidad, inteligibilidad, legibilidad, auditabilidad, testabilidad, comprobabilidad, simulabilidad, descomponibilidad, verificabilidad, replicabilidad, comunicación, código abierto*, así como referencias a la *transparencia técnica en sentido estricto*, o la importante

³ Con mayor detalle Cotino, 2022 d. Esta interesante distinción transparencia externa-interna, parte del Ministerio del Interior Países Bajos (2021: apartado, 2B.4), su conocido en estudio de impacto de algoritmos.

⁴ *Ibidem*, con remisión al *Toetsingskader Algemene Rekenkamer*, esto es, Marco de pruebas de algoritmos: primeros pasos. Disponible en <https://bit.ly/3kpbuAo>



diferencia que se ha seguido entre *transparencia externa e interna*. Me remito a la concreción de estos conceptos.

También, en dicho estudio introducen algunos elementos generales sobre ¿cuánta transparencia? Así, se afirma que el grado de transparencia habrá de ser mayor a impacto y riesgo por el uso de la IA. También habrá de ser mayor según la mayor opacidad y el tipo de algoritmo. Del mismo modo, a mayor participación, control y supervisión humana en las decisiones de la IA es posible exigir menor transparencia y explicabilidad del sistema de IA utilizada. Asimismo y como principio, también mayor transparencia se ha de dar cuando haya un uso de algoritmos por el poder público. Lamentablemente, el muy mejorable artículo 41 Ley 40/2015 y otras regulaciones no parecen ir en esa dirección y al menos en España el sector público no tiene casi interés en exigir garantías de transparencia y explicabilidad cuando es él quien utiliza los algoritmos y sistemas de IA.

Pues bien, a partir de estas premisas, el presente estudio se centra en qué transparencia algorítmica, esto es, qué datos, información o conocimiento debe facilitarse. La disposición o liberación de esta información no absorbe totalmente qué transparencia algorítmica, pues la transparencia también se puede traducir en la apertura y acceso al sistema para poder operar, testear y *jugar* con él y poder conocerlo. Asimismo se hace algún apunte de cómo hay que disponer estos datos, información y conocimiento. En la descripción de los variados contenidos de la transparencia se menciona si se trata de exigencias normativas presentes o futuras. Hoy por hoy las exigencias de transparencia se dan bien por normativa de protección de datos en general, bien por sus exigencias especiales respecto de decisiones automatizadas (art. 22 RGPD y afines). También hay algunas normativas sectoriales (como la *ley rider* -art. 64.4.d) Estatuto de los Trabajadores). Especial importancia tiene la normativa aplicable al uso público de algoritmos e IA (una clara visión de conjunto de interés, Jiménez, 2022; Boix, 2022), pese a que en España -a diferencia de Francia, por ejemplo- es una regulación muy mejorable. En el contexto público la transparencia algorítmica se da por exigencias constitucionales (arts. 9.3, 24 CE) y legales generales, así como por algunas regulaciones específicas aplicables (art. 41 Ley 40/2015, arts. 13 y 11 artículo 13 Real Decreto 203/2021). Más positivas son regulaciones novedosas como la avanzada Ley 1/2022 valenciana de transparencia o el artículo 23 Ley 15/2022 inspirada en el apartado XVIII de la Carta de Derechos Digitales, a la que también se presta singular atención. En todo caso, aunque no está aprobado, se tiene especialmente en cuenta la transparencia exigida en razón del futuro Reglamento de IA de la UE ("Artificial Intelligence Act", en adelante AIA). Además de la normativa, en este estudio se acude a las autoridades de referencia (normalmente internacionales, europeas y comparadas) que van detallando las exigencias normativas o van proponiendo los estándares de transparencia deseables en general o en particular para el sector público. En este punto hay que destacar desde ya los especiales aportes para la transferencia autolimita pública,



por el ICO inglés (2021 a, b, ICO y Alan Turing Institute, 2022), desde Países Bajos (en especial M^o Justicia, 2021) y desde el Etalab francés (2022). Por su parte, los organismos de normalización técnica NIST (EEUU, 2022 a y b), ISO, CEN CENELEC (UE) por lo general abordan -e incipientemente- la transparencia interna, no la transparencia dirigida al público en general⁵.

2.- La premisa: la “comunicación” de la existencia del sistema IA y notificaciones de su uso. Registros y bases de datos

El Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial de la Comisión Europea (HLEG, 2019) incluye en la transparencia a la “comunicación”, esencialmente como “derecho a saber que [las personas] están interactuando con un sistema de IA” (n^o 78) y para ello los sistemas de IA se deben identificar como tales a los humanos.⁶ Siguiendo el estudio de referencia de Harvard, la comunicación se desdobra en primer lugar, como un principio de “notificación cuando se interactúa con un sistema de IA” (J. Fjeld, *et al.*, 2020), de modo que “los seres humanos siempre deben ser conscientes cuando se relacionan con la tecnología en lugar de hacerlo directamente con otra persona”. Ello es especialmente respecto de “interacciones de chatbots, los sistemas de reconocimiento facial, los sistemas de puntuación de crédito y, en general, “cuando los sistemas de aprendizaje automático se utilizan en la esfera pública”. Igualmente se afirma un “principio de interacción” para advertir que existe, por ejemplo, un sistema de reconocimiento facial.

A este respecto cabe destacar el artículo 52 del futuro AIA, que no está reservado a los sistemas de alto riesgo: “Obligaciones de transparencia para determinados sistemas de IA”. Se dispone que “Los proveedores garantizarán que los sistemas de IA destinados a interactuar con personas físicas estén diseñados y desarrollados de

⁵ En el contexto de ISO cabe remitir ISO/IEC TR 24027 *Sesgo en sistemas de IA y toma de decisiones asistida por IA* ISO/IEC TR 24029-1 *Evaluación de la robustez de la red neuronal* ISO 26000 *Guía sobre responsabilidad social*. Y en desarrollo: ISO/IEC DIS 23894 *Gestión de riesgos*, ISO/IEC AWI TS12791 *Tratamiento del sesgo no deseado en la clasificación y tareas de aprendizaje automático de regresión* así como ISO/IEC AWI 12792 *Taxonomía de transparencia de los sistemas de IA*. En el contexto de CEN CENELEC de la Unión Europea, se sigue en 2022 el proceso de adopción de las primeras normas, cabe destacar la CEN/CLC/JTC 21 - *Artificial Intelligence*.

⁶ N^o 78: “Comunicación. Los sistemas de IA no deberían presentarse a sí mismos como humanos ante los usuarios; las personas tienen derecho a saber que están interactuando con un sistema de IA. Por lo tanto, los sistemas de IA deben ser identificables como tales. Además, cuando sea necesario, se debería ofrecer al usuario la posibilidad de decidir si prefiere interactuar con un sistema de IA o con otra persona, con el fin de garantizar el cumplimiento de los derechos fundamentales. Más allá de lo expuesto, se debería informar sobre las capacidades y limitaciones del sistema de IA a los profesionales o usuarios finales; dicha información debería proporcionarse de un modo adecuado según el caso de uso de que se trate y debería incluir información acerca del nivel de precisión del sistema de IA, así como de sus limitaciones.”



forma que dichas personas estén informadas de que están interactuando con un sistema de IA" (1º). Asimismo se debe informar del funcionamiento del sistema -y obviamente de su existencia- a las personas "expuestas" a "un sistema de reconocimiento de emociones o de un sistema de categorización biométrica" 2º). También se debe conocer cuando un "sistema de IA que genere o manipule contenido de imagen, sonido o vídeo que se asemeje notablemente a personas, objetos, lugares u otras entidades o sucesos existentes, y que pueda inducir erróneamente a una persona a pensar que son auténticos o verídicos (ultrafalsificación), harán público que el contenido ha sido generado de forma artificial o manipulado."

También la transparencia como "comunicación" se desdobra en un "principio de notificación" de la adopción de decisiones por un sistema de IA (J. Fjeld, *et al.*, 2020), que vendría acompañado en su caso de la opción de no utilizar esos sistemas o aportar la información a los mismos. Se conecta especialmente respecto de las decisiones de sistemas IA sobre individuos y las garantías frente a las mismas, a las que se hará especial referencia más adelante.

El AIA para algunos sistemas de alto riesgo crea la "Base de datos de la UE para sistemas de IA de alto riesgo independientes" y la información de esta base de datos "será accesible para el público" (art. 60.2º). En dicha base de datos tienen que registrarse algunos sistemas de alto riesgo⁷ (art. 51 AIA) y todo el mundo puede acceder a información bastante básica que se detalla en el Anexo VIII. Se trata de un esquema de transparencia activa que se daba antiguamente con las bases de datos personales, que habían de registrarse y su información básica se hacía pública. No obstante, el sistema se abandonó con el RGPD pero se retoma en el caso de los sistemas de alto riesgo.

3.- Hacia la transparencia activa sobre la existencia de sistemas de IA públicos y su grado de incidencia en las decisiones públicas

En el caso del sector público es especialmente preocupante que tan siquiera sepamos qué sistemas de IA existen e impactan en la actuación y prestación de servicios así como en nuestros derechos (en especial, Cotino, 2021).

Desde 2019 la Red de Derecho Administrativo de la IA (Red DAIA, 2019 a y b) hemos criticado y mucho que no haya un mapeo de los usos de IA en el sector público. De hecho, al menos en los casos en los que el sistema de IA trate datos personales considero que sería una obligación legal difundir estos tratamientos en razón del artículo 6 bis de la Ley 19/2013, de 9 de diciembre, de transparencia, respecto del

⁷ En concreto, respecto de los sistemas de alto riesgo según el listado del Anexo III (art. 6. 2º), esto es, los otros sistemas de alto riesgo en razón del artículo 6.1º y Anexo II no tienen que registrarse en la base de datos.



inventario de actividades de tratamiento en aplicación del artículo 31 de la Ley orgánica 8/2018. Y hoy por hoy no se conoce un caso en el que se cumpla dicha obligación.

La situación actual es ciertamente lamentable por cuanto sólo se conocen los sistemas IA que los poderes públicos tienen a bien exhibir y en modo alguno sistemas masivamente utilizados por la Agencia Tributaria o la Tesorería General de la Seguridad Social, entre otros, que únicamente se intuyen. Es casi imposible conocer los sistemas algorítmicos a través de anuncios en plataformas de contratación, ciertamente la función de transparencia de estas plataformas es otra y no hay homogeneidad alguna (Gutiérrez, 2021: 29-30). De hecho, cuando se conoce la existencia pública de un sistema algorítmico relevante, resulta casi imposible conocer los detalles técnicos que pueden tener importancia para conocer el impacto en derechos e intereses. Así sucede por ejemplo con el sistema *Viogen*⁸. El ejercicio del derecho de acceso a la información pública puede ser útil pero tiene sus límites⁹. A nivel europeo se presentó un primer mapeo de sistemas públicos de IA (Misuraca, 2020), que ha sido completado recientemente (AI Watch, 2022 a y también 2022 b). En cualquier caso y pese a la utilidad de estos documentos, se trata esencialmente de los usos que el propio poder público quiere dar a conocer.

Es especialmente difícil conocer los sistemas de IA que se están usando para asistir o apoyar decisiones públicas, especialmente en fases previas o preparatorias. Ello queda en general bajo el régimen parco y difuso de la "información y actuaciones previas" del artículo 55 Ley 39/2015 de procedimiento. En esta fase, la luz de la transparencia es muy tenue y difícilmente alcanza ni el derecho de acceso al expediente del propio interesado (artículo 53 Ley 39/2015) ni la legislación de transparencia. El ámbito tributario utiliza intensivamente estos sistemas, pero su opacidad está especialmente protegida por el artículo 170.7 Real Decreto 1065/2007, de 27 de julio que dispone que "los medios informáticos de tratamiento de información y los demás sistemas de selección de los obligados tributarios que vayan a ser objeto de actuaciones inspectoras tendrán carácter reservado, no serán objeto de publicidad o de comunicación ni se pondrán de manifiesto a los obligados tributarios ni a órganos ajenos a la aplicación de los tributos." Asimismo, se trata de una fase de la actuación administrativa en buena medida inmune al control judicial.

⁸ Véase el reportaje publicado en el diario El País (elpais.com) el 10 de abril de 2022 bajo el título "VioGén: visita a las tripas del algoritmo que calcula el riesgo de que una mujer sufra violencia machista". Amnistía Internacional y la Universidad de Valencia están intentando recabar información sobre el sistema. Cabe también remitir al informe presentado por Fundación Éticas (2022).

⁹ Quien suscribe el 29 de julio de 2020 ejerció un derecho de acceso a la información pública. Así, ante la Dirección General de Planificación Estratégica, Calidad y Modernización de la Conselleria de Justicia, que fue resuelto favorablemente el 20 de agosto de 2020 (GVRTE/2020/1163199). Se formularon 17 preguntas. He dispuesto un acceso a la resolución que incluye las preguntas formuladas en <https://bit.ly/37f2apR>



No en vano, la determinación de los obligados tributarios que vayan a ser objeto de comprobación "tiene el carácter de acto de mero trámite y no será susceptible de recurso o reclamación económico-administrativa" (art. 170.8º Real Decreto 1065/2007). Así pues, estos usos que son realmente importantes hoy día quedan fuera de la actuación formal y por ello se facilita una *huida del Derecho*: garantías de transparencia, respeto del derecho a una buena administración, garantías del procedimiento, etc. y también prácticamente blindados al control judicial. Debería proclamarse una clara trazabilidad y transparencia del uso de sistemas algorítmicos en cualquier fase y tipo de actuación administrativa -formal o informal- que, entre otras cosas, permita determinar el grado real de intervención humana en la toma de decisiones (Red DAIA, 2019a, conclusión nº 3).

La Carta de Derechos digitales en su apartado XVIII (Cotino, 2022 b) respecto de la Administración y en concreto su punto 5º señala que "el poder público autor de una actividad en el entorno digital deberá identificar a los órganos responsables de la misma." Se sigue así lo dispuesto en el artículo 41. 2º Ley 40/2015 cuando impone que en las actuaciones automatizadas "se indicará el órgano que debe ser considerado responsable a efectos de impugnación." Poco parece aportar pues la Carta a este respecto. Bien es cierto que es de interés que la letra b del apartado 6 de este XVIII, afirma como derecho la transparencia, que incluye entre otros elementos la obligación de informar "su carácter decisorio o no decisorio" del uso de IA público.

