



VNIVERSITAT DE VALÈNCIA

La implantación de la transparencia algorítmica

Informe para la implantación del registro de algoritmos públicos y el cumplimiento de las obligaciones de publicidad activa de la Ley 1/2022 valenciana de transparencia

Grupo de investigación:

Andrés Boix Palop
Jorge Castellanos Claramunt
Lorenzo Cotino Hueso
Adrián Palma Ortigosa
Alba Soriano Aranz

Coordinación: Lorenzo Cotino Hueso

Informe realizado en el marco del convenio de la Cátedra PAGODA específico para esta actividad y Proyecto de Investigación del Plan Nacional de I+D+I «Registro público de algoritmos» (Ref: PDC2022-133890-I00) 2022-2023.



MINISTERIO
DE CIENCIA
E INNOVACIÓN



Financiado por
la Unión Europea
NextGenerationEU



Financiación del Informe por parte de Presidencia de la Generalitat Valenciana



GENERALITAT
VALENCIANA

Presidencia

CONTENIDO GENERAL

RESUMEN EJECUTIVO.....	6
PRIMERO. CUÁNDO DEBE FACILITARSE INFORMACIÓN DE LOS SISTEMAS.....	20
SEGUNDO. QUÉ INFORMACIÓN HAY QUE FACILITAR DE LOS SISTEMAS ALGORÍTMICOS Y DE INTELIGENCIA ARTIFICIAL EN RAZÓN DE LA LEY VALENCIANA	47
TERCERO. CÓMO GESTIONAR LA GOBERNANZA INTERNA DEL REGISTRO DE ALGORITMOS.....	95
CONCLUSIONES Y RECOMENDACIONES PARA LA PUESTA EN MARCHA DEL SISTEMA DE PUBLICIDAD ACTIVA DE ALGORITMOS EMPLEADOS POR LA GENERALITAT VALENCIANA	104

ÍNDICE

RESUMEN EJECUTIVO.....	6
I. PRESUPUESTOS DE LA CREACIÓN DE REGISTROS Y PUBLICIDAD ACTIVA. CUÁNDO DEBE FACILITARSE INFORMACIÓN DE LOS SISTEMAS.....	8
II. QUÉ INFORMACIÓN HAY QUE FACILITAR DE LOS SISTEMAS ALGORÍTMICOS Y DE INTELIGENCIA ARTIFICIAL EN RAZÓN DE LA LEY VALENCIANA	13
III. GOBERNANZA INTERNA.....	17
PRIMERO. CUÁNDO DEBE FACILITARSE INFORMACIÓN DE LOS SISTEMAS	20
I. UNA APROXIMACIÓN NORMATIVA.....	20
1. <i>Texto de la Ley 1/2022 valenciana de transparencia (art. 16.1.I)</i>	20
2. <i>Carta de Derechos Digitales 2021</i>	22
3. <i>Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación..</i>	23
4. <i>Decreto 203/2021</i>	24
5. <i>Reglamento (UE) 2016/679, de 27 de abril de 2016 (de protección de datos)</i>	25
6. <i>Decreto-ley 2/2023, de 8 de marzo, de medidas urgentes de impulso a la inteligencia artificial en Extremadura</i>	25
7. <i>Ley “Rider”, estatuto de los trabajadores</i>	26
8. <i>Regulación municipal de Barcelona</i>	26
II.- CUANDO SE EMPLEEN “SISTEMAS ALGORÍTMICOS O DE INTELIGENCIA ARTIFICIAL”	27
1. <i>La definición de IA en el futuro Reglamento IA de la UE</i>	29
2. <i>La ley valenciana opta por un concepto amplio e inclusivo de sistemas automatizados, algorítmicos y de inteligencia artificial para el sector público</i>	33
3. <i>En “procedimientos administrativos o la prestación de los servicios públicos”. Los usos y procesos en los que se aplica la inteligencia artificial en la UE y EEUU</i>	34
III. CRITERIOS Y VARIABLES PARA DETERMINAR EL “IMPACTO”, NIVEL DE RIESGO Y RELEVANCIA JURÍDICA DE LOS SISTEMAS ALGORÍTMICOS PÚBLICOS	36
1. <i>El concreto al ámbito de actuación pública y los sistemas inteligencia artificial públicos de “alto riesgo”</i>	37
2. <i>Más allá del futuro RIA. El “impacto” o relevancia y los efectos significativos del sistema inteligencia artificial</i>	39
3. <i>Usos públicos de IA para tomar decisiones individualizadas respecto de personas, decisiones internas administrativas o para la elaboración de políticas y su impacto colectivo</i>	41
4. <i>Usos masivos o perpetuadores de ilegalidad y la necesidad de “recalibrar” las garantías exigibles</i>	42
5. <i>La mayor o menor automatización de la actuación administrativa</i>	42
Niveles de automatización de la actuación administrativa.....	42
Mayor riesgo en decisiones “únicamente”, “íntegramente” automatizadas, sin intervención directa humana o que no sean “accesorias” para la decisión humana	44
6. <i>Una escala de usos públicos de inteligencia artificial más o menos impactantes</i>	45
SEGUNDO. QUÉ INFORMACIÓN HAY QUE FACILITAR DE LOS SISTEMAS ALGORÍTMICOS Y DE INTELIGENCIA ARTIFICIAL EN RAZÓN DE LA LEY VALENCIANA.....	47
I. A MODO DE INTRODUCCIÓN	47
II. CÓMO HA DE SER LA INFORMACIÓN: “LA DESCRIPCIÓN DE MANERA COMPRENSIBLE”. EL NIVEL DE PROFUNDIDAD Y COMPRENSIBILIDAD DE LA INFORMACIÓN QUE SE OFRECE. LA INFORMACIÓN POR CAPAS...	50
III. NORMATIVA DE REFERENCIA SOBRE LA INFORMACIÓN A FACILITAR	53