No obstante, dada la relevancia de la cuestión podría haberse ido más allá. Precisamente desde DAIA quien suscribe propuso que este apartado 5º se completara (Red DAIA, 2020). En concreto, propuse que dicho apartado incluyera "De igual modo, en el caso de tratamientos automatizados o uso de sistemas de IA habrá de detallarse si se trata de una decisión plenamente automatizada. Igualmente se concretará en qué medida el sistema o tratamiento utilizado en las diversas fases de la actuación predetermina el contenido de actuaciones administrativas, como pueda ser el caso de actuaciones de contratación, adjudicación de ayudas o subvenciones, excepto en las actuaciones plenamente regladas.

Asimismo, en el apartado relativo al reconocimiento de derechos, en general propuse que los mismos será de aplicación "en el marco de la actuación administrativa *en todas sus fases, incluidas las fases información y actuaciones previas*". Es más, entre los derechos, propuse añadir otra letra relativa a esta cuestión, entre los derechos que se reconocen: "x) En el caso de que la decisión administrativa adoptada reproduzca esencialmente la propuesta del sistema automatizado o inteligente, o de IA, habrá de garantizarse que el interesado conozca dicha situación para poder ejercer sus derechos."

A este respecto, en las aportaciones que lanzamos desde OdiselA (2021), Gutiérrez recuerda que como mínimo el interesado tiene derecho a conocer de forma explícita



desde el inicio del procedimiento o de la actuación material que la decisión será objeto de automatización total o parcial y, en su caso, en qué fase o fases del procedimiento se aplica dicha automatización. Esa información permitiría individualizar el posible "vicio del acto" o de un "elemento concreto del acto" de cara a su impugnación. La profesora se preguntaba, con relación a los derechos del interesado en el procedimiento administrativo previsto en el art. 53.1.a) Ley 39/2015, o a la puesta de manifiesto del expediente en el trámite de audiencia ¿cómo se materializaría dada la concepción restrictiva actual del expediente administrativo (art. 70.4 Ley 39/2015), al no formar parte del expediente la información contenida en "aplicaciones, bases de datos o informes internos" (supongamos que esos informes se han generado de forma automatizada por un algoritmo).

Esta situación de clara opacidad se está remediando sólo timidamente en España. Las insuficiencias del artículo 41 de la Ley 40/2015 son palmarias y el artículo 13 del Decreto 203/2021 sólo las atempera. Así, en razón de este artículo 13 sólo para la AGE un sistema que adopte decisiones íntegramente automatizadas debe ser autorizado por resolución. Por lo que ahora interesa, dicha resolución, expresando los recursos procedentes, se publicará en la sede de la AGE. El contenido de la resolución que ha de publicarse puede adivinarse en razón del artículo 11. 1. *i*, pues habrá de incluir una "relación actualizada de las actuaciones administrativas automatizadas vinculadas a los servicios, procedimientos [...]. Cada una se acompañará de 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". Posiblemente algún organismo de la AGE ya cumpla con esta obligación, pero lo cierto es que no ha sido posible localizar un solo ejemplo en las sedes electrónicas.

Especialmente novedosa es la Ley 1/2022, de 13 de abril, de la Generalitat, de Transparencia y Buen Gobierno de la Comunitat Valenciana. Su artículo 16.1 l) impone la publicidad activa "de sistemas algorítmicos o de inteligencia artificial que tengan impacto en los procedimientos administrativos o la prestación de los servicios públicos"¹⁰. La implantación de esta obligación no es sencilla¹¹ y se pretende superar experiencias voluntarias como las del City of Helsinki AI Register¹² o el de

¹⁰ "l) La relación de sistemas algorítmicos o de inteligencia artificial que tengan impacto en los procedimientos administrativos o la prestación de los servicios públicos con la descripción de manera comprensible de su diseño y funcionamiento, el nivel de riesgo que implican y el punto de contacto al que poder dirigirse en cada caso, de acuerdo con los principios de transparencia y explicabilidad."

¹¹ Quien suscribe dirige el grupo de investigación de la Universidad de Valencia que suscribe un Convenio de colaboración Científica con la Dirección General de Transparencia desde 2022 para los estudios e investigación de la implantación de la publicidad activa de los sistemas algorítmicos y de inteligencia artificial usados por el sector público, exigida por la Ley 1/2022, de 13 de abril.

¹² Helsinki: <https://ai.hel.fi/en/ai-register/>



Ámsterdam¹³. En Francia destaca el inventario de algoritmos de la ciudad de Antibes Juan-les-Pins¹⁴ y las listas de algoritmos de Nantes-Métropole¹⁵ y de Pôle Emploi¹⁶. Pese a que se pueda señalar que hay *más* ruido que nueces (Benítez, 2022), ciertamente en Reino Unido, Francia o Países Bajos se trata de acciones de gran profundidad. Aún sin que sea obligatorio normativamente quizá el mejor ejemplo en España lo brinda el ámbito de salud catalán¹⁷.

4.- Transparencia sobre quiénes son los proveedores de sistemas de IA y quiénes los utilizan y un contacto

No es momento ahora de detallar la compleja cadena de valor en la IA. Entre los distintos "operadores" de los sistemas IA¹⁸, basta distinguir ahora esencialmente las figuras del o proveedor del sistema (por ejemplo, la empresa que diseña y desarrolla el sistema y lo pone en el mercado para su uso), en su caso del importador o distribuidor¹⁹, así como la posición del "usuario" que utiliza el sistema desarrollado (por ejemplo, una empresa o administración que ha contratado el uso del sistema). Hay que distinguir estas posiciones del proveedor y del usuario quienes son los colectivos o personas afectadas por el uso del sistema de IA. Y es que hay que llamar la atención de un equívoco habitual. En el caso de la IA, los consumidores o "usuarios" no son las personas interesadas o afectadas, que son normalmente personas físicas o jurídicas a cuyos derechos impacta el uso de la IA por los usuarios (por ejemplo, el trabajador o cliente de esa empresa o el ciudadano ante la Administración usuaria de IA).

Así las cosas, a la hora de informar sobre quién, se tratará de informar sobre quién es la entidad que ha desarrollado o es el proveedor del sistema de IA de alto riesgo. Para los usuarios del sistema, el futuro AIA en su artículo 13. 3º a) obliga al proveedor

¹³ Ámsterdam: <https://algorithmeregister.amsterdam.nl/en/ai-register/>

¹⁴ Antibes Juan-les-Pins: <https://bit.ly/3wekNFO>

¹⁵ Nantes-Métropole: <https://bit.ly/3wfoKqC>

¹⁶ Pôle Emploi: <https://www.pole-emploi.fr/candidat/algorithmes.html>

¹⁷ *Observatori IA en salut, Identificació d'algorismes d'IA* <https://bit.ly/3ilnyvT>

¹⁸ EL artículo 3 AIA introduce la noción inclusiva de "Operador" (nº 8), esto es, "el proveedor, el usuario, el representante autorizado, el importador y el distribuidor."

¹⁹ Artículo 3, AIA "2) Proveedor»: toda persona física o jurídica, autoridad pública, agencia u organismo de otra índole que desarrolle un sistema de IA o para el que se haya desarrollado un sistema de IA con vistas a introducirlo en el mercado o ponerlo en servicio con su propio nombre o marca comercial, ya sea de manera remunerada o gratuita."

[...] 7) Distribuidor: toda persona física o jurídica que forme parte de la cadena de suministro, distinta del proveedor o el importador, que comercializa un sistema de IA en el mercado de la Unión sin influir sobre sus propiedades."



a especificar la información sobre "a) la identidad y los datos de contacto del proveedor y, en su caso, de su representante autorizado".

Ya de cara al público en general (y por ende potenciales interesados o afectados) y como se ha anticipado el futuro AIA obliga a registrar determinados sistemas de alto riesgo y la información básica sobre los mismos crea la "Base de datos de la UE para sistemas de IA de alto riesgo (art. 51 AIA) y la información de esta base de datos "será accesible para el público" (art. 60.2º). Por lo que ahora interesa, se incluye información del proveedor del sistema IA, en su caso del representante²⁰.

Para el caso del uso público de IA, el ICO (2021; 2022) ha concretado la información que se ha de dar respecto de los sujetos y para todo el público en su guía de transparencia algorítmica pública del ICO para el Reino Unido. Así, se solicita que se enumere quién es responsable de implementar su herramienta, detallando la organización, el equipo responsable de la herramienta; propietario principal responsable; proveedor externo o cualquier tercero involucrado. Se ha de informar también si la herramienta ha sido desarrollada externamente. Incluso hay que identificar el número de Registro Mercantil de su proveedor externo y el papel desarrollado y los términos de su acceso a cualquier dato del gobierno. En Francia el Etalab impone informar sobre "la administración responsable de la decisión" (2022: 15) en la mención explícita obligatoria y para el registro de algoritmos: Nombre de la administración Dirección/departamento responsable; Contacto en la administración de que se trate (2022: 18).

La reciente Ley 1/2022, de 13 de abril, valenciana que impone la publicidad activa en el portal de transparencia sobre algoritmo y sistemas IA obliga a informar de "el punto de contacto al que poder dirigirse en cada caso, de acuerdo con los principios de transparencia y explicabilidad." (art. 16 l).

En el caso de que el sistema IA trate datos personales, la información que obligatoriamente debe facilitarse al público (de cara a los propios interesados) debe conjugarse con las figuras de responsable, corresponsables o encargados. En esta dirección, el artículo 13 RGPD obliga a facilitar información sobre "a) la identidad y los datos de contacto del responsable y, en su caso, de su representante; b) los datos de contacto del delegado de protección de datos, en su caso;" Y, también, "e) los destinatarios o las categorías de destinatarios de los datos personales, en su caso;". Piénsese que, en principio, el "usuario" del sistema de IA que trata datos personales será el "responsable" del tratamiento, mientras que el proveedor o desarrollador del sistema de IA en principio será un encargado del tratamiento. Así, en principio, por

²⁰ Así: "1. El nombre, dirección y datos de contacto del proveedor"; "2. Cuando sea otra persona la que presente la información en nombre del proveedor, el nombre, dirección y datos de contacto de dicha persona; 3. El nombre, dirección y datos de contacto del representante autorizado, en su caso."



normativa de protección de datos, habrá de informarse de la identidad del usuario del sistema y no del desarrollador o proveedor. Ello será así siempre que el desarrollador o proveedor del sistema de IA no haga un tratamiento con esos datos personales para sus finalidades propias, cosa que no es en modo alguno inhabitual en la cadena de valor de la IA. En este último caso, sería también responsable del tratamiento y también habría de informar en razón del RGPD. En la misma dirección, si el usuario del sistema IA, que es el responsable, comunica datos personales a un tercero (como por ejemplo, datos inferidos por el sistema IA relacionados con el afectado), debe facilitarse la información del destinatario.

5.- Transparencia sobre por qué existe el sistema IA y explicación de que no había mejores alternativas

Resulta de interés destacar que las exigencias de la transparencia, al menos en el ámbito público, parecen exigir también dar explicaciones sobre "cómo y bajo qué criterios" se han adoptado decisiones sobre el uso de sistemas algorítmicos por las instituciones públicas. Y, de especial interés, "por qué dicha implementación tecnológica innovadora es la mejor alternativa frente a otras soluciones". Se requiere también que se "enumere las alternativas no algorítmicas que consideró, si esto se aplica a su proyecto, o una descripción de su proceso de toma de decisiones antes de presentar la herramienta (2.3)" (ICO, 2021 b, detallando la información de segundo nivel).

Se trataría de información valiosa que queda estrechamente ligada al principio de proporcionalidad como alternatividad tanto en general para el poder público como en especial cuando haya restricciones o impactos en derechos fundamentales. Es decir, se parte de los peligros que implica el uso de algoritmos públicos y sobre esta base se requiere que, al menos, se hayan valorado alternativas para no utilizar un sistema de IA. Ello puede especialmente ser importante en los casos en los que se opte por el uso de sistemas de IA intrínsecamente opacos y, en particular, por el poder público.

En una línea similar y para el sector público o privado, el caso de sistemas de IA que traten con datos personales, la AEPD apunta en su lista de evaluación de auditoría que "Teniendo en cuenta criterios de eficiencia, calidad y precisión del componente IA, se ha elegido el modelo más adecuado (usando criterios de simplicidad e inteligibilidad), entre varios componentes concurrentes, y desde el punto de vista de su codificación para facilitar la legibilidad, comprensión de su lógica, la coherencia interna y la explicabilidad" (AEPD, 2021: 15-16). En esta línea, las Directrices para el sector público de Países Bajos (M^o Justicia, 2021) también señalan que se facilite información de por qué se realiza análisis de datos y, por lo que más interesa, "por qué el uso del análisis de datos es proporcionado, y no había mejores alternativas para lograr el objetivo".



6.- Transparencia sobre la finalidad prevista del sistema IA

Además de la información sobre la existencia misma del sistema de IA, resulta esencial la información sobre para qué sirve, esto es, detallar los fines, generales y concretos, del modelo y en su caso la prioridad de los mismos (Gutiérrez, 2021: 55-56). Un elemento esencial es la finalidad prevista del sistema de IA (artículo 3. 12^o²¹), esto es, el uso para el que un proveedor concibe un sistema de IA. Precisamente la finalidad prevista es la que determina que un sistema se considere o no de alto riesgo y, por tanto, quede sujeto a un régimen jurídico muy exigente por su peligrosidad e impacto en los afectados. Como claro ejemplo, el Anexo III AIA determina como sistemas de alto riesgo los que se utilicen con las finalidades de identificación y categorización de personas físicas con sistemas biométricos e IA; componentes de seguridad en la gestión y explotación del tráfico rodado y el suministro de agua, gas, calefacción y electricidad; para controlar las emisiones y la contaminación; para el acceso, admisión o la asignación de personas programas educativos; para evaluar los resultados del aprendizaje o dirigir el proceso de aprendizaje. En el ámbito laboral la IA, para la contratación, evaluación de candidatos, promoción y la terminación de las relaciones contractuales, asignación de tareas, etc.

Hay que destacar que en términos del AIA, el proveedor debe valorar cuál es el "rendimiento" del sistema, esto es "la capacidad de un sistema de IA para lograr su objetivo", art. 3. 18). Debe señalarse que, esencialmente, se requiere que la finalidad prevista o propósito del sistema IA y la capacidad del mismo sirva principal y directamente a la finalidad considerada de alto riesgo por el anexo III. Es decir, un sistema de IA no deberá considerarse de alto riesgo si es utilizado de manera meramente instrumental o auxiliar para las finalidades del anexo III. Es decir, que su utilización para tal finalidad "no es puramente accesorio" (art. 6. 3 b) versión Presidencia checa julio 2022), sino que en su caso se adoptan decisiones a partir de los datos de salida del sistema de IA. La finalidad prevista por el proveedor o desarrollador es la que tiene que seguir el usuario que utilice el sistema. Si el usuario del sistema IA lo destina a otras finalidades, pasa a quedar obligado como proveedor. Pues bien, en el caso de los sistemas de alto riesgo que regula el AIA, el artículo 13. 3^o sobre transparencia obliga a que el proveedor informe al usuario del sistema (no al afectado) de "b) las características, capacidades y limitaciones del funcionamiento del sistema de IA de alto riesgo, y en particular: i) su finalidad prevista; iii) cualquier circunstancia conocida o previsible, asociada a la utilización del sistema de IA de alto riesgo conforme a su finalidad prevista o a un uso indebido razonablemente

²¹ Cabe recordar la definición del artículo 1. 12) AIA "Finalidad prevista: el uso para el que un proveedor concibe un sistema de IA, incluido el contexto y las condiciones de uso concretas, según la información facilitada por el proveedor en las instrucciones de uso, los materiales y las declaraciones de promoción y venta, y la documentación técnica."



previsible, que pueda dar lugar a riesgos para la salud y la seguridad o los derechos fundamentales; iv) su funcionamiento en relación con las personas o los grupos de personas en relación con los que se pretenda utilizar el sistema”.

Se trata de información del proveedor al usuario. Y debe tenerse en cuenta en general, que la transparencia a la que obliga el artículo 13 AIA del proveedor al usuario, es “un tipo y un nivel de transparencia adecuados para que el usuario y el proveedor cumplan las obligaciones oportunas previstas” En este punto, pese a que la regulación podría ser más detallada, habría que incluir las obligaciones de protección de datos por las que el usuario del sistema IA, como responsable del tratamiento de datos, debe informar a los afectados e interesados en materia de protección de datos. El usuario del sistema debe poder contar con información suficiente para ello por parte de su proveedor de IA que, como se ha indicado, es el “encargado” en términos de protección de datos.

La información de los proveedores sobre la finalidad prevista del sistema IA no es sólo para los usuarios (art 13) o a las entidades de certificación o supervisoras, sino que también debe exponer la finalidad del sistema en la base de datos pública (art. 60.2º, art. 51 AIA, Anexo VIII). Así, se obliga a que estos registros de acceso público incluyan información sobre “5. La descripción de la finalidad prevista del sistema de IA, Anexo VIII).

Respecto del uso público de IA, cabe seguir al ICO, por cuanto se debe informar sobre el alcance de la herramienta, para qué ha sido diseñada y para qué no está destinada. Se incluye “describir para qué se ha diseñado y para qué no se ha diseñado la herramienta, incluidos los fines para los que las personas pueden pensar erróneamente que se utilizará la herramienta” (2.1). Igualmente, “proporcionar una lista de beneficios: valor por dinero, eficiencia o facilidad para el individuo” (2.2). De igual modo, el ICO (2021 b) requiere que se explique “qué problema está tratando de resolver con la herramienta y cómo está resolviendo el problema” y “su justificación o justificación para usar la herramienta”. Resulta también muy relevante el grado de detalle para que se facilite información sobre “con qué frecuencia se usa la herramienta”²².

Sea o no un sistema de alto riesgo, público o privado, si el sistema de IA trata datos personales, según el artículo 13 RGPD hay que facilitar al público información sobre “c) los fines del tratamiento a que se destinan los datos personales y la base jurídica del tratamiento”. No hay que olvidar la importancia de los “Principios relativos al tratamiento”, en particular que los datos son “b) recogidos con fines determinados, explícitos y legítimos” (art. 5 RGPD). Así, habrá que facilitar información respecto de las finalidades “explícitas” y con suficiente grado de concreción respecto del

²² “Por ejemplo, la cantidad de decisiones tomadas por mes o la cantidad de ciudadanos que interactúan con la herramienta (2.5)”. ICO (2021 b).



tratamiento de datos (AEPD, 2018: 7.2 "Finalidad", 10 y ss.). Deben evitarse prácticas tales como incluir finalidades demasiado genéricas o inespecíficas, que puedan conducir a tratamientos ulteriores que excedan las expectativas razonables del interesado. Hay que recordar que para el caso de tratamientos de datos personales, la finalidad o finalidades del tratamiento es un elemento esencial que hay que detallar para poder explicar y ponderar la proporcionalidad, adecuación y necesidad, además de evaluar toda la legitimación.

Es más, en el caso de que el sistema IA trate datos personales y además adopte decisiones esencialmente automatizadas con impacto en las personas (art. 22 RGPD), las autoridades de protección de datos señalan que es obligatorio informar *por qué es necesario ese tratamiento y cuál es su legitimación*, a saber: por qué realiza análisis de datos (cuál es su finalidad y qué se hace con los resultados); por qué el uso del análisis de datos es proporcionado, y no había mejores alternativas para lograr el objetivo. Asimismo, cuál es la base legal para realizar estos análisis (M^o Interior, 2021). También, en clave de protección de datos, debe haber detalle del tratamiento, los perfilados realizados, cómo se elaboran, su pertinencia para el proceso de decisiones automatizadas y cómo se utilizan para una decisión relativa al interesado (G29, 2018 y AEPD, 2020: 22), así como su impacto consecuencias del análisis para los ciudadanos afectados (M^o Justicia, 2021). De nuevo cabe recordar que el proveedor (encargado) habrá de facilitar al usuario del sistema (responsable de protección de datos) información suficiente para que pueda cumplir sus obligaciones de transparencia de protección de datos (art. 13 AIA).

El Etalab francés además de la finalidad del tratamiento en la mención explícita (2022: 15), respecto del inventario de algoritmos señala que hay que informar "si se toma una decisión administrativa, cuál es el objetivo de este algoritmo?"; "¿Cómo y cuándo interviene el algoritmo en la toma de decisiones? Para realizar estas tareas, ¿en qué momento entra el algoritmo en el proceso de decisión? ¿Cuál es el proceso completo y de qué parte o partes se encarga el algoritmo?" (Etalab, 2022: 18).

7.- Transparencia sobre los datos del sistema de IA (de entretenimiento, de entrada y los inferidos por el sistema)

Un elemento esencial sobre la transparencia y la información es la relativa a los datos de entrenamiento del sistema de IA, así como los datos de entrada para su funcionamiento, validación y prueba. Se trata de un elemento clave que incide directamente en la calidad y robustez del sistema, así como respecto de la posibilidad de controlar sesgos, errores o posibles discriminaciones. De igual modo, en el caso nada infrecuente de que el sistema de IA trate datos personales, es una exigencia para la protección de este derecho y otros conexos. Genéricamente, en el estudio de impacto IA en Países Bajos se hace referencia a la transparencia de los datos utilizados, las variables de entrada y los parámetros y umbrales utilizados (M^o Interior, 2021).



Respecto de los sistemas de alto riesgo, el artículo 13 del futuro AIA obliga a los proveedores a informar a los usuarios de "b) las características, capacidades y limitaciones del funcionamiento del sistema de IA de alto riesgo, y en particular: iv) su funcionamiento en relación con las personas o los grupos de personas en relación con los que se pretenda utilizar el sistema;" y por lo que más interesa ahora "v) cuando proceda, especificaciones relativas a los datos de entrada, o cualquier otra información pertinente en relación con los conjuntos de datos de entrenamiento, validación y prueba usados, teniendo en cuenta la finalidad prevista del sistema de IA;".

De cara a las entidades de evaluación de los sistemas de alto riesgo así como a las autoridades supervisoras, el AIA impone a los proveedores de sistemas que quede claramente documentada una "descripción detallada" "d) sobre datos [...] así como los conjuntos de datos de entrenamiento utilizados, incluida la información acerca de la procedencia de dichos conjuntos de datos, su alcance y sus características principales; cómo se obtuvieron y seleccionaron los datos; los procedimientos de etiquetado (p. ej., para el aprendizaje supervisado), las metodologías de depuración de datos (p. ej., la detección de anomalías);", también sobre "g) datos de validación y prueba empleados y sus características principales" (Anexo IV. 1).

El ICO de manera bastante precisa detalla para el sector público que ha de haber transparencia de:

- el nombre de los conjuntos de datos que utilizó, si corresponde (4.1)
- una descripción general de los datos utilizados para entrenar y ejecutar la herramienta, incluida -una descripción de las categorías que se usaron para entrenar, probar u operar el modelo, por ejemplo, 'edad', 'dirección', etc. (4.2)
- la URL de los conjuntos de datos que ha utilizado, si está disponible (4.3)
- cómo y por qué recopila datos, o cómo y por qué los datos fueron recopilados originalmente por otra persona (4.4)
- los acuerdos de intercambio de datos que tiene vigentes (4.5)
- detalles sobre quién tiene o tendrá acceso a estos datos y durante cuánto tiempo se almacenan y en qué circunstancias (4.6, ICO, 2021 b).

Mº Justicia (2022) de Países Bajos entran si cabe en mayor detalle. Así, requiere que se documente y registre datos de entrada y su uso efectivo y relevancia²³, se

²³ En los Criterios, el Mº Justicia (2022) afirma "Registre los datos de entrada (datos/conjuntos de datos de origen) utilizados y utilice sólo los datos necesarios. Documente y registre esto: La documentación técnica especifica las fuentes de datos utilizadas. - Se han definido los datos de entrada (datos



describa la calidad de los datos y cómo se examinó²⁴. De igual modo que se describa, registre y documenten las decisiones sobre la inclusión o no de determinados datos por su calidad o suficiencia²⁵. Y el Etalab (2022: 19) francés en su guía para los registros de algoritmo detalla la información sobre "datos procesados", "Fuente de los datos": "Quién proporciona los datos (el usuario, otra administración, etc.)? ¿Cómo se proporcionan (un archivo, una API, etc.)? [...]" o "Cómo se recogen los datos procesados".

La escasa regulación para el sector público en España, respecto de los sistemas automatizados sí que impone difundir información sobre "los datos utilizados en su configuración y aprendizaje" (artículo 13 del Decreto 203/2021 y en conexión con éste y respecto de la publicidad en sede electrónica, el artículo 11. 1. /).

Señala Gutiérrez al respecto de los datos utilizados por el modelo de IA que debe detallarse también la preparación de los datos, esto es, la selección, ampliación, depuración y eliminación de datos redundantes o erróneos, combinación de datos de varias fuentes, determinación de variables relevantes, tratamiento de datos no estructurados, partición de los *datasets* en datos de entrenamiento, de prueba y validación del modelo. Y en la línea de lo anterior, la verificación de la calidad de los datos (fiabilidad de las fuentes de procedencia, representatividad de los datos, ausencia o minimización de los sesgos, actualización de los datos).

En el contexto de la protección de datos personales, cuando los datos personales no se hayan obtenido del interesado se ha de facilitar información sobre "las categorías de datos personales de que se trate" (artículo 14.1º d) y "la fuente de la que proceden los datos personales y, en su caso, si proceden de fuentes de acceso público" (artículo 14.2º f). Como señala la AEPD también ha de haber "especial indicación de los datos especialmente protegidos" (AEPD, 2018: 15). También la AEPD (2021: 15-16) señala que hay que evaluar si "el origen de los datos está documentado y existe un mecanismo para informar" y "Las características de los datos usados para entrenar al componente IA están identificadas, documentadas y adecuadamente justificadas."

fuentes/conjuntos de datos) que se utilizan. - Los datos utilizados son relevantes. - Los datos utilizados y su relevancia están documentados."

²⁴ *Ibidem* "Describa la calidad de la(s) fuente(s) de datos utilizada(s) y si la(s) fuente(s) de datos es(son) de calidad suficiente para el propósito para el que se utiliza. En una "inmersión profunda en los datos" y también durante la anterior contención del riesgo, se examinó y explicó a fondo la calidad de las fuentes de datos. La documentación técnica también describe cómo se examinó la calidad de los datos. Sólo se incluyen en el modelo fuentes de datos de suficiente calidad."

²⁵ *Ibidem*, "Registre las suposiciones/elecciones utilizadas. No se trata de hacer transparentes todos los supuestos durante la programación, sino de las decisiones que se tomaron: por ejemplo, no incluir ciertos datos en el análisis porque su calidad era insuficiente o porque un conjunto de datos no contiene datos suficientes para llevar a cabo un análisis estadístico."



Las anteriores exigencias se dan siempre que el sistema de IA trate datos personales, pero si además el sistema IA adopta decisiones automatizadas con datos personales (art. 22 RGPD), como luego se detalla, las autoridades de datos subrayan que se informe sobre *los datos empleados y su calidad*. En concreto, el detalle de los datos empleados para la toma de decisión, más allá de la categoría, y en particular información sobre los plazos de uso de los datos (su antigüedad) (AEPD, 2020: 23): qué fuentes de datos de qué organizaciones se utilizan para este fin, y cuál es su calidad (Países Bajos), las categorías de datos que se han utilizado o se utilizarán en la elaboración de perfiles o el proceso de toma de decisiones y por qué estas categorías se consideran pertinentes (G29, 2018). La importancia relativa que cada uno de ellos tiene en la toma de decisión y la calidad de los datos de entrenamiento y el tipo de patrones utilizados. Se añade asimismo respecto de los datos que "En el caso de que el sistema IA contenga información de terceros identificables, la prohibición de tratar esa información sin legitimación y de las consecuencias de realizarlo." (AEPD, 2020: 23). El proveedor habrá de facilitar información suficiente al usuario del sistema para que pueda cumplir con estas obligaciones.

Mención aparte merece el tema de los *datos inferidos*, esto es, los datos, información o conocimiento sobre una persona física que se infieren u obtienen a partir de tratar sus datos, por lo general con sistemas automatizados y de IA. Se trata de un tema respecto del que quien suscribe (Cotino, 2022 a: ap. 9) y Palma (2021 y 2022) hemos prestado especial atención. Por lo que ahora interesa, todos los datos e información generada gracias al procesamiento de datos con sistemas de IA por lo general no son cognoscibles. Considero con el Grupo del Artículo 29 (2018: 10) que estos datos generados por la IA que se vinculen a una persona particular son datos personales aunque no hayan sido facilitados el interesado. En consecuencia, el interesado o afectado sí que tendría derecho de acceso respecto de los mismos, sin perjuicio de posibles límites del artículo 23 RGPD. Sin embargo, el TJUE no es tan claro que respecto de estos datos inferidos pueda ejercerse el derecho de acceso o rectificación²⁶. Es por ello que es necesaria una regulación o criterios interpretativos de las autoridades regulatorias o judiciales.