1. La información básica que exige el Decreto 203/2021	53
2. La información a detallar según el Estatuto de los trabajadores	54
3. Información a proporcionar en el caso del artículo 22 RGPD	56
4. Información básica de la base de datos del futuro Reglamento de IA	56
5. Ayuntamiento de Barcelona.....	58
IV. QUÉ INFORMACIÓN HAY QUE FACILITAR: SOBRE “SU DISEÑO Y FUNCIONAMIENTO”	61
1. Transparencia sobre los datos del sistema de IA (de entrenamiento, de entrada y los inferidos por el sistema)	61
2. Transparencia sobre la tipología y especificaciones técnicas del sistema IA.....	64
Transparencia del tipo de sistema y algoritmos, sus módulos, grado de madurez o sistema de entrenamiento	65
El acceso al “código fuente”	67
V. INFORMACIÓN SOBRE “EL NIVEL DE RIESGO QUE IMPLICAN”: LA EXISTENCIA, FINALIDAD, INCIDENCIA EN LAS DECISIONES PÚBLICAS, LA LÓGICA Y MOTIVACIÓN COMPRENSIBLE	68
1. Información de que existe el sistema IA y de que no había mejores alternativas	68
2. Transparencia sobre la finalidad prevista del sistema IA.....	69
3. 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.....	72
4. Grado de influencia del sistema automatizado en la decisión que se adopte	76
5. Transparencia y motivación comprensible de las decisiones públicas con IA.....	76
VI. INFORMACIÓN SOBRE LAS GARANTÍAS DEL SISTEMA, ESTUDIOS DE IMPACTO, MEDIDAS DE MITIGACIÓN DE RIESGOS, GARANTÍAS DE SUPERVISIÓN	81
Transparencia de las evaluaciones de impacto y su confidencialidad.....	82
Transparencia de las medidas y mitigaciones frente a riesgos.....	83
Transparencia sobre supervisión humana y auditorías	83
VII. INFORMACIÓN SOBRE “EL PUNTO DE CONTACTO AL QUE PODER DIRIGIRSE EN CADA CASO” Y OTRA INFORMACIÓN SOBRE RESPONSABLES, PROVEEDORES Y USUARIOS, CONTRATACIÓN DEL SISTEMA	84
1. Información de responsable de tratamiento de datos y del delegado de protección de datos	85
2. Divulgación del responsable del sistema y aplicación y en su caso, del informe preceptivo realizado en razón del Decreto 220/2014, de 12 de diciembre de Administración Electrónica valenciana.....	85
3. En su caso, identificación de los responsables en razón del Decreto 130/2012, de 24 de agosto, del Consell, por el que se establece la organización de la seguridad de la información de la Generalitat.....	87
4. Debería incluirse información sobre quiénes son los proveedores de sistemas de IA y quiénes los utilizan en la Administración.....	88
5. Además, información sobre transparencia en la contratación pública de IA	89
VIII. LÍMITES.....	90
RESUMEN COMPARADO DE LA INFORMACIÓN PROPORCIONADA EN LOS REGISTROS ALGORÍTMICOS YA ESTABLECIDOS O PROPUESTOS	91

TERCERO. CÓMO GESTIONAR LA GOBERNANZA INTERNA DEL REGISTRO DE ALGORITMOS ..95

I. ÓRGANOS DE GOBERNANZA INTERNA DEL REGISTRO DE ALGORITMOS	95
II. REGISTRO DE ALGORITMOS QUE YA SE ESTÁN UTILIZANDO	96
1. Búsqueda de potenciales algoritmos.....	96
2. Solicitud de información algorítmica	97
3. Evaluación y asignación del riesgo algorítmico.....	98
3.1) Riesgo algorítmico alto	98
3.2) Riesgo algorítmico medio	99
3.3) Riesgo algorítmico bajo	99
4. Registro de algoritmos.....	100

III. REGISTROS DE ALGORITMOS QUE SE INCORPORARÁN EN EL FUTURO	100
1. <i>Principio de coordinación entre órganos administrativos.</i>	100
2. <i>Solicitud de información algorítmica.</i>	101
3. <i>Evaluación y asignación del riesgo</i>	101
4. <i>Registro de algoritmos públicos.</i>	101
IV. GOBERNANZA DEL REGISTRO DE ALGORITMOS PÚBLICOS.....	101
1. <i>Medidas de gobernanza general aplicable a los sistemas automatizados de riesgo alto</i>	102
2.) <i>Medidas de gobernanza general aplicable a los sistemas automatizados de riesgo medio</i>	102
CONCLUSIONES Y RECOMENDACIONES PARA LA PUESTA EN MARCHA DEL SISTEMA DE PUBLICIDAD ACTIVA DE ALGORITMOS EMPLEADOS POR LA GENERALITAT VALENCIANA	104

Resumen ejecutivo

La creciente implantación y empleo de sistemas algorítmicos o de Inteligencia Artificial avanzada que van más allá de la mera automatización de actuaciones administrativas totalmente determinadas y regladas, en la medida en que introduce en el proceso de toma de decisiones por parte de los poderes públicos el uso de tecnologías hasta ahora no habituales y con un indudable potencial de afección a los derechos de los ciudadanos, ha de ser regulado con mucha más profundidad y detalle de lo que se ha realizado hasta la fecha. Uno de los elementos clave para esta regulación es la debida transparencia respecto del uso de estas herramientas, su funcionamiento concreto, qué elementos, información y parámetros tienen en cuenta y en qué medida, de qué manera, para la toma de esas decisiones y hasta qué punto ha de exigirse, así como de qué manera y con qué límites ha de llevarse a cabo por las Administraciones Públicas.