²⁶ Así, no consideró que procediera reconocer estos derechos respecto de un análisis jurídico que se había realizado sobre un permiso de residencia; STJUE de 17 de julio de 2014, asuntos acumulados, C-141/12 y C-372/12, caso YS. y M. y S. (ver apartados 45 y 48). En cambio, respecto de los datos de evaluación de un examen sí que considero que sí que procedía el acceso, pero la rectificación no amparaba, obviamente, a rectificar a posteriori las respuestas incorrectas. STJUE de 20 de diciembre de 2017, asunto, C-434/16, caso Nowak, apartados 43, 44, 52 y 54.



8.- Transparencia sobre la tipología y especificaciones técnicas del sistema IA

La transparencia de la IA gira especialmente alrededor de los elementos y especificaciones técnicas del sistema de IA. En principio, esta información de carácter técnico parece que habría de estar más reservada a la llamada transparencia interna, esto es, información dirigida a los usuarios del sistema IA, así como para organismos de verificación y autoridades supervisoras. No en vano son quienes requerirán las especificaciones técnicas para el funcionamiento correcto del sistema, su control y supervisión. Por el contrario, y también en principio, estas especificaciones técnicas del sistema quedarían lejos del entendimiento y comprensión por el público en general e incluso por los afectados por los sistemas de IA. Es así que puede tener especial sentido la información por niveles o capas de profundidad según se señala en apartado relativo al cómo dispensar la información.

Sin embargo, ello no siempre ha de ser así por cuanto la importancia de la transparencia descansa en muy buena medida en la potencialidad de que se pueda controlar externamente. Así, pese a que la ciudadanía media no tenga capacidades de entender e interpretar la información técnica sobre un sistema de IA, es bien posible que en la sociedad civil existan expertos que potencialmente puedan analizar los datos e información. De igual modo, no hay que obviar el papel de la comunidad científica, en muchos supuestos aliada con la sociedad civil. De la misma manera, en el caso de interesados o afectados por las decisiones de IA, la defensa de sus derechos e intereses se hará pasar en muchas ocasiones por el análisis de la información que se les facilite por especialistas. Así las cosas, sin perjuicio de que hay datos o información que ha de quedar reservada a usuarios o autoridades, especialmente para proteger los derechos de los proveedores, no hay que excluir la facilitación o liberación de información y datos técnicos sobre el sistema de IA para el público en general.

8.1.- Transparencia del tipo de sistema y algoritmos, sus módulos, grado de madurez o sistema de entrenamiento

Los sistemas automatizados que adopten decisiones en el sector público únicamente han de informar en sede electrónica con una "descripción de su diseño y funcionamiento" arts. 13 y 11. 1. /) Decreto 203/2021). Poco más aporta la Ley 1/2022, de 13 de abril, valenciana que exige la publicidad activa de una "descripción de manera comprensible de su diseño y funcionamiento, el nivel de riesgo que implican [...] de acuerdo con los principios de transparencia y explicabilidad." (art. 16 l).

Desde diferentes organismos se afirma la necesidad de dar información del tipo de modelos y algoritmos, sus módulos, grado de madurez o sistema de entrenamiento.

Así, respecto del tipo de algoritmo, el ICO (2021 b) para el sector público indica la publicación del "tipo de modelo, por ejemplo, un sistema experto o una red neuronal profunda (2.4)". El estudio de impacto IA en Países Bajos (M^o Justicia, 2022) detalla "la



identificación del modelo algorítmico utilizado": Los tipos de algoritmos seleccionados para generar el modelo a partir de los datos (e.g. modelos de regresión, árboles de decisión, bosques aleatorios, redes neuronales; justificación del modelo seleccionado en términos de rendimiento/interpretabilidad (en particular, si se trata de modelos de caja negra) frente a otros más interpretables u otras alternativas tecnológicas que permitan alcanzar los mismos fines establecidos por la organización; los parámetros del modelo, y los hiperparámetros seleccionados con carácter previo al entrenamiento del modelo; las métricas de rendimiento utilizadas para la validación del modelo y análisis de errores de entrenamiento y de generalización según el modelo utilizado." Se trata de un elemento importante, que también se puede vincular a los análisis de proporcionalidad en la alternatividad en la elección del sistema y de los algoritmos concretos.

La Directrices del Mº Interior (2021: 2.8) de Países Bajos, respecto de la arquitectura del sistema señala que la documentación técnica describe el "método de análisis utilizado" y "la precisión del método analítico". Y el referido estudio de impacto (Mº Justicia, 2022) afirma que hay que dar información sobre cómo se entrena el algoritmo. El ICO (2022 b) para el sector público también requiere que se facilite información sobre el grado de madurez del sistema, esto es, si la herramienta está en la etapa de idea, diseño, desarrollo, producción o retirada, incluida la fecha y hora en que se creó y cualquier actualización. Las referidas Directrices sí que tienen en cuenta medidas específicas para la evaluación del sistema y la organización de la información para que sea posible. Así, se señala que se "organice el código en módulos que puedan ser evaluados por separado y combinados." Asimismo, que "se compruebe que estos módulos funcionan correctamente tanto por separado como en combinación".

Y el Etalab (2022: 19) francés para el registro de algoritmos dedica un apartado sobre "Información sobre el funcionamiento interno del algoritmo" (4.2.4), con indicación de "Tipo de algoritmo", "si se trata de un sistema de reglas (las reglas de cálculo están codificadas por personas) o de un algoritmo basado en aprendizaje automático." "Detalle las operaciones técnicas realizadas por el algoritmo. Esta categoría puede ser simple o compleja, según el tipo de algoritmo utilizado", así como "enlaces a repositorios de código fuente o un diagrama de flujo".

8.2.- El acceso al "código fuente"

El acceso al "código fuente" del sistema IA puede ser especialmente conflictivo. Cerrillo (2020: 68) recuerda que el acceso al código fuente del algoritmo no conlleva necesariamente que el algoritmo sea más transparente y, en particular, que se pueda conocer cómo funciona o cómo ha llegado a un determinado resultado, tampoco que la información pueda ser comprendida por cualquier persona. Por lo que la difusión del código fuente debería ir acompañada de una descripción del algoritmo en lenguaje natural comprensible por cualquier persona.



Se menciona el acceso al código fuente del sistema IA en el estudio de impacto de Países Bajos (Mº Interior, 2022). La Recomendación sobre la ética de la IA UNESCO 2021 lo reserva para supuestos especiales en los que se "puede requerir también que se compartan códigos o conjuntos de datos" (nº 39)²⁷. La propuesta de normativa de la ELI (2022) regula las reservas al acceso al código fuente. Su artículo 8.3º señala que "El acceso al código fuente y a los conjuntos de datos de formación y ensayo podrá limitarse o restringirse totalmente cuando sea necesario para salvaguardar los intereses y derechos legítimos de la autoridad de ejecución, del proveedor del sistema o de terceros."²⁸ Al respecto, pese a que es una de sus propuestas para la transparencia (en el sector público). Respecto del acceso al código fuente, el programa o software que en su caso acompaña al sistema IA, cabe remitir a la rica experiencia comparada especialmente de Francia e Italia que recientemente ha analizado Gutiérrez (2022). De manera bastante avanzada la CADA, autoridad de acceso a la información francesa, como recuerda esta autora ha venido calificando como "documentos administrativos", no sólo el código fuente o los algoritmos implementados por una Administración, sino también la documentación técnica relativa al código fuente, como puede ser el documento de especificación de requisitos de software. Según se acaba de señalar, la guía del Etalab hace referencia a la posible remisión al código fuente.

8.3.- Transparencia técnica sobre el sistema IA que impone el AIA para el público en general, usuarios o entidades y autoridades supervisoras

Para los sistemas de alto riesgo el futuro AIA impone fuertes obligaciones de facilitar información técnica del sistema, aunque variadas ya se trate del público en general, del usuario del sistema o de las entidades evaluadoras o autoridades supervisoras.

El AIA exige al proveedor que en las bases de datos públicas se difunda al público información sobre el nombre comercial del sistema (nº 4), si está comercializado o puesto en servicio (nº 6), datos del certificado expedido (nº 7) así como de la declaración UE de conformidad (nº 10)²⁹. También en general ha de haber

²⁷ Conferencia General 41ª reunión - París, 41 C/73, 22 de noviembre de 2021. Anexo. Se puede consultar el documento en la Biblioteca Digital de la UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000379920_spa

²⁸ Se indica que "los intereses que se protegen están relacionados con los "secretos" en sentido amplio: información personal identificable, privacidad, propiedad intelectual, secretos comerciales (u otros)". El ELI se apoya para ello en Directiva 96/9/CE sobre la protección jurídica de las bases de datos; Directiva 91/250/CEE del Consejo sobre la protección jurídica de los programas de ordenador; Reglamento (UE) 2018/1807 sobre un marco para la libre circulación de datos no personales en la Unión Europea.

²⁹ "4. El nombre comercial del sistema de IA y toda referencia inequívoca adicional que permita su identificación y trazabilidad. [...] 6. La situación del sistema de IA (comercializado o puesto en servicio; ha dejado de comercializarse/estar en servicio, recuperado). 7. El tipo, el número y la fecha de caducidad del certificado expedido por el organismo notificado y nombre o número de identificación de dicho organismo notificado, cuando proceda. 8. Una copia escaneada del certificado a que se refiere el punto 7, cuando proceda. 9. Los Estados miembros donde el sistema de IA esté o se haya introducido



información sobre las "instrucciones de uso electrónicas". No obstante, se contemplan algunas restricciones de esta información al público³⁰.

Desde el punto de vista de la transparencia "interna", el AIA obliga al proveedor que informe al usuario del sistema (art. 13) de "las características, capacidades y limitaciones del funcionamiento del sistema de IA de alto riesgo", y en particular sobre "nivel de precisión, solidez y ciberseguridad" y "circunstancias" que puedan afectar estos niveles o "cualquier circunstancia [...] que pueda dar lugar a riesgos", información sobre "su funcionamiento", "vida útil" o "medidas de mantenimiento y cuidado necesarias"³¹. También en este ámbito de la transparencia "interna" hacia el usuario -no al afectado- se exige poner información a disposición de órganos supervisores. Especialmente, debe también tenerse en cuenta todas las especificaciones que se detallan respecto de la documentación técnica en el articulado y anexos. La misma se detalla en los artículos 11, 18 y especialmente en el Anexo IV del futuro AIA. Esta información debe elaborarla el proveedor desarrollador del sistema IA (de alto riesgo) para bien mantenerla a disposición de los órganos supervisores, bien para facilitarla a los mismos. Por su extensión cabe remitir a dicho Anexo IV, recordando que incluye documentación detallada del sistema de IA, con descripción detallada de los elementos y de su proceso de desarrollo, del seguimiento, el funcionamiento y el control del sistema de IA o de cambios introducidos en el sistema. Se incluye la descripción, cómo interactúa con soportes físicos o software externo, versiones, microprogramas, actualizaciones, instrucciones de uso e instalación, seguimiento, funcionamiento y el control del sistema, sus capacidades y limitaciones de funcionamiento, niveles de precisión, vigilancia humana, sistema de gestión de riesgos, etc. También de cara a los evaluadores y autoridades de supervisión de los sistemas de IA hay que tener en cuenta especialmente las obligaciones de trazabilidad. Así, se impone que se diseñen los sistemas de alto riesgo "con capacidades que permitan registrar automáticamente eventos («archivos de registro») mientras están en funcionamiento" y los sistemas

en el mercado o puesto en servicio o esté disponible en la Unión. 10. Una copia de la declaración UE de conformidad contemplada en el artículo 48."

³⁰ Así, hay reservas respecto de anexo III, puntos 1, 6 y 7 por lo que se refiere a sistemas IA para la aplicación de la ley y la gestión de la migración, el asilo y el control fronterizo.

³¹ "ii) el nivel de precisión, solidez y ciberseguridad mencionado en el artículo 15 con respecto al cual se haya probado y validado el sistema de IA de alto riesgo y que puede esperarse de este, así como las circunstancias conocidas o previsibles que podrían afectar al nivel de precisión, solidez y ciberseguridad esperado; iii) cualquier circunstancia conocida o previsible, asociada a la utilización del sistema de IA de alto riesgo conforme a su finalidad prevista o a un uso indebido razonablemente previsible, que pueda dar lugar a riesgos para la salud y la seguridad o los derechos fundamentales; iv) su funcionamiento en relación con las personas o los grupos de personas en relación con los que se pretenda utilizar el sistema; [...] e) la vida útil prevista del sistema de IA de alto riesgo, así como las medidas de mantenimiento y cuidado necesarias para garantizar el correcto funcionamiento de dicho sistema, también en lo que respecta a la actualización del software.



"garantizarán un nivel de trazabilidad del funcionamiento del sistema de IA durante su ciclo de vida" que permita controlar el funcionamiento y facilitarán el seguimiento posterior (art. 12 sobre "registros")³². De este modo, se generan toda una serie de datos sobre el funcionamiento real del sistema a disposición tanto del propio proveedor por las entidades supervisoras.

9.- Transparencia sobre el funcionamiento y lógica del sistema IA y su incidencia en la toma de decisiones y la transparencia en razón del artículo 22 RGPD

La Unesco en su Recomendación (nº 39) apunta que se dé "información sobre los factores que influyen en una predicción o decisión específicas". Estas exigencias se consideran principios éticos de transparencia y en particular como obligaciones de comunicación y "notificación" "sobre los datos personales utilizados en el proceso de toma de decisiones, "acceso a los factores, la lógica y las técnicas que produjeron el resultado" de un sistema de IA y, en general, "cómo se llega a los procesos de toma de decisiones automatizados y de aprendizaje automático" (Fjeld, *et al.* 2020). Concreta Gutiérrez (2021: 69) respecto del fundamento y lógica del modelo que debe haber transparencia sobre las inferencias, patrones o correlaciones que justifican por qué unos datos de entrada generan unos resultados concretos, explicadas de forma inteligible, sencilla y no técnica.

El ICO (2021 b) para el sector público señala la necesidad de informar sobre "cómo se integra la herramienta en el proceso y qué influencia tiene la herramienta en el proceso de toma de decisiones". A este respecto resulta de interés recordar para todo tipo de organizaciones cómo el HLEG hace referencia a la "transparencia del modelo de negocio" y se debe dar "explicaciones sobre la medida en que el sistema de IA condiciona e influye en el proceso de toma de decisiones de la organización, sobre las decisiones de diseño del sistema y sobre la lógica subyacente a su despliegue"³³. Señala para el sector público Gutiérrez que la transparencia alcanza a determinar en qué ámbitos concretos de la actividad pública se adoptan decisiones basadas total o parcialmente en tratamientos algorítmicos y en qué medida esas decisiones algorítmicas afectan a los derechos y libertades de los ciudadanos, individual o colectivamente.