Hasta la fecha, el marco normativo que debía ser tenido en cuenta para el desarrollo, y sus debidas garantías y transparencia, de la toma de decisiones administrativas empleando estos instrumentos no pasaba de las previsiones, necesariamente genéricas, que las leyes generales de procedimiento administrativo o de régimen jurídico han venido disponiendo sobre el encuadre de las decisiones administrativa automatizadas, que en ningún caso preveían, por ejemplo en España (ley 40/2015, de Régimen Jurídico del Sector Público) garantías de transparencia o publicidad activa reforzadas, sino exigencias de calidad de la actuación, imputación de la responsabilidad a un órgano administrativo o responsable, que debía serlo del tratamiento, de la manera en que se realizaba y de la corrección y calidad del mismo, tanto a efectos de dar explicaciones como para centralizar la posible exigencia de responsabilidades. Junto a ello, las normas en materia de protección de datos, tanto en el ámbito estatal como europeo, y especialmente el Reglamento General de Protección de Datos, han venido tradicionalmente estableciendo exigencias sobre la debida transparencia de los tratamientos automatizados de datos que ha de ser en todo caso garantizada para que todo ciudadano pueda tener efectivamente garantizados sus derechos a conocer los tratamientos que le afectan y poder reaccionar frente a ellos. Por esta vía, de hecho, hasta la fecha, es cómo se han podido desarrollar en la práctica los más eficaces mecanismos de garantía para lograr la debida información y transparencia mínima respecto de los tratamientos de esta índole que puedan afectar a los derechos de los ciudadanos.

Ello no obstante, la manifiesta insuficiencia de este marco como debida garantía de los ciudadanos en un contexto de creciente utilización de estas herramientas ha provocado la aparición de muchas propuestas de textos normativos, recomendaciones y aportaciones dogmáticas para la mejora de la situación, tanto en la esfera internacional (con propuestas de Naciones Unidas, de la UNESCO, de colectivos académicos o de expertos en Inteligencia Artificial) como nacional (el Gobierno de España ha propuesto una Carta de Derechos Digitales de valor no normativo que, entre otras cosas, afronta estos problemas e incide en la necesidad de mayor controles, garantías y transparencia), a lo que se une la manifiesta necesidad de adoptar regulaciones de carácter normativo que disciplinen el empleo de la IA tanto por entes públicos como por ciudadanos y empresas de una manera mucho más detallada, completa y rigurosa, línea en la que ya se está avanzando por parte de la Unión Europea, que como es sabido

está desarrollando los trabajos para la aprobación de un reglamento europeo de Inteligencia Artificial.

En paralelo a esa evolución, todos los países de nuestro entorno se han visto obligados a mejorar y avanzar en el desarrollo de medidas de transparencia para los casos en que ya se están utilizando estas herramientas, en ocasiones por medio de la elaboración de reformas legales o normativas o, más habitualmente estableciendo protocolos y guías para su despliegue de tipo administrativo. Porque el hecho de que todo el conjunto de garantías y derechos o de medidas de transparencia garantizadas por un marco normativo completo y desarrollado no esté todavía desarrollado no justifica que no se puedan realizar mejoras concretas en la acción pública en este dominio por medio de medidas de tipo ejecutivo, acompañando una mayor ambición en el despliegue del empleo de la IA con, de manera natural, medidas de esta índole. En estos momentos existen ya numerosos ejemplos de esta evolución en el Derecho comparado.

En España, y en concreto en la Comunitat Valenciana, ha sido hasta la fecha mucho más prudente que en otros países europeos, en parte porque también el avance de la aplicación de IA a los procesos de toma de decisión de los poderes públicos españoles ha sido menor, aunque como es obvio su despliegue también ha obligado a tomar decisiones sobre transparencia algorítmica que, dada la poca exigencia del marco legal vigente, han sido por lo general poco ambiciosas. Esta situación ha cambiado, al menos respecto de la Administración autonómica valenciana, con la aprobación de la Ley 1/2022, de 13 de abril, de la Generalitat, de Transparencia y Buen Gobierno de la Comunitat Valenciana, que en su artículo 16.1 I) ha establecido por primera vez en España la obligación jurídica estricta de que los poderes públicos (en concreto, y de momento, sólo la Generalitat valenciana y resto de administraciones institucionales autonómicas valencianas) deban incluir en sus obligaciones de transparencia activa medidas precisas y concretas de transparencia algorítmica cuando se haga uso de estos instrumentos o de IA para la toma de decisiones, o cuando afecten de manera sustancial en el proceso, referidas al despliegue de políticas o servicios públicos o a la adopción de decisiones administrativas de cualquier tipo.

El presente documento trata, a partir de la referida previsión legal del art. 16.1 I) de la ley valenciana 1/2022, de exponer cómo pueda ser desarrollada esta previsión y, a la luz de las experiencias comparadas más avanzadas, se expone cómo podría y debería realizarse el desarrollo ejecutivo concreto de estas exigencias de transparencia algorítmica, con la vocación de proporcionar una guía y concretas propuestas para la aplicación de estas medidas y desarrollo de la obligación legal. A estos efectos, el presente documento, cuyo resumen ejecutivo se expone a continuación, cuenta además de con esta primera parte, con sendos trabajos sobre las medidas de fondo de transparencia algorítmica que han de ser adoptado, incluyendo los contenidos de la obligación que han de desarrollarse y sus límites, como las medidas organizativas que, de acuerdo con la experiencia comparada y el marco legal vigente en España, parecen más adecuadas para un despliegue exigente y efectivo de la obligación legal, ya vigente.

I. Presupuestos de la creación de registros y publicidad activa. Cuándo debe facilitarse información de los sistemas

La Ley 1/2022, de 13 de abril, de la Generalitat, de Transparencia y Buen Gobierno de la Comunitat Valenciana en su artículo 16, sobre “Información de relevancia jurídica”, dispone que:

1. *Las administraciones públicas del artículo 3.2 deben publicar:*

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.

Procede delimitar el presupuesto de aplicación de la ley, esto es, cuándo deben publicarse por tratarse “sistemas algorítmicos o de inteligencia artificial” y especialmente tener en cuenta los criterios y variables para determinar el impacto, nivel de riesgo y relevancia jurídica de los sistemas algorítmicos públicos.

Ley valenciana se refiere a “sistemas algorítmicos o de inteligencia artificial” y obliga a definir criterios y variables para determinar el impacto, nivel de riesgo y relevancia jurídica de los sistemas algorítmicos públicos

Interesa ahora centrarse en la delimitación del ámbito de aplicación de este precepto, en particular, en qué hayamos de entender por:

- “sistemas algorítmicos o de inteligencia artificial”

- “que tengan impacto en los procedimientos administrativos o la prestación de los servicios públicos”.

Elementos centrales y estructurales de la normativa actualmente existente.

Así, 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).

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 concretas y exigentes son, como ya se ha avanzado, las regulaciones novedosas que ha introducido recientemente la referida ley valenciana, pero también algún otro instrumento normativo estatal o de otras CCAA como son el artículo 23 de la Ley 15/2022 y el Decreto-ley 2/2023, de 8 de marzo, de medidas urgentes de impulso a la inteligencia artificial en Extremadura, inspirados expresamente en el apartado XVIII de la Carta de Derechos Digitales aprobada por el Gobierno de España (que, como ya se ha referido, es un instrumento de *soft-law*, pero de indudable importancia por fijar orientaciones programáticas claras en la materia)

En todo caso, y aunque no esté aún aprobado, han de ser además tenidas en cuenta las medidas en materia de transparencia interna en el empleo de IA y algoritmos (entre proveedor y usuario del sistema) que se prevé y será exigida en razón del el futuro Reglamento de IA de la UE, ya referido.

Otros ejemplos, en este caso a nivel local, de desarrollos que pueden tomarse como referencia son, en Barcelona, el Protocolo de Definición de metodologías de trabajo y protocolos para la implementación de sistemas algorítmicos de 15 de diciembre de Comisión”. En todo caso, tampoco a escala local la implantación de las obligaciones de transparencia es sencilla y muchas veces se ha llevado a cabo sin marco legal previo, debido a la necesidad de ponerla en práctica, como es el caso de experiencias voluntarias como las del City of Helsinki AI Register o el de Ámsterdam . En Francia, por su parte, en esta misma línea, destacan el inventario de algoritmos de la ciudad de Antibes Juan-les-Pins y las listas de algoritmos de Nantes-Métropole y de Pôle Emploi. 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 Barcelona.

Respecto de la delimitación de los “sistemas algorítmicos o de inteligencia artificial”

La obligación de registros y publicación de información activa gira sobre el concepto de “sistemas algorítmicos o de inteligencia artificial”. La ley valenciana opta por un concepto amplio e inclusivo de sistemas automatizados, algorítmicos y de inteligencia artificial para el sector público: en el ámbito de aplicación de la ley valenciana hay que incluir como regla general, a los sistemas automatizados, aunque no fueran de IA.

Y es que todos los sistemas automatizados o algorítmicos potencialmente pueden generar impactos y afecciones a los derechos de las personas u otros bienes e intereses protegibles. Ahora bien, los riesgos serán mayores conforme el sistema sea más complejo y especialmente en razón de la autonomía del sistema propiamente de IA. Y, en particular, los riesgos se dan con los sistemas predictivos, así como los sistemas de autoaprendizaje y aprendizaje profundo por falta de explicabilidad, que será difícil justificar su uso público.

El concepto de sistema algorítmico queda vinculado al riesgo. Los criterios y variables para determinar el impacto, nivel de riesgo y relevancia jurídica de los sistemas algorítmicos públicos

Según se viene proponiendo desde 2022, el concepto de sistema algorítmico que maneja la ley, debe quedar vinculado a la evaluación del riesgo e impacto sobre los derechos, bienes e intereses en juego.

Uso público de sistemas de “alto riesgo” según el futuro RIA

Sin duda alguna, para considerar el impacto y la relevancia en los casos de personalización de servicios e información del sector público, es un buen punto de partida los usos de IA públicos de “alto riesgo” del futuro RIA, siguiendo su artículo 6 combinado con los anexos II y III.

En primer lugar, a partir del el artículo 6.3º RIA en combinación con el Anexo III. Se trata de sistemas IA utilizados *no sólo de manera accesoria o meramente instrumental, sino de manera decisiva o sustantiva* para las finalidades previstas del Anexo III. Se considera de alto riesgo el uso de:

- sistemas de identificación biométrica en el contexto de la seguridad pública¹.
- el uso de sistemas de IA para la gestión y priorización de emergencias y fronteras,
- los usos relacionados con la aplicación policial y judicial de la ley: evaluación de riesgos individuales de delinquir y reincidir, evaluación de perfiles, pruebas y polígrafos, detección estado emocional, análisis de documentación.
- usos de IA pública para determinar la admisión y programas en el ámbito de la educación, evaluación y seguimiento de aprendizaje.
- los sistemas relativos a los servicios de empleo y evaluación de rendimiento.