En el contexto laboral, la *ley rider* -art. 64.4.d) Estatuto de los Trabajadores- reconoce al comité de empresa el "derecho" a "Ser informado por la empresa de los parámetros, reglas e instrucciones en los que se basan los algoritmos o sistemas de

³² En el caso de sistemas biométricos los registros también han de añadir cada uso, la fecha y la hora, base de datos de referencia, datos de entrada e identificación de las personas físicas.

³³ "Además, debería ser posible disponer de explicaciones sobre la medida en que el sistema de IA condiciona e influye en el proceso de toma de decisiones de la organización, sobre las decisiones de diseño del sistema y sobre la lógica subyacente a su despliegue (garantizando así la transparencia del modelo de negocio)." N° 77, p. 22. (HLEG, 2019)



inteligencia artificial que afectan a la toma de decisiones que pueden incidir en las condiciones de trabajo, el acceso y mantenimiento del empleo, incluida la elaboración de perfiles." Y el MTES (2022) ha detallado la información precedente.

Estas referencias están claramente vinculadas al artículo 22 RGPD (interpretado por G29, 2018) o el artículo 11 de la Directiva (UE) 2016/680 para el ámbito judicial, policial y de seguridad reconocen el "derecho" a no ser sometido a decisiones automatizadas. En mayo de 2018 este derecho se ha extendido internacionalmente con la modificación del Convenio 108 del Consejo de Europa que lo reconoce, también como "derecho", en su nuevo artículo 9. 1º³⁴, al igual que los Estándares de Protección de Datos para los Estados Iberoamericanos de 2017 (nº 29). Asimismo, también está reconocido como derecho en el artículo 20 de la reciente Lei nº 13.709, de 14 de agosto de 2018 de Brasil, o más recientemente artículo 20 el "Derecho a no ser objeto de una decisión basada única o parcialmente en valoraciones automatizadas" en la Ley orgánica de protección de datos personales de Ecuador de 2021³⁵.

Como en otros lugares he tenido ocasión de examinar (Cotino, 2019 b) y especialmente Palma (2021, 2022), la clara intención de este "derecho" es que las decisiones automatizadas relevantes, por su sensibilidad o particularidad, tienen ser compensadas con garantías especiales. Y por lo que ahora interesa, particulares deberes de transparencia analizados recientemente por Medina (2022) y por Cerrillo (2022: 62 y ss.). Así, se da la concreta obligación de informar sobre "la existencia de decisiones automatizadas, incluida la elaboración de perfiles, a que se refiere el artículo 22, apartados 1 y 4, y, al menos en tales casos, información significativa sobre la lógica aplicada, así como la importancia y las consecuencias previstas de dicho tratamiento para el interesado" (art. 13. 2º f y 14. 2º g) RGPD). De igual modo, el derecho de acceso permite solicitar dicha información (artículo 15. 1º. h).

³⁴ En concreto, a cada individuo se le reconoce el "derecho" a " no estar sujeto a una decisión que le afecte significativamente basándose únicamente en un procesamiento automatizado de datos sin que se tengan en cuenta sus opiniones", Consejo de Europa - No. [223] Protocolo por el que se modifica el Convenio para la Protección de las Personas con respecto al procesamiento automático de datos personales [Estrasburgo, 10.X.2018] <https://www.coe.int/en/web/conventions/new-treaties>

³⁵ Ecuador Quinto Suplemento del Registro Oficial No.459 , 26 de Mayo 2021, artículo 20 en el que expresamente se señala que el titular podrá: "a. Solicitar al responsable del tratamiento una explicación motivada sobre la decisión tomada por el responsable c. encargado del tratamiento de datos personales; b. Presentar observaciones; c. Solicitar los criterios de valoración sobre el programa automatizado; o, d. Solicitar al responsable información sobre los tipos de datos utilizados y la fuente de la cual han sido obtenidos los mismos; e. Impugnar la decisión ante el responsable o encargado del tratamiento".



En G29 (2018: 35) detalló algo que la información que debe facilitarse³⁶. La AEPD (2020: 23) lo perfiló más para la IA en España³⁷, matizando en todo caso que "dependerá del tipo de componente IA utilizado".

Para el sector público, el M^o Justicia (2021) de Países Bajos también ha detallado la información a facilitar³⁸, así como en Alemania las del Comisionado Federal para la Protección de Datos³⁹. Más adelante se conectan estas exigencias con las obligaciones de motivación en lenguaje natural de las decisiones públicas.

De la síntesis de estos referentes, puede afirmarse que hay que informar detallando el tratamiento y su impacto, explicando por qué es necesario cuál es su legitimación, especificando los datos empleados y su calidad, así como la calidad y robustez del

³⁶ Así: -las categorías de datos que se han utilizado o se utilizarán en la elaboración de perfiles o el proceso de toma de decisiones; -por qué estas categorías se consideran pertinentes; -cómo se elaboran los perfiles utilizados en el proceso de decisiones automatizadas, incluidas las estadísticas utilizadas en el análisis; -por qué este perfil es pertinente para el proceso de decisiones automatizadas; y cómo se utiliza para una decisión relativa al interesado. También se recomienda informar en general respecto de toda decisión automatizada, aunque no sean las protegidas por el artículo 22 y este derecho, *Ibidem*, pág. 27.

³⁷ Lo ha concretado, y afirma "un ejemplo de información que podría tener relevancia de cara al interesado", a saber: "El detalle de los datos empleados para la toma de decisión, más allá de la categoría, y en particular información sobre los plazos de uso de los datos (su antigüedad). La importancia relativa que cada uno de ellos tiene en la toma de decisión. La calidad de los datos de entrenamiento y el tipo de patrones utilizados. Los perfilados realizados y sus implicaciones. Valores de precisión o error según la métrica adecuada para medir la bondad de la inferencia. La existencia o no de supervisión humana cualificada. La referencia a auditorías, especialmente sobre las posibles desviaciones de los resultados de las inferencias, así como la certificación o certificaciones realizadas sobre el sistema de IA. En el caso de sistemas adaptativos o evolutivos, la última auditoría realizada. En el caso de que el sistema IA contenga información de terceros identificables, la prohibición de tratar esa información sin legitimación y de las consecuencias de realizarlo."

³⁸ Se debe informar sobre "qué hace el análisis de datos; por qué realiza análisis de datos (cuál es su finalidad y qué se hace con los resultados); por qué el uso del análisis de datos es proporcionado, y no había mejores alternativas para lograr el objetivo; cuáles son las posibles consecuencias del análisis para los ciudadanos afectados y cómo se han tenido en cuenta las posibles repercusiones en los derechos fundamentales; posible aplicación del *aprendizaje automático* y su explicación; cuál es la base legal para realizar estos análisis; qué fuentes de datos de qué organizaciones se utilizan para este fin, y cuál es su calidad; qué persona de la organización gubernamental es responsable del análisis; el papel de terceros en estos análisis; con qué organizaciones (públicas o privadas), si las hay, se comparten los datos fuente y/o los resultados de los análisis y sobre qué base; qué garantía de calidad se lleva a cabo (qué riesgos se identifican y qué medidas se toman contra ellos, y cómo se comprueban); cómo se produce la intervención humana entre el análisis y una posible decisión, y qué marcos de evaluación existen."

³⁹ "Transparenz der Verwaltung beim Einsatz von Algorithmen für gelebten Grundrechtsschutz unabdingbar", en *36 Konferenz der Informationsfreiheitsbeauftragten in Deutschland*, págs. 3-4, <https://n9.cl/ystpv>



sistema. Finalmente se debe informar también sobre las garantías, supervisión y auditorías.

A lo largo de este estudio y en cada apartado se han precisado estas obligaciones de transparencia e información derivadas del artículo 22 RGPD. Ahora sólo corresponde centrarse la información que detalla el tratamiento de datos realizado y su impacto. De este modo, la transparencia alcanza a los perfilados realizados y sus implicaciones (AEPD, 2020: 23); cómo se elaboran los perfiles utilizados en el proceso de decisiones automatizadas, incluidas las estadísticas utilizadas en el análisis; por qué este perfil es pertinente para el proceso de decisiones automatizadas; y cómo se utiliza para una decisión relativa al interesado (G29, 2018). En el estudio de impacto IA de M^o Interior (2021) de Países Bajos se añade la posible aplicación del *aprendizaje automático* y su explicación; cuáles son las posibles consecuencias del análisis para los ciudadanos afectados y cómo se han tenido en cuenta las posibles repercusiones en los derechos fundamentales. El mencionado Comisionado Federal alemán señala la información sobre "la lógica contenida en el mismo, en particular las fórmulas de cálculo utilizadas, incluida la ponderación de los datos de entrada, información sobre el conocimiento especializado subyacente y la configuración individual por parte del usuario y el alcance de las decisiones basadas en ellas y el posible impacto de los procedimientos."

Toda esta información se ha de sumar a los deberes de transparencia generales sobre cualquier tratamiento de protección de datos (arts. 13 y 14 RGPD) que también han sido expuestos. En el terreno de exigencias en razón de protección de datos, especialmente relevantes a este respecto, la AEPD indica que ha de haber una evaluación sobre "la información sobre los metadatos del componente IA, su lógica y las consecuencias que pueden derivarse de su empleo están accesibles para las partes interesadas junto con los medios o mecanismos disponibles para ejercer sus derechos en caso de objeción respecto de los resultados." Asimismo, ha de controlarse que "existe la documentación suficiente para comprender la lógica del componente IA utilizado y realizar la trazabilidad de su comportamiento respecto a cada conjunto de datos de entrada, el uso que hace de estos y de los datos intermedios, y cómo entrega los datos de salida." (AEPD, 2021: 14-15)

10.- Transparencia sobre las garantías del sistema, estudios de impacto, medidas de mitigación de riesgos, garantías de supervisión

Un apartado también esencial respecto de la información que debe brindarse por transparencia algorítmica es la relativa a las garantías del sistema de IA. Así, en general la Recomendación Unesco (n^o 39) apunta la transparencia "sobre la existencia o no de garantías adecuadas (como medidas de seguridad o de equidad)." En este conjunto de información a facilitar relativo a garantías, se podría incluir la información sobre adecuadas, mitigación de riesgos y peligros, medidas de seguridad o de equidad, análisis de riesgos, sistemas de evaluación o análisis de



impacto, mantenimiento y revisión, información sobre la supervisión humana y auditorías, sistema de gestión de la calidad del sistema.

10.1.- Transparencia de las evaluaciones de impacto y su confidencialidad

Desde la Recomendación UNESCO 2021 se afirma el principio de la transparencia de las evaluaciones de impacto (nº 51 y 53). Así, se afirma que "deberían aplicarse protocolos de transparencia ejecutables, que correspondan al acceso a la información, incluida la información de interés público en poder de entidades privadas" (nº 51) y en concreto, "las evaluaciones del impacto ético deberían ser transparentes y abiertas al público, cuando proceda" (nº 53).

Respecto de las evaluaciones de impacto, el ICO (2021 b) para el sector público señala que se facilite información sobre los estudios de impacto que se hayan realizado (de protección de datos, algorítmica, ética, igualdad) y que se detalle nombre, descripción, fecha, y enlace a la evaluación o un resumen, si está disponible. Ciertamente, la transparencia de los estudios de impacto es una cuestión compleja y normalmente no se trata de información que se ponga a disposición del público. En esta dirección el ELI (2022) en su propuesta de normas afirman que "evaluación de impacto del sistema se llevará a cabo de conformidad con la obligación de respetar y proteger la confidencialidad de los datos y la información relativa o perteneciente a las personas y entidades implicadas en el proceso de evaluación, de acuerdo con las leyes y requisitos pertinentes" (art. 8.1º). Y se señala que se puede considerar reservada más información sobre "datos y la información compartidos, utilizados o adquiridos a lo largo de la evaluación de impacto". No obstante, señala que esta reserva deberá "estar debidamente justificada y sopesada en relación con los requisitos de transparencia de la evaluación de impacto, los intereses de la autoridad de ejecución y del proveedor del sistema." (art. 8.2º)⁴⁰. En cualquier caso, "Esta restricción de la confidencialidad debe sopesarse siempre con el principio de transparencia" y "En cualquier caso, dicha decisión debe estar debidamente justificada".

⁴⁰ Se indica que "los intereses que se protegen están relacionados con los "secretos" en sentido amplio: información personal identificable, privacidad, propiedad intelectual, secretos comerciales (u otros)". El ELI se apoya para ello en Directiva 96/9/CE sobre la protección jurídica de las bases de datos; Directiva 91/250/CEE del Consejo sobre la protección jurídica de los programas de ordenador; Reglamento (UE) 2018/1807 sobre un marco para la libre circulación de datos no personales en la Unión Europea.



10.2.- Transparencia de las medidas y mitigaciones frente a riesgos

Respecto de las medidas y mitigaciones frente a riesgos, también el ICO (2021 b) apunta que se dé "una descripción detallada de los riesgos comunes para su herramienta, que incluya: los nombres de los riesgos comunes (5.5); una descripción de cada riesgo identificado (5.6)⁴¹. Y que se "proporcione una descripción detallada de las acciones que ha tomado para mitigar los riesgos" (5.7). En esta línea, la AEPD (2021: 14-15) en su lista de evaluación incluye si "Ante un comportamiento erróneo por parte del componente IA que pueda ocasionar perjuicios a los interesados, se han previsto mecanismos para minimizar esos perjuicios dar soporte a la comunicación a las partes interesadas y facilitar la comunicación entre todas las partes involucradas en el proceso."

10.3.- Transparencia sobre supervisión humana y auditorías

En este apartado de garantías, cabe incluir la información sobre supervisión humana. Así, siguiendo al ICO, se deben dar detalles específicos sobre cuándo y cómo un ser humano revisa o verifica la decisión automatizada (2.7). Hay que explicar "cómo los humanos tienen la supervisión de la herramienta, incluyendo: cuánta información proporciona la herramienta al tomador de decisiones y cuál es la información (3.2); las decisiones que toman los humanos en el proceso general, incluidas las opciones para los humanos que revisan la herramienta (3.3); capacitación que deben tomar las personas que implementan y usan la herramienta, si esto se aplica a su proyecto (3.4).

Y vinculado a las garantías frente a decisiones algorítmicas, la supervisión expost que implica informar sobre el "proceso de apelación y revisión. Describa cómo permite que los miembros del público revisen o apelen una decisión" (ICO, 2021b: 3.5).