Adicionalmente, y referido a los aspectos más próximos si cabe a la actividad administrativa y que por ello hemos de tener mucho más presentes:

- los sistemas de evaluación, concesión, revocación de prestaciones y servicios esenciales de asistencia pública.
- los usos en el ámbito de justicia y penitenciario.

Se dice que un tercio de los sistemas de IA públicos serán de alto riesgo, mientras que sólo un 10% de los públicos y privados lo serían.

El segundo conjunto de sistemas de alto riesgo para el RAI es relativo a productos cubiertos por la legislación de armonización de la Unión enumerada en el anexo II o componentes de seguridad de éstos. Especialmente hay que tener cualquier uso público de IA en:

- dispositivos médicos, automóviles, maquinaria, así como el software independiente de dispositivos médicos.
- También los robots con fines de asistencia y cuidado personal o en el sector sanitario.
- De igual modo, cuando la IA se incorpora como componentes de seguridad de productos o sistemas en los sectores de aviación civil, vehículos agrícolas o forestales, equipos marinos, sistema ferroviario, vehículos de motor y sus remolques.

Aunque no se considere de alto riesgo en el RIA, otros usos y sistemas públicos de IA especialmente impactantes como sistemas para la persecución del fraude, blanqueo, impuestos, seguridad social, inspección de trabajo, sanción de tráfico o similares que pasan por ser especialmente conflictivos.

Variables de impacto y riesgo a tener en cuenta

Métricas de impacto y error: Es esencial utilizar métricas de impacto además de métricas de error al evaluar el desempeño de los sistemas de IA. Un pequeño error puede tener un impacto significativo en áreas críticas, como la salud y la seguridad.

Consideración de los efectos en grupos protegidos: Es importante tener en cuenta si los sesgos o errores en los sistemas de IA tienen efectos especialmente perjudiciales para ciertos grupos protegidos.

Impactos jurídicos y significativos: Los sistemas de IA en el sector público pueden tener consecuencias jurídicas o impactos significativos en los individuos.

Sobre la base de ***criterios de artículo 22 del Reglamento General de Protección de Datos (RGPD)***: los sistemas que producen efectos jurídicos o

¹ Sistemas de inteligencia artificial con reconocimiento facial y datos biométricos. Mejor regular bien que prohibir mal", en *El Cronista del Estado Social*, IUSTEL, *monográfico Inteligencia artificial*, nº 100, septiembre-octubre, 68-79

afectan significativamente al individuo se deben considerar cuidadosamente. El Grupo del Artículo 29 (ahora CEPD) considera que: Las decisiones automatizadas en áreas como el crédito, los servicios sanitarios, las oportunidades laborales y el acceso a la educación son consideradas "significativas". Estos criterios, en su caso, deben aplicarse también a los sistemas de IA en el sector público.

Impacto de la personalización 360° de Servicios e Información: Aunque podría parecer que estos sistemas de IA no tienen impactos jurídicos o efectos significativos, es importante tener en cuenta que pueden definir perfiles personales, profesionales, de consumo y de crédito, lo cual puede tener un impacto significativo en la vida de los individuos. Es más, el G29 considera que el perfilado en la publicidad en línea sí que puede ser "significativo".

Individualización de decisiones: Un elemento crucial para valorar el impacto de los sistemas de IA es el grado en que sus decisiones se individualizan y aplican a una persona específica. Por ejemplo, En esta dirección el artículo L311-3-1 de la ley francesa hace referencia a "una decisión individual" y el Etalab, tiene en cuenta que "se adopten decisiones administrativas individuales con respecto a personas físicas o jurídicas, de derecho público o privado, designadas nominativamente". Las decisiones que se individualizan y se toman directamente por el sistema de IA, como la decisión de intervenir médicamente a una persona o de aprobar o no su examen, pueden tener un mayor impacto.

Impacto menor en la mejora de sistemas administrativos: El uso de IA para la mejora general de los sistemas administrativos o para la configuración de los servicios públicos puede tener un impacto menor en comparación con decisiones que se individualizan.

Mayor impacto en decisiones individualizadas:

Importancia del impacto estructural y colectivo: No se puede ignorar el impacto estructural y masivo del uso de estos sistemas en decisiones políticas o normativas generales. La Recomendación de la UNESCO insiste en la necesidad de un enfoque colectivo al evaluar el impacto de los sistemas de IA.

Generalización y masividad: Es fundamental considerar el alcance de la aplicación de los sistemas de IA. Un error o sesgo en un sistema que se utiliza ampliamente puede tener consecuencias mucho más graves que en uno de uso limitado.

Peligro de errores masivos: Un error en un sistema de IA puede replicarse en miles o millones de decisiones si no se controla. Por lo tanto, es crucial monitorizar y corregir los errores para evitar su propagación a gran escala.

Perpetuación de sesgos: Si los errores no se controlan y corrigen, pueden convertirse en parte del "big data" que alimenta a futuros algoritmos, lo que puede cronificar el sesgo o la discriminación. Hay que valorar si los resultados del sistema IA público que se generan pueden reforzar y perpetuar estos errores y sesgos, resultando en decisiones cada vez más perjudiciales para ciertos grupos.

Niveles de mayor o menor automatización e intervención o supervisión humana: Siguiendo los estándares 2016 para el vehículo autónomo, recientemente Roehl ha afirmado seis niveles de automatización de la actuación administrativa:

- "automatización mínima";
- "recuperación y tratamiento de datos";
- cuando el sistema sugiere los "pasos procedimentales a seguir";

- “decisiones asistidas”;
- “decisiones automatizadas”,
- “decisiones autónomas” con sistemas dinámicos de machine learning no supervisado.

El grado de supervisión real humana: puede, por ejemplo, determinar si es aplicable o no el actual artículo 41 Ley 40/2015. Igualmente, las especiales garantías del artículo 22 RGPD sólo se dan respecto de las decisiones “únicamente” automatizadas.

La intervención humana debe ser significativa y llevada a cabo por una persona autorizada y competente para que no se apliquen estas garantías especiales. El estudio de impacto que debe hacerse ha de registrar el grado de intervención humana.

El grado en el que la decisión totalmente automatizada se integra en una decisión, actuación, proceso o procedimiento Administrativo.

Escala de menor a mayor impacto o peligro en razón del grado de incidencia del sistema automatizado en la decisión administrativa:

- los sistemas de uso meramente instrumental (procesadores, traductores, etc.);

- también sería meramente instrumental el uso de sistemas automatizados para actos reglados, siempre que el sistema se limite a seguir la decisión normativa.

- Más riesgos en el uso de IA pública se dan en ámbitos de discrecionalidad técnica,

- y más si cabe en el uso para ámbitos de discrecionalidad política o discrecionalidad técnica no parametrizable.

- También podría añadirse a esta escala entre los niveles de mayor riesgo, las decisiones IA administrativas para la aplicación de conceptos jurídicos indeterminados.

Finalmente, se establece *una escala de mayor a menor impacto y riesgo:*

- Sistemas calificables como de alto riesgo por el futuro RIA.

- Sistemas que identifican o priorizan objetivos para la aplicación de la ley o para realizar inspecciones en el ámbito de infracciones penales, administrativas y persecución de ilicitudes y fraude, cada vez más habituales en ámbitos de seguridad, mercado y competencia, trabajo, salud, cuidado del medio ambiente, etc.

- Sistemas para la adjudicación de contratos, subvenciones, privilegios del gobierno.

- Sistemas IA para personalizar, priorizar o apoyar la prestación de servicios a los ciudadanos en salud, educación, empleo, servicios sociales, etc.

- Sistemas de extracción de información, investigación, recopilación, supervisión y el análisis de datos, para la elaboración de políticas, la toma de decisiones, monitoreo general y análisis de riesgos.

- Sistemas para la gestión de la organización interna, recursos humanos y las adquisiciones y la gestión de los recursos tecnológicos.

- Sistemas para interactuar y comunicarse con el público sobre sus derechos y obligaciones y su participación.

II. Qué información hay que facilitar de los sistemas algorítmicos y de inteligencia artificial en razón de la Ley valenciana

El bloque del informe dedicado a la información de los sistemas algorítmicos empleados por las Administraciones públicas que debería ser facilitada pretende realizar una aproximación y recomendaciones relativas a los datos que deberían incluirse, con diferentes niveles de publicidad y transparencia, en el registro valenciano de sistemas algorítmicos y de IA. Las propuestas realizadas parten tanto del marco normativo vigente y aplicable en España y, en particular, en la Comunidad Valenciana, como de los modelos comparados de registros algorítmicos que, en los últimos años se han establecido en algunos Estados y ciudades europeas.

Este presente bloque se divide en siete apartados. El primer apartado realiza una introducción general al tema tratado en el informe. De esta forma, partiendo de la obligación de relacionar los sistemas algorítmicos empleados por las Administraciones públicas fijado en el art. 16.2.1) de la Ley 1/2022, de 13 de abril, de la Generalitat, de Transparencia y Buen Gobierno de la Comunitat Valenciana, se pasa a justificar las razones que justifican la mayor transparencia y publicidad posibles de las actividades llevadas a cabo por los poderes públicos. En definitiva, una mayor transparencia contribuye a mejorar la calidad democrática de los Estados pues, entre otras cuestiones, facilita el control de y la confianza en la actuación de los poderes públicos.

La parte introductoria se refiere, asimismo, a la necesidad de establecer una mayor nivel de transparencia en determinados casos. En particular, la transparencia debería ser mayor cuanto más complejo sea el sistema y cuanto más intervenga el sistema en la toma de la decisión que afecte a la ciudadanía. En todo caso, tal como se indica ya en la introducción, no siempre resulta deseable una transparencia absoluta pues pueden generarse problemas de privacidad, seguridad, confidencialidad y propiedad intelectual. Por ello, resulta necesario encontrar un equilibrio entre las diferentes finalidades que se pretenden perseguir en el contexto de uso de sistemas de IA.

La segunda parte de esta parte del presente informe se refiere a cómo ha de ser la información, esto es, el nivel de profundidad y comprensibilidad de la información que se ofrece. En este sentido, el presente informa aboga por proporcionar la información por capas, tal como ya se hace por parte del *Information Commissioner's Office* en el Reino Unido o en el Registro de sistemas algorítmicos de los Países Bajos. Así, se parte de la idea de que la información que se hace pública debe estructurarse de una forma que permita su comprensión por personas que no tengan un elevado nivel de especialización en esta clase de tecnologías, sin perjuicio de proporcionar, además, información algo más compleja y con un nivel de detalle mucho mayor cuya comprensión quizás quede limitada a un sector más especializado de la población.