En el caso del artículo 22 RGPD respecto de decisiones IA con tratamientos de datos personales con IA, las autoridades incluyen la obligación de dar información sobre la "existencia o no de supervisión humana cualificada" y "La referencia a auditorías, especialmente sobre las posibles desviaciones de los resultados de las inferencias, así como la certificación o certificaciones realizadas sobre el sistema de IA". En el caso de sistemas adaptativos o evolutivos, la última auditoría realizada." (AEPD, 2020: 23)

⁴¹ Y señala como ejemplo: "daño potencial de la herramienta que se usa de una manera para la que no fue diseñada o construida; creación de resultados sesgados, incluso a través de datos de entrenamiento que no son representativos o contienen sesgos; arbitrariedad y funcionalidad, como la herramienta que proporciona decisiones injustas o incorrectas".



10.4.- Transparencia "interna" sobre garantías del sistema en razón del AIA

En el caso de los sistemas de alto riesgo, el futuro AIA obliga a que el proveedor dé información al usuario (no en general al público) sobre "d) las medidas de vigilancia humana a que se hace referencia en el artículo 14 [relativo a la supervisión humana], incluidas las medidas técnicas establecidas para facilitar la interpretación de la información de salida de los sistemas de IA por parte de los usuarios" (art. 13 AIA). De este modo, los usuarios del sistema deben entender por completo las capacidades y limitaciones del sistema, si hay indicios de anomalías, problemas y saber y ponerles solución. Información que permita a los usuarios interpretar correctamente la información de salida y decidir en su caso, no utilizar el sistema o desestimar, invalidar o revertir la información de salida, así como intervenir accionando un botón específicamente.

También en el contexto de los sistemas de alto riesgo y el futuro AIA hay que destacar especialmente el artículo 17 AIA relativo a la obligación de un sistema de gestión de la calidad del sistema IA. En razón del mismo se requiere que se documenten "las técnicas, los procedimientos y las actuaciones sistemáticas" bien para "el diseño y el control y la verificación del diseño del sistema de IA", bien "en el desarrollo y el control y el aseguramiento de la calidad del sistema". Debe documentarse "una estrategia para el cumplimiento reglamentario, incluido el cumplimiento de los procedimientos de evaluación de la conformidad y de los procedimientos de gestión de las modificaciones de los sistemas de IA de alto riesgo". También "los procedimientos de examen, prueba y validación que se llevarán a cabo antes, durante y después del desarrollo [...] así como la frecuencia"; del mismo modo se han de documentar "las especificaciones técnicas, incluidas las normas, que se aplicarán". Igualmente, deben documentarse "los procedimientos asociados a la notificación de incidentes graves y defectos de funcionamiento"⁴². La regla general es que la "inclusión de los aspectos mencionados [...] será proporcional al tamaño de la organización del proveedor" (art. 17.2º AIA). Se trata de generar todo

⁴² Apartado 1: "a) una estrategia para el cumplimiento reglamentario, incluido el cumplimiento de los procedimientos de evaluación de la conformidad y de los procedimientos de gestión de las modificaciones de los sistemas de IA de alto riesgo; b) las técnicas, los procedimientos y las actuaciones sistemáticas que se utilizarán en el diseño y el control y la verificación del diseño del sistema de IA de alto riesgo; c) las técnicas, los procedimientos y las actuaciones sistemáticas que se utilizarán en el desarrollo y el control y el aseguramiento de la calidad del sistema de IA de alto riesgo; d) los procedimientos de examen, prueba y validación que se llevarán a cabo antes, durante y después del desarrollo del sistema de IA de alto riesgo, así como la frecuencia con que tendrán lugar; [...] g) el sistema de gestión de riesgos que se menciona en el artículo 9; i) los procedimientos asociados a la notificación de incidentes graves y defectos de funcionamiento con arreglo al artículo 62; j) la gestión de la comunicación con las autoridades nacionales competentes; las autoridades competentes, incluidas las sectoriales, que permiten acceder a datos o facilitan el acceso a ellos; los organismos notificados; otros operadores; los clientes, u otras partes interesadas; [...] l) la gestión de los recursos, incluida la seguridad de las medidas relacionadas con el suministro; m) un marco de rendición de cuentas que defina las responsabilidades del personal directivo y de otra índole en relación con todos los aspectos enumerados en este apartado."



un conjunto de datos e información destinado y a disposición esencialmente a los organismos evaluadores y autoridades supervisoras.

11.- Transparencia de los algoritmos públicos y la motivación comprensible de las decisiones públicas con IA

A lo largo del estudio se ha ido haciendo referencias específicas a los deberes de transparencia particulares del sector público. No obstante, procede una reflexión particular con relación a la motivación comprensible de las decisiones públicas con IA. Aunque no es norma jurídica, la Carta de Derechos Digitales de 2021 en su apartado XVIII 6 c) afirma el derecho a "obtener una motivación comprensible en lenguaje natural de las decisiones que se adopten en el entorno digital, con justificación de las normas jurídicas relevantes, tecnología empleada, así como de los criterios de aplicación de las mismas al caso. El interesado tendrá derecho a que se motive o se explique la decisión administrativa cuando esta se separe del criterio propuesto por un sistema automatizado o inteligente" (Cotino, 2022 b).

El artículo 35 Ley 39/2015 obliga a motivar prácticamente todos los actos administrativos, siendo además exigencia constitucional en todo acto limitativo de derechos e intereses (art. 9.3 CE en conexión con el art. 24 CE). Señala Gamero (2021) que no se trata de que existan motivos para dictar el acto, sino de que esos concretos motivos sean transparentes y visibles, que se hagan constar en el propio acto. También puede señalarse que las exigencias de motivación y explicación administrativa pueden quedar en simbiosis, en su caso reforzadas o apuntaladas por las garantías del artículo 22 RGPD proyectadas para el ámbito del sector público en el caso habitual de que supongan un tratamiento de datos personales. Pese a la deficiente técnica del artículo 22 RGPD, siguiendo el Considerando 71 RGPD deben considerarse como garantías mínimas también para el sector público "la información específica al interesado y el derecho a obtener intervención humana, a expresar su punto de vista, a recibir una explicación de la decisión tomada después de tal evaluación y a impugnar la decisión."

La obligación de motivación administrativa obliga a hacer públicas las razones de hecho y de derecho que formalmente fundamentan el acto administrativo. Aunque el acto tenga sus motivos en sí mismos para dictarse, si tales motivos no se expresan en el propio acto, éste carece de motivación. Como recuerda Gamero, la motivación administrativa implica una garantía formal, aunque con la finalidad material no padecer indefensión. La motivación permite el control del acto administrativo, pues el destinatario puede demostrar la ilegalidad de las razones de la Administración. Además, la motivación permite precisar mejor el contenido de la voluntad administrativa y, en general, justifica la actividad administrativa frente a todos (Gamero y Fernández, 2016: 526-527). Bajo esta comprensión, para Gamero no haría falta que una regulación expresa obligue a motivar los actos dictados mediante algoritmos o IA. La Agencia de Derechos fundamentales de la UE (FRA, 2020: 81)



ubica la obligación de motivación algorítmica en el derecho a la buena Administración

Como Martín Delgado apunta también con Red DAIA (2020 b) para la Carta, "el foco de atención de debe ponerse únicamente en el resultado final (hemos de tener en cuenta que en algunos casos no podrá trazarse debidamente cómo se ha llegado a ese output), sino en la forma en la que "piensa" el sistema". Se trata del proceso lógico que conduce a adoptar el acto, la programación del algoritmo, puesta a disposición de los destinatarios del acto, para que puedan si los criterios en cuestión son objetivos y no incurrir en la arbitrariedad proscrita por los artículos 9.1 y 103.2 CE.

Siguiendo la Carta, los sistemas de aprendizaje no pueden incorporarse en el sector público si no se puede obtener una motivación comprensible en lenguaje natural de las decisiones. De lo contrario, no podría verificarse que se dictó con criterios objetivos y sin incurrir en arbitrariedad.

Pues bien, la exigencia de motivación administrativa es variable en razón de la naturaleza de la actuación administrativa, así como de la intensidad de impacto en los derechos e intereses de los interesados. En el contexto de la actuación algorítmica los peligros potenciales exigen que se den garantías de transparencia y explicabilidad, que se agudizan, entre otras variables, en el caso de que haya un tratamiento automatizado de datos. Además, en el caso de las actuaciones administrativas con sistemas IA no parece muy sencilla esta diferenciación entre la motivación formal y los motivos en sí mismos que llevan a la decisión. Se requiere poder comprobar que el funcionamiento del sistema es el correcto, pues puede no serlo pese a que motive formalmente sus decisiones y con aparente corrección que incluso pudiera superar fácilmente los requerimientos de no arbitrariedad y garantías frente a la indefensión. Es por ello que considero que sí que resulta importante una regulación específica para el sector público que además de recordar la exigencia general de motivación formal de los actos adoptados con IA, reconduzca las exigencias de transparencia y explicabilidad algorítmica como elementos necesarios para la garantía de los diversos derechos fundamentales implicados. De este modo, el amplio abanico de exigencias de transparencia que aquí se han recogido en muy buena medida se incorporan como instrumentos de garantía de derechos frente a



los poderes públicos. Gutiérrez⁴³ o Ponce⁴⁴ detallan estas exigencias vinculadas a garantías de defensa o buena Administración.

Añade Gamero (2021) que no puede descartarse que la motivación habrá de incluir también la acreditación de que el sistema de información mediante el que se ha adoptado la decisión haya superado exitosamente la correspondiente y periódica auditoría que verifique su adecuado funcionamiento. Y ello lo ubica también en la motivación siguiendo supuestos de cinemómetros con relación a las multas de velocidad⁴⁵. Se tratará pues de la existencia de informes y dictámenes bien incorporados al acto o que se puedan conocer por el interesado, constando en el expediente o en otras fuentes de información a su alcance.

Por otra parte, considero especialmente plausible la afirmación de la Carta de que "El interesado tendrá derecho a que se motive o se explique la decisión administrativa cuando esta se separe del criterio propuesto por un sistema automatizado o inteligente." Personalmente me congratulo de su inclusión dado que personalmente propuse su inclusión en la fase inicial de la Carta. Bien es cierto que hay que evitar el "sesgo de automatización" (art. 14 AIA relativo a la vigilancia humana, apartado 4 b), esto es, "la posible tendencia a confiar automáticamente o en exceso en la información de salida generada por un sistema de IA de alto riesgo ("sesgo de automatización"), en particular con aquellos sistemas que se utilizan para aportar información o recomendaciones con el fin de que personas físicas adopten una decisión". No obstante, la ciudadanía también merece conocer las razones humanas que llevan a no seguir las propuestas o decisiones automatizadas. Se trata de una extensión necesaria precisamente del control o supervisión humana. A este respecto, también acierta la Carta en el XXIII relativo al ámbito de salud cuando dispone que "4. El empleo de sistemas digitales de asistencia al diagnóstico, y en

⁴³ Y desde OdiselA (2021), Gutiérrez David proponía añadir que "En particular, la motivación contendrá una referencia sucinta al modelo o modelos algorítmicos concretos implementado, a la tipología de datos de entrenamiento y datos de validación empleados, a las variables y parámetros aplicados por el modelo, al porcentaje de aciertos y errores resultantes de la validación del modelo o modelos empleados y a las herramientas complementarias de explicación del modelo si éste se caracteriza por su baja interpretabilidad, y en su caso, a la fase o fases del procedimiento o actividad material concreta en que dicho modelo o modelos han sido aplicados."

⁴⁴ En la consulta para la Carta, desde Red DAIA (2020 b) Ponce propuso que se añadiera que "Dicha motivación deberá ser suficiente para permitir la defensa de los derechos afectados y comprobar la buena administración desarrollada, incluyendo: información sobre el tipo de uso de Inteligencia Artificial, las variables usadas, su ponderación en la decisión final y la manera como deberían haber cambiado para que la persona obtuviera un resultado distinto; el código fuente, siempre que sea legalmente posible, justificando en su caso la imposibilidad jurídica de su inclusión.

⁴⁵ Según recuerda el autor, si no se acreditan las pruebas del buen funcionamiento, no existe prueba de cargo que permita sancionar. Sentencia de la Sala de lo Contencioso-Administrativo del Tribunal superior de Justicia de Castilla y León 1465/2001, de 29 de septiembre; y otras muchas concordantes. E. Gamero Casado, "Necesidad de motivación... cit.



particular de procesos basados en inteligencia artificial no limitará el derecho al libre criterio clínico del personal sanitario." Se trata, precisamente de una proyección del derecho a la revisión humana de las decisiones automatizadas del que habla el artículo 22 RGPD y, tal acción humana en el ámbito administrativo requiere su motivación, máxime cuando un sistema inteligente propone una decisión diferente.

Inspirado expresamente en la Carta, cabe mencionar el muy reciente artículo 23 Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación. Tiene un valor jurídico muy relativo aunque sí simbólico y de carácter promocional. Respecto del uso público de "algoritmos", el mismo hace referencia a la "transparencia y rendición de cuentas, siempre que sea factible técnicamente." También incluye el mandato de priorizar "la transparencia en el diseño y la implementación y la capacidad de interpretación de las decisiones adoptadas por los mismos."

No es momento de detallar las experiencias y regulación comparada, por todos, Gutiérrez (2022), en el mismo libro Palumbo (2022), también Galetta (2022). Pero ciertamente se da un grado bastante más avanzado que la regulación y la práctica españolas en la transparencia de los algoritmos públicos.

Así, por ejemplo, la sentencia del Consejo de Estado italiano núm.8472 de 13 de diciembre de 2019 exige el "*pleno conocimiento del módulo utilizado y de los criterios aplicados*", así como que "este conocimiento del algoritmo debe garantizarse en todos los aspectos: desde sus autores hasta el procedimiento utilizado para su elaboración, el mecanismo de decisión, las prioridades asignadas en el procedimiento de evaluación y toma de decisiones y los datos seleccionados como relevantes (...) a fin de poder verificar que los criterios, condiciones y resultados del procedimiento robótico cumplen con las prescripciones y las finalidades establecidas por la ley o por la propia administración sobre dicho procedimiento y para que queden claras —y, por lo tanto, sean cuestionables— las modalidades y reglas a partir de las que se haya programado" (Cerrillo, 2020 b: 31 y ss.). Las decisiones en Italia ciertamente han sido diversas.

En Francia el artículo 4 de la Ley n° 2016-1321 del 7 de octubre de 2016 para una República digital (*sur la république numérique*) ha regulado los elementos básicos de las decisiones algorítmicas administrativas. Y los mismos han sido desarrollados Decreto del 16 de marzo de 2017.⁴⁶ El Consejo Constitucional ha revisado dicha legislación en su Decisión N° 2018-765 DC, 12 de junio de 2018 (n° 70) y somete a tres condiciones el "mero uso de un algoritmo para basar una decisión administrativa individual", la primera de ellas directamente relacionada con la transparencia

⁴⁶ Decreto N° 2017-330, de 14 de marzo de 2017, relativo a los derechos de las personas que sean objeto de decisiones individuales adoptadas sobre el fundamento de un tratamiento algorítmico. JORF n°0064 du 16 mars 2017



algorítmica: la decisión explícitamente ha de señalar que ha sido adoptada sobre la base de un algoritmo, bajo sanción de nulidad por incumplimiento. Y a petición del interesado deben comunicarse las principales características de la implementación del algoritmo. El Reglamento francés precisa que se ha de detallar el grado y modo de contribución del procesamiento algorítmico a la toma de decisiones; los datos procesados y sus fuentes; los parámetros de tratamiento y, en su caso, su ponderación, aplicados a la situación de la persona interesada y las operaciones realizadas por el tratamiento. Asimismo cabe recordar que el artículo 6 de la Ley obliga a publicar en línea las reglas que definen los principales tratamientos algorítmicos utilizados en sus decisiones⁴⁷. El artículo R311-3-1-1 del mencionado *Decreto* impone que decisión administrativa individual deberá contener una "mención explícita" de la "finalidad perseguida por el tratamiento algorítmico". El Etalab (2022) en agosto ha concretado estas obligaciones y la información de esta mención y pone el ejemplo de cómo cumplirlo: "Esta decisión se tomó sobre la base de un tratamiento algorítmico. Este tratamiento permite *Imencione la finalidad, por ejemplo, calcular el importe del impuesto adeudado*] y cuyas reglas se definen aquí [*Enlace a las reglas que definen el tratamiento algorítmico principal utilizado en la ejecución de las tareas de la administración cuando son la base de decisiones individuales, cf. L. 312-1-3 de la CRPA*]. En aplicación de los artículos R. 311-3-1-1 y R. 311-3-1-2 del código de relaciones entre el público y la administración, usted puede solicitar la comunicación de las normas que definen este tratamiento y su aplicación en su caso a [*Nombre de la administración, señas*]. En caso de no recibir respuesta en el plazo de un mes a partir de la recepción de su solicitud por nuestros servicios, dispone de un plazo de dos meses para remitir el asunto a la Comisión de Acceso a los Documentos Administrativos (CADA), de conformidad con los procedimientos descritos en el sitio web www.cada.fr."

La normativa reconoce el "derecho a obtener la comunicación de las reglas que definen el tratamiento y las principales características de su aplicación, así como las modalidades de ejercicio de este derecho a la comunicación y de revisión, si corresponde, ante la Comisión de Acceso a los Documentos Administrativos". Y el artículo R311-3-1-2 concreta la información a facilitar de "forma inteligible", en tanto que no sean "secretos protegidos por la ley":

"1° El grado y modo de contribución del procesamiento algorítmico a la toma de decisiones; 2° Los datos tratados y sus fuentes; 3° Los parámetros de tratamiento y,

⁴⁷ Las otras dos condiciones son que (2º) la decisión debe ser recurrible administrativamente. Y (3º) el Consejo Constitucional confirma la prohibición específica de decisiones exclusivamente automatizadas si incluyen el uso de datos especialmente protegidos. Es más, el nuevo artículo 10.1º Loi n° 78-17 de 6 de enero de 1978 prohíbe las decisiones semiautomatizadas en el ámbito de la justicia.



en su caso, su ponderación, aplicados a la situación del interesado; 4° Las operaciones realizadas por el tratamiento".

Y como recuerda Gutiérrez, la información que se solicita lo es tanto respecto de sistemas algorítmicos como los propiamente basados en IA.

12.- Transparencia en la contratación pública de IA

En los referentes sobre ética de la IA se afirma como principio en el ámbito de la transparencia a la "contratación pública abierta" (Fjeld, *et al.* 2020). Se sigue en este sentido a *Access Now* (2018: nº 9) cuando recomienda que: "Cuando un organismo gubernamental pretenda adquirir un sistema de IA o sus componentes, la contratación debe realizarse de forma abierta y transparente, de acuerdo con las normas de contratación pública. Esto incluye la publicación del propósito del sistema, los objetivos, los parámetros y otra información para facilitar la comprensión del público. La contratación debe incluir un periodo para los comentarios del público, y los Estados deben ponerse en contacto con los grupos potencialmente afectados cuando sea pertinente para garantizar la oportunidad de hacer aportaciones." El tema de la contratación por el sector público de sistemas de IA bien merece un análisis jurídico detallado (entre otros, Vestri, 2021a o Miranzo, 2020) y no procede ahora entrar en detalle de las obligaciones de transparencia en el ámbito de la contratación pública proyectada a la innovación y en particular a sistemas de IA. No obstante, siguiendo a Gutiérrez (2021: 48), cabe ahora referir que procedería dar información sobre cómo se han adquirido e implementado esos sistemas de decisión algorítmica (mediante desarrollos *in house* o mediante licitación). Igualmente, sería preciso lograr vías para poder hacer un seguimiento de las contrataciones públicas de sistemas algorítmicos o de IA, pues hoy por hoy es una labor casi imposible.

Desde la Red DAIA para la elaboración de la Carta (2020 a y b) hemos subrayado que las entidades públicas han de asumir un papel activo en la gestión contractual referida al uso de la Inteligencia Artificial, valorando con carácter previo qué condiciones deberían respetarse por parte de las entidades privadas contratistas teniendo en cuenta las singularidades del sector público. En concreto, debe reflexionarse sobre cuál debería ser el régimen de transparencia y acceso a la programación y en qué condiciones se podría disponer de la titularidad o, en su caso, de los derechos de uso de los algoritmos. Es por ello que podría haber sido positiva la inclusión que propusimos: "Las garantías de este párrafo habrán de darse cuando el desarrollo de sistemas o aplicaciones se efectúen en el marco de la colaboración y contratación público privada y habrán de garantizarse en los pliegos y condiciones u otros instrumentos."



13.- ¿Y cómo ha de ser la información?

Como se ha adelantado, la disposición de la información depende en muy buena medida de la naturaleza del sujeto al que va destinada y la función concreta que tiene informar.

Resultan de especial interés las Directrices para el sector público neerlandés (M^o Justicia, 2021) cuando se señala la importancia de un acceso global no segmentado o por silos a la información. Así, cabe "garantizar que los equipos tengan pleno acceso y conocimiento de la documentación, las decisiones y el código de los demás. Cuando las decisiones sobre las características, las especificaciones, el diseño, la construcción y las pruebas se reparten entre varios equipos, las diferencias de interpretación pueden pasar desapercibidas y no ser intencionadas durante el traspaso. La transparencia y la explicabilidad están entonces en riesgo.". Estas afirmaciones de todo interés pueden valer tanto para la transparencia interna a especialistas (usuarios, entidades evaluadoras o supervisoras) como para el público, en particular para la comunidad de especialistas.

De cara a la información de los proveedores a los usuarios de sistemas, el AIA señala que los sistemas han de ir acompañados de instrucciones con "información concisa, completa, correcta y clara que sea pertinente, accesible y comprensible para los usuarios" (art. 13. 2^o AIA).

La lista de evaluación del HLEG, en su apartado de "comunicación" tiene en cuenta especialmente la naturaleza de los sujetos y su perspectiva. Así, incluye "si hay procesos que tengan en cuenta las opiniones de los usuarios y que utilicen dichas opiniones para adaptar el sistema"; "Si se informa hacia otras audiencias, hacia terceros o hacia el público en general." Y "Según el caso de uso, ¿ha tenido en cuenta la psicología humana y sus posibles limitaciones, como el riesgo de confusión, el sesgo de confirmación o la fatiga cognitiva?".

Como es conocido especialmente en el ámbito de la protección de datos, la información muchas veces, especialmente la dirigida al público general o afectados por decisiones, debe estar estructurada en capas, de un nivel general más accesible y simple y una capa que permite profundizar en el conocimiento (AEPD, 2018: 5 y ss.). Y ello debe replicarse respecto de la transparencia para el público de la IA y así se aprecia en las mejores prácticas internacionales respecto de la transparencia de la IA. Las Directrices para el sector público de Países Bajos (M^o Justicia: 32) para el caso de información sobre decisiones automatizadas se detalla que "Ante todo, debe ser concisa y transparente, comprensible y de fácil acceso." De ello deriva que la información "está "estratificada", donde uno puede navegar rápidamente a los pasajes relevantes sin tener que desplazarse por toda la declaración de privacidad. "Comprensible" implica que un representante medio de la audiencia a la que se dirige debe ser capaz de entender la información. En este contexto, puede ser útil



comprobar de vez en cuando con la audiencia real si esto es así, por ejemplo mediante un panel de usuarios. "Fácilmente accesible" significa que uno necesita hacer poco esfuerzo para acceder a la información. Incluirlo en una *declaración de privacidad* en el propio sitio web con un enlace claro en la *página de inicio* [...] y además] en un lenguaje claro y sencillo [...] No debe dejar lugar a diferentes interpretaciones y, en cualquier caso, debe aportar claridad sobre el objetivo y la base jurídica de los análisis", con una información "estratificada".

También en el terreno concreto de la IA, los estándares del ICO (2021 b) para el sector público están claramente estructurados en estos dos niveles. De igual modo, desde ahí se indica también que se ha de informar de cómo se puede obtener más información sobre la herramienta o cómo pueden hacer cuestiones, incluidas las opciones fuera de línea y una dirección de correo electrónico de contacto de la organización, el equipo o la persona de contacto responsable.

Las Directrices del M^o Justicia (2021) para el sector público holandés inciden en este apartado y apuntan que "cuando la información retenida sirve para la "explicabilidad", suele adoptar la forma de documentos "normales", como documentos de Word o PDF o presentaciones (digitales)." Y se apunta especialmente la mejor gestión en forma de código fuente, repositorios de código y bases de datos"⁴⁸.

14.- Para concluir: Recapitulación del detalle de la información algorítmica, ¿Quién accederá a ella?

Se inició el estudio recordando que respecto de la transparencia y explicabilidad algorítmicas, los datos, información y conocimiento en los que se sustancian dependen en buena medida de definir para qué se quiere la transparencia y a quién o para quién va dirigida. Sobre estas premisas se ha pretendido detallar *qué transparencia*. Así, el punto de partida y premisa es la "comunicación" general de la existencia del sistema IA, así como en su caso "notificaciones" o informaciones particulares de su uso a los afectados o interesados. De igual modo se han señalado los futuros registros y bases de datos públicas de sistemas de IA de alto riesgo regulados por el AIA. También, y con creciente importancia se han mencionado las obligaciones de difusión activa de información sobre la existencia de sistemas de IA públicos y su grado de incidencia en las decisiones públicas. El futuro ejemplo valenciano con la ley de 2022 puede ser un referente para España y Europa, muy por encima de las escasas garantías normativas estatales en España. La transparencia también ha de ser sobre quiénes son los proveedores de sistemas de IA y quiénes

⁴⁸ "En lo que respecta a la transparencia técnica, es más probable que la información que debe conservarse sea en forma de código (fuente) y datos. Es más probable que esta información se gestione en otros entornos, como *los repositorios de código* y las bases de datos. Hay que prestar especial atención a que esta información esté también en el panorama de los responsables de la gestión de la información para que ésta sea también accesible y cumpla con la Ley de Archivos, entre otras cosas."



los usuarios. De igual forma se han recordado las obligaciones de informar de responsables de protección de datos y otras figuras. Asimismo la transparencia algorítmica lo ha de ser sobre por qué existe el sistema IA y una explicación de que no había mejores alternativas. Ello puede ser de especial relevancia jurídica en algunos contextos en los que el uso mismo de un sistema de IA pueda considerarse restrictivo de derechos. Igualmente, la transparencia supone informar sobre la finalidad prevista, general y concreta del sistema IA y en su caso la prioridad de los mismos. La transparencia e información sobre los datos de entrenamiento del sistema de IA, así como los datos de entrada para su funcionamiento, validación y prueba es también esencial y exigible en razón de diferentes normativas. Es más, de nuevo he insistido en que el afectado o interesado tiene derecho a los datos inferidos sobre él por el sistema de IA, sin perjuicio de que puedan concurrir límites específicos.

Otro grueso de información a disponer en razón de la transparencia algorítmica es relativa a la tipología y especificaciones técnicas del sistema IA. En principio quedaría más próxima a la transparencia "interna" por su naturaleza, si bien, los interesados o afectados por los sistemas así como la comunidad científica y especializada también requieren de diferente información "técnica". De ahí que se detalle la diferente transparencia sobre el tipo de modelos y algoritmos, sus módulos, grado de madurez o sistema de entrenamiento. Del mismo modo se hace referencia al sensible acceso al "código fuente", programación y software del sistema. De manera específica se analiza la transparencia técnica sobre el sistema IA que impone el AIA para el público en general, así como la transparencia "interna" técnica del proveedor para usuarios y para las entidades evaluadoras y autoridades supervisoras.

La transparencia y explicabilidad algorítmicas han venido ligadas hasta la fecha especialmente a la regulación de protección de datos y en concreto al ámbito de las decisiones automatizadas -o de IA- del artículo 22 RGPD y afines. Lamentablemente son exigencias normativas bastante limitadas, si bien este precepto inspira y se proyecta también en otros contextos en forma de exigencias de transparencia o "notificaciones" sobre el fundamento y lógica del sistema IA, como las más recientes del ámbito laboral (*ley rider*) como especialmente para el sector público. Así las cosas se ha descrito en lo posible la información que procede, a partir de los referentes internacionales así como el alto grado de detalle de la AEPD en 2020.

Un bloque bien importante de transparencia algorítmica es la relativa a los datos, información y conocimiento sobre las garantías del sistema de IA, sobre la existencia y contenido de las evaluaciones de impacto y su confidencialidad. En este grupo no es poca la información potencial a facilitar sobre las medidas y mitigaciones frente a riesgos, sobre la supervisión humana del sistema IA así como sobre auditorías. Finalmente en este bloque se ha detallado la numerosa información sobre garantías del sistema exigida por el futuro AIA, de nuevo transparencia "interna" no dirigida al público en general, sino a usuarios, organismos evaluadores y autoridades.