La necesidad de proporcionar la información en estos dos niveles se explica, por una parte, porque la publicación de explicaciones comprensibles del sistema algorítmico, indicando los derechos y posibilidades de defensa de sus intereses de que disponen las personas que pueden resultar afectadas aquel, facilita que las personas interesadas en cualquier procedimiento en el que se emplee un sistema automatizado puedan ejercer una protección efectiva de sus derechos e intereses legítimos. Por otra parte, al proporcionar un mayor nivel de

detalle en la información sobre los sistemas algorítmicos que se publica, se facilita que organizaciones y profesionales externos con suficiente nivel de especialización y conocimiento puedan llevar a cabo un control efectivo sobre los sistemas de inteligencia artificial o, en general, automatizados, empleados por las Administraciones. En este punto, como es evidente, el marco que pueda finalmente establecer para toda la UE el RIA, cuyo proyecto ya es conocido y está siendo objeto de un amplio debate, ha de ser muy tenido en cuenta.

La tercera sección del presente bloque se centra en la Normativa de referencia sobre la información a facilitar. En primer lugar, nos referimos al art. 11.1.i) del Real Decreto 203/2021, por el que se aprueba el Reglamento de actuación y funcionamiento del sector público por medios electrónicos, pues establece que, en la relación de las actuaciones administrativas automatizadas, la indicación de cada una de estas 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”. A continuación se menciona el Estatuto de los Trabajadores ya que esta norma también realiza una referencia, si bien genérica, de la necesidad de publicar información acerca del funcionamiento de los sistemas algorítmicos (art. 64.4.d). Ahora bien, en desarrollo de dicho precepto, el Ministerio de Trabajo y Economía Social ha desarrollado una Guía que, de manera muy detallada, nos indica exactamente la información que debe publicarse en relación con los sistemas algorítmicos que se emplean.

Resulta ineludible, asimismo, referirse a las obligaciones de información derivadas del art. 22 del Reglamento General de Protección de Datos (en adelante, RGPD), el cual limita el uso de sistemas puramente automatizados de toma de decisiones. En relación con este precepto, la Agencia Española de Protección de Datos nos indica, por ejemplo, en la página 24 de su Informe sobre “Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción” cuál podría ser el tipo de información que debe proporcionarse en algunos casos.

Asimismo, también en relación con el marco normativo aplicable en España es necesario que nos refiramos al futuro Reglamento de Inteligencia Artificial, cuya aprobación definitiva se encuentra cada vez más cercana. En la versión actual de esta norma encontramos que, el art. 60 establece que deberá crearse una “Base de datos de la UE para sistemas de IA de alto riesgo independientes” en la que deberán inscribirse algunos sistemas de IA de alto riesgo antes de ponerse en funcionamiento o introducirse en el mercado (art. 51). Con respecto a la información sobre cada sistema que deberá incluirse en la base de datos, es el Anexo VIII de la norma el que detalla la información que debe incluirse en la base de datos.

También en España, dentro del marco normativo catalán local y regional debemos referirnos al registro de algoritmos que pretende poner en marcha el Ayuntamiento de Barcelona y al registro de sistemas automatizados del observatorio de IA en salud de Catalunya. Entre otras cuestiones, resulta de especial interés, en relación con estos dos registros, destacar la incidencia que se hace, por parte del Ayuntamiento de Barcelona, de la necesidad de facilitar “la participación de la sociedad civil, empresas y entidades [...] dando la oportunidad a todas las partes implicadas de debatir la aplicación de los sistemas algorítmicos y de conocer sus resultados”.

La sección cuarta del presente bloque del informe se ocupa de la información que hay que facilitar sobre el diseño y funcionamiento. En primer lugar, nos referimos a la finalidad prevista del sistema de IA ya que, entre otras cuestiones, esta información determinará si el sistema se considera de alto riesgo o no en la aplicación del nuevo Reglamento Europeo de Inteligencia Artificial, teniendo esta consideración un efecto determinante en el marco jurídico que resultará de aplicación y, por tanto, también la información relativa al sistema que deberá ser publicada. En el marco del RGPD, cuando los sistemas automatizados empleen datos personales, debe proporcionarse información sobre la finalidad del tratamiento, asegurando que estos fines son legítimos (arts. 5, 13 y 22).

En el ámbito comparado, hemos estimado conveniente realizar una breve referencia al registro de algoritmos del ICO en el Reino Unido ya que, no solamente se prevé la publicación de información relativa a la finalidad, sino también para qué tipo de usos no se ha pensado el sistema. Asimismo, también se analiza la Ley francesa n° 2016-1321 (*sur la république numérique*), que reformó el Código de relaciones entre el público y la Administración y el principal [informe](#) que concreta la publicidad de la que se debe dotar el uso de algoritmos públicos, que se refieren a la necesidad de publicar los objetivos para los que se emplean estos.

Además de las finalidades para las que se emplea el sistema, también resulta esencial proporcionar información acerca de los datos que utiliza el sistema algorítmico, tanto los empleados para su entrenamiento y funcionamiento, como los datos inferidos por el sistema. La información debe proporcionarse si atendemos a lo dispuesto en el futuro Reglamento Europeo de Inteligencia Artificial (art. 13). Además, se establece en dicho proyecto normativo que la información acerca de los datos de entrenamiento que deberá proporcionarse a las autoridades supervisoras será mucho más detallada con el objetivo de facilitar el control de los sistemas de IA.