Se ha hecho referencia a las obligaciones de transparencia sobre el fundamento y lógica del sistema IA y su incidencia en la toma de decisiones y la transparencia en razón del artículo 22 RGPD. Pues bien, se ha dedicado un apartado más concreto sobre la transparencia de los algoritmos públicos y la exigencia de una motivación comprensible de las decisiones públicas con IA. Se trata de ámbitos muy conectados respecto de los que he formulado concreciones jurídico normativas que pueden ser relevantes. También en este contexto público se han detallado exigencias de transparencia en la contratación pública de IA. Tras el detalle de qué transparencia se ha concluido este estudio con una breve referencia a cómo se ha de brindar la información en la que se sustancia.

No es posible en este estudio adentrarse en el terreno del ejercicio del derecho de acceso a la información pública en materia de algoritmos. Se trata de una herramienta en manos de la sociedad civil y la ciudadanía en general así como la comunidad investigadora. Cabe seguir, Boix (2022), Gutiérrez (2021), Cerrillo (2020 a), Vestri (2021 b y 2022). También este acceso puede darse a través del derecho de acceso por los afectados de los tratamientos de datos (RGPD), o en razón de los derechos procesales en caso de un conflicto especialmente judicializado. Ahora sólo quiero llamar la atención en que las importantes las obligaciones de transparencia "interna" (del proveedor hacia el usuario o las autoridades) generan documentación e información que puede llegar a ser accesible por el público en general o los interesados en particular. Se hace referencia a la información dirigida a sujetos específicos, como puedan ser los usuarios del sistema IA, organismos evaluadores o autoridades supervisoras. Ello tiene especial interés respecto de la documentación técnica o de gestión de calidad exigida por el futuro AIA. Esta documentación y toda la información que se solicite ha de facilitarse a las autoridades competentes (art. 23 y otros AIA) y a la misma acceden entre otros los organismos notificados, autoridades notificantes y otros sujetos implicados en los procedimientos de evaluación de conformidad. La documentación se ha de mantener a disposición y conservar, así como la documentación relativa a la gestión de calidad (art. 17) durante diez años tras puesta en servicio o introducción en el mercado (art. 50). Su régimen de acceso está regulado en el futuro artículo 64. Si bien sólo está previsto el acceso a esta información por sujetos cualificados, no hay que excluir totalmente el acceso esta información por interesados y afectados, e incluso por la comunidad científica, sociedad civil y el público en general bajo una serie de circunstancias. En muchos casos se tratará de información pública (art. 13 Ley 19/2013, "contenidos [...] cualquiera que sea su formato o soporte [...] elaborados o adquiridos en el ejercicio de sus funciones") a la que potencialmente se puede acceder a través del derecho de acceso a la información pública. Ello será así especialmente en todos los casos de uso público de IA, ya sean desarrollos propios o de desarrollos para el sector público. También hay que tener en cuenta que se trata de información disponible para el ejercicio de funciones públicas por las autoridades supervisoras de IA u otras competencias. Es por ello que en línea de principio no es difícil que quede también



en el ámbito de la información pública del referido artículo 13, sin perjuicio de que puedan concurrir causas de inadmisión (art. 18) o límites aplicables (arts. 14 y 15). Y las autoridades en su caso judiciales también podrán requerir la información cuando se considere preciso, sin perjuicio de las fórmulas de secreto y reserva que puedan concurrir.

Bibliografía

ACCESS NOW. 2018. "Human Rights in the Age of Artificial Intelligence", Access Now, nº 9. <https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf>

AEPD:

-2018. *Guía para el cumplimiento del deber de informar*, <https://www.aepd.es/es/media/guias/guia-modelo-clausula-informativa.pdf>

-2020. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, de febrero, <https://www.aepd.es/media/guias/adecuacion-rgpd-ia.pdf>

-2021. *Requisitos para Auditorías de Tratamientos que incluyan IA*, <https://www.aepd.es/es/media/guias/requisitos-auditorias-tratamientos-incluyan-ia.pdf>

AI WATCH:

-2022 a. *European landscape on the use of Artificial Intelligence by the Public Sector*, (Tangi, L. *et al*) EUR 31088 EN, Publications Office of the European Union, Luxembourg, doi:10.2760/39336, <https://publications.jrc.ec.europa.eu/repository/handle/JRC129301>

-2022 b. *Road to the adoption of artificial intelligence by the public sector: a handbook for policymakers, public administrations and relevant stakeholders*, European Commission, Joint Research Centre, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2760/288757>

BENÍTEZ PALMA, E. .2022. "Mucho ruido y pocas nueces: algunas consideraciones sobre los registros públicos de algoritmos", VESTRI, G. *Disrupción tecnológica en la administración pública retos*, Aranzadi, Cizur, pp. 83-96.

BOIX PALOP, A. 2022. "Transparencia en la utilización de inteligencia artificial por parte de la Administración", *El Cronista del Estado Social y Democrático de Derecho*, N.º. 100, 2022. 90-105.

CERRILLO I MARTÍNEZ, A.:



-2020 a. "La transparencia de los algoritmos que utilizan las administraciones públicas", *Anuario de Transparencia Local*, nº. 3, 2020, pp. 41-78.

-2020 b. ¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?, *European Review of Digital Administration & Law* Volume 1 – Issue 1-2 –:18-36, DOI: 10.4399/97888255389603.

COTINO HUESO, L.:

-2019 a. "Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho" en *Revista Catalana de Derecho Público* nº 58 (junio).

<http://revistes.eapc.gencat.cat/index.php/rcdp/issue/view/n58>

<http://dx.doi.org/10.2436/rcdp.i58.2019.3303>

-2019 b. "Derechos y garantías ante el uso público y privado de inteligencia artificial, robótica y big data", en Bauzá, Marcelo (dir.), *El Derecho de las TIC en Iberoamérica*, Obra Colectiva de FIADI (Federación Iberoamericana de Asociaciones de Derecho e Informática), La Ley- Thompson-Reuters, Montevideo, pp. 917-952,

<http://links.uv.es/BmO8AU7>

-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 RAMIÓ, C. (coord.), *Repensando la administración digital y la innovación pública*, Instituto Nacional de Administración Pública (INAP), Madrid.

<https://links.uv.es/FUW2pz6>

-2022 a. "Nuevo paradigma en la garantías de los derechos fundamentales y una nueva protección de datos frente al impacto social y colectivo de la inteligencia artificial", en COTINO HUESO, L. (editor), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Thompson-Reuters Aranzadi, FIADI, Cizur.

-2022 b. "Derechos ante la Administración digital y la inteligencia artificial (apartado XVIII)" en L. Cotino Hueso, (editor), *La Carta de Derechos Digitales*, Tirant Lo Blanch, Valencia, 2022.

-2022 c. "Transparencia y explicabilidad de la inteligencia artificial y "compañía" (comunicación, interpretabilidad, inteligibilidad, auditabilidad, testabilidad, comprobabilidad, simulabilidad...). Para qué, para quién y cuánta", en COTINO HUESO L. y CASTELLANOS J. (coords.), *Transparencia y explicabilidad de la inteligencia artificial*, Tirant lo Blanch, Valencia.

EUROPEAN LAW INSTITUTE (ELI), 2022. *Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration*, European Law Institute, Universidad de Viena, , <https://acortar.link/46WRX6>



- FJELD, J. *et al.* 2020. "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI", Berkman Klein Center for Internet & Society Research at Harvard University. Enero., <https://dash.harvard.edu/handle/1/42160420>
- FLORIDI, L., *et al.*, "AI4People —An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations", *Minds and Machines* 28(4), 2018, pp. 689-707, <https://doi.org/10.1007/s11023-018-9482-5>
- FRA. 2020. Agencia de los Derechos Fundamentales de la Unión Europea. *Getting the future right. artificial intelligence and fundamental rights*, <https://n9.cl/np9yhe>
- FUNDACIÓN ETICAS. 2022. *The External Audit of the VioGén System*, mayo, <https://eticasfoundation.org/wp-content/uploads/2022/03/ETICAS-FND-The-External-Audit-of-the-VioGen-System.pdf>
- G29 Grupo del Artículo 29. 2018. *Directrices sobre decisiones automatizadas de 6 de febrero de*, <https://www.aepd.es/sites/default/files/2019-12/wp251rev01-es.pdf>
- GALETTA, D-U. 2022. "Derechos y garantías concretas respecto del uso por los poderes públicos de decisiones automatizadas e inteligencia artificial en el procedimiento administrativo", en Cotino Hueso, L. (editor), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Thompson-Reuters Aranzadi, Cizur, pp. 79-93.
- GAMERO CASADO E. 2021. "Necesidad de motivación e invalidez de los actos administrativos sustentados en inteligencia artificial o en algoritmos", *Almacén de Derecho*, 4 febrero.
- GAMERO CASADO E. y FERNÁNDEZ RAMOS, S. 2016. *Manual básico de Derecho Administrativo*, (13ª ed.), Tecnos.
- GUTIÉRREZ DAVID, M.E.:
- 2021. "Administraciones inteligentes y acceso al código fuente y los algoritmos públicos. Conjurando riesgos de cajas negras decisionales", en *Derecom*, nº 31, pp. 19-105, <http://www.derecom.com/derecom/>
 - 2022. "Acceso al código fuente y a los algoritmos de las Administraciones inteligentes. Lecciones a partir de experiencias comparadas", en COTINO HUESO L. y CASTELLANOS J. (coords.), *Transparencia y explicabilidad de la inteligencia artificial*, Tirant lo Blanch, Valencia.



HLEG (Comisión Europea - Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial). 2019. *Directrices éticas para una IA fiable*, <https://n9.cl/q5hyc>

ICO (Information Commissioner Office):

-2021 a). *What do we need to do to ensure lawfulness, fairness, and transparency in AI systems?*, ICO, <https://n9.cl/kz4nm>

-2021 b). *Algorithmic transparency data standard*, ICO, 2021 (última versión julio 2022), <https://www.gov.uk/government/collections/algorithmic-transparency-standard>

ICO y Alan Turing Institute. 2020. *Explaining decisions made with Artificial Intelligence*, <https://n9.cl/jw5ys>

JIMÉNEZ LÓPEZ, J. 2022. "Oscuridad algorítmica en el sector público", en VESTRI, G. *Disrupción tecnológica en la administración pública retos*, Aranzadi, Cizur.

JOBIN, A., IENCA, M. y VAYENA, E. 2019. "The Global Landscape of AI Ethics Guidelines", en *Nature Machine Intelligence* 1, 389-399.

MANTELERO, A. 2022. *Beyond Data. Human Rights, Ethical and Social Impact Assessment in AI*, Information Technology and Law Series, IT&LAW 36.

MEDINA GUERRERO, M. 2022. "El derecho a conocer los algoritmos utilizados en la toma de decisiones. Aproximación desde la perspectiva del derecho fundamental a la protección de datos personales", en *Teoría y realidad constitucional*, nº 49, pp. 141-171, <https://dialnet.unirioja.es/descarga/articulo/8450038.pdf>

MIRANZO DÍAZ, J. 2020. "Inteligencia artificial y contratación pública", en MARTÍN I. (Dir.), *Administración electrónica, transparencia y contratación pública*, INAP, pp. 105-142.

MISURACA, G. 2020. *AI Watch. Artificial Intelligence in public services. Overview of the use and impact of AI in public services in the EU*, Joint Research Centre, Unión Europea. <https://n9.cl/vmdnf>

Mº INTERIOR Y RELACIONES DEL REINO (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties). 2021. *Impact Assessment. Mensenrechten en Algoritmes (Evaluación de impacto. Derechos humanos y algoritmos)*, julio, <https://n9.cl/rnqqd>

Mº JUSTICIA (Ministerie van Justitie en Veiligheid, Rijksoverheid, Gobierno central). 2021. *Richtlijnen voor het toepassen van algoritmen door overheden en publieksvoorlichting over data-analyses (Pautas para la aplicación de algoritmos*



por parte de los gobiernos y educación pública sobre análisis de datos), Directiva (Richtlijn), de 08-03-, <https://acortar.link/6C226N>

MTES (Ministerio de Trabajo y Economía Social). 2022. *Información algorítmica en el ámbito laboral. Guía práctica y herramienta sobre la obligación empresarial de información sobre el uso de algoritmos en el ámbito laboral*, Gobierno de España, Mayo, https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/trabajo14/Documents/2022/100622-Guia_algoritmos.pdf

NIST; SCHWARTZ, R. *et al.* 2022. *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence*, Special Publication 1270 (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, <https://doi.org/10.6028/NIST.SP.1270>

NIST. 2022. *AI Risk Management Framework: Initial Draft*, marzo, <https://www.nist.gov/document/ai-risk-management-framework-initial-draft>

ODISEIA. 2021. (COTINO HUESO, L., coord.), *Documento propuestas mejora Carta Derechos Digitales*, diciembre, <https://links.uv.es/Twqnt2U>

PALMA ORTIGOSA, A.:

-2021. *Régimen jurídico de la toma de decisiones automatizadas y el uso de sistemas de inteligencia artificial en el marco del derecho a la protección de datos personales*, Tesis doctoral Universidad de Valencia.

-2022. *Decisiones automatizadas y protección de datos personales. Especial atención a los sistemas de inteligencia artificial*, Dykinson,

PALUMBO, S. 2022 "Algoritmo y transparencia de la administración pública en el ordenamiento jurídico italiano. Algunas reflexiones críticas a partir del caso "Buona Scuola".

RED DAIA:

-2019 a. *Conclusiones de Toledo de 1 de abril*, en línea <https://bit.ly/2LvW77M>

-2019 b. *Declaración final de Valencia de 24 de octubre*, en línea https://www.dropbox.com/s/zlth1wqon2z8cob/DAIA_Valenciadeclaracion.pdf?dl=0

-2020 a. *Aportación a consulta de derechos digitales de julio, 2020* <https://www.dropbox.com/s/w9xlb64o1zemaca/DAIACONSULTAvenida.pdf?dl=0>



-2020 b. *Carta de Derechos digitales y sector público: propuestas de mejora*. diciembre 2020. <https://www.dropbox.com/s/gsbirh4kcn6v1ko/DAIAcartadic20.pdf?dl=0>

VESTRI. G.:

-2021 a). "Sistemas de inteligencia artificial en la contratación pública: entre códigos fuente y datos abiertos", *Actualidad administrativa*, N° 12.

-2021 b). "La inteligencia artificial ante el desafío de la transparencia algorítmica: Una aproximación desde la perspectiva jurídico-administrativa", *Revista Aragonesa de Administración Pública*, nº 56, Zaragoza, 2021, pp. 368-398

-2022. "Denegación vs. Derecho de acceso al código fuente en los sistemas algorítmicos. Una perspectiva jurídico-administrativa", en COTINO HUESO L. y CASTELLANOS J. (coords.), *Transparencia y explicabilidad de la inteligencia artificial*, Tirant lo Blanch, Valencia.