Nos referimos asimismo al gran nivel de concreción con el que, tanto el Reino Unido como los Países Bajos y Francia, han regulado la información que debe proporcionarse en relación con los datos que resultan necesarios para el funcionamiento de cada herramienta algorítmica y la escasa regulación en este ámbito existente en España que, sin embargo, se suple parcialmente para aquellos sistemas que emplean datos personales, con la normativa aplicable en este ámbito. Resulta esencial incidir, en este sentido, en que no solo es esencial dotar de cierto nivel de transparencia a los datos que se emplean para el tratamiento y funcionamiento del sistema, sino que también deben publicarse las técnicas empleadas para procesar esos datos, etiquetarlos, depurarlos, etc. Además, resulta conveniente prestar especial atención y concretar la protección jurídica de los datos inferidos.

Con respecto a la transparencia sobre la tipología y especificaciones técnicas del sistema, esta información va especialmente dirigida a las entidades de supervisión y a los usuarios del sistema, esto es, las organizaciones públicas o privadas que emplean la herramienta algorítmica en el marco de sus procesos o procedimientos. Sin embargo, no es descartable que esta información también se haga pública con el objetivo de facilitar un mayor control externo por parte de terceros, como pueden ser personas expertas en el ámbito de la sociedad civil.

También nos referimos a la necesidad de dar información del tipo de modelos y algoritmos, sus módulos, grado de madurez o sistema de

entrenamiento, como ya se exige, por ejemplo, en el Reino Unido, Países Bajos o Francia.

Particularmente controvertida resulta la cuestión de si se debe dar o no pleno acceso al código fuente de los sistemas algorítmicos. En este sentido, la [Recomendación](#) de la UNESCO sobre la ética de la IA prevé la posibilidad de guardar algunas reservas a la publicación completa del código fuente. Ahora bien, debemos tener en cuenta, entre otras cuestiones, que la autoridad de acceso a la información francesa ha calificado como documento administrativo y, por tanto, publicable, el código fuente de los algoritmos públicos.

Un tema absolutamente esencial en relación con el uso de sistemas automatizados por parte de las Administraciones públicas es la necesidad de motivar las decisiones tomadas con base en los análisis efectuados por estas herramientas. La necesidad de motivar gran parte de los actos administrativos provoca la obligación de dotar de un mínimo de transparencia a una parte muy significativa de las decisiones públicas automatizadas. Esto puede suponer un problema en aquellos casos en los que los sistemas empleados sean especialmente complejos. Por ello, resulta necesario requerir un mayor grado de explicabilidad de los algoritmos empleados por las Administraciones públicas en aras a asegurar que las personas interesadas en procedimientos en los que intervenga una herramienta de IA puedan defender, de manera efectiva, sus derechos e intereses legítimos. En esta motivación o explicación proporcionada, debe hacerse constar que el sistema ha superado los correspondientes procesos de evaluación. Asimismo, deben necesariamente motivarse aquellas decisiones en los que una persona decida separarse del criterio fijado por la herramienta algorítmica. En este punto, además, hay que atender a la importancia de las medidas de autorregulación o de autolimitación que puedan ser establecidas por las propias Administraciones públicas para, en paralelo a la extensión del empleo de herramientas algorítmicas y de IA, disciplinar internamente las medidas de control y garantía en materia de transparencia que en cada caso se estimen necesarias

La quinta sección se ocupa de la información que debe facilitarse sobre el nivel de riesgo de las herramientas de IA empleadas por las Administraciones. En este contexto, la información más relevante que debe proporcionarse es la referida a la metodología empleada para identificar los riesgos así como los riesgos identificados y las medidas adoptadas para mitigarlos.

En relación con el riesgo del sistema hay también otros elementos que tienen particular relevancia. Por ejemplo, incidimos en que deben explicarse las razones por las que resulta necesario emplear el sistema de IA y por qué, a pesar de los riesgos que la herramienta concreta que estamos tratando pueda entrañar, no exista una mejor alternativa a introducir esa herramienta en el procedimiento concreto en que se pretende emplear. Además, debe también indicarse el grado de influencia del sistema automatización en la decisión adoptada.

Asimismo, en aquellos contextos en los que el uso de sistemas de IA puede generar especiales riesgos, como puede ser en el ámbito laboral, resulta necesario no solo informar de la existencia de la decisión automatizada o semi-automatizada, sino también de la forma en que dicha decisión se integra en el procedimiento y sus implicaciones para los derechos fundamentales de las personas afectadas, para poder así detectar y defenderse frente a los posibles daños que pueda provocar.

En aras a poder detectar y gestionar los posibles riesgos que puedan surgir del uso de herramientas de IA deben llevarse a cabo evaluaciones de impactos de los sistemas empleados. Resulta particularmente complicado establecer un régimen de transparencia aplicable a las evaluaciones de impacto ya que su publicación puede poner en riesgo otros intereses. Por esta razón, el European Law Institute (ELI) aboga por respetar aquellos datos que puedan ser confidenciales, pero siempre justificando debidamente la ausencia de transparencia en estos casos.

En relación con el nivel de riesgo, también debe publicarse el nivel de supervisión humana al que es sometido el sistema, incluyendo la información que reciben del sistema las personas que intervienen en el procedimiento, la capacidad que tienen para modificar las decisiones adoptadas por el sistema, así como, y esto es especialmente importante, la capacitación requerida y formación impartida a las personas que supervisan el sistema algorítmico.

La sección sexta se ocupa de la información sobre el punto de contacto al que poder dirigirse en cada caso. Este es uno de los requisitos que precisamente establece el art. 16.1.I) de la Ley 1/2022 de la Comunidad Valenciana y que también se contiene en el Decreto 220/2014, de 12 de diciembre, del Consell, por el que se aprueba el Reglamento de Administración Electrónica de la Comunitat Valenciana (arts. 6.2, 25, 78 y 96). Además, en todo caso, la información sobre el punto de contacto nos remite necesariamente a lo dispuesto en el art. 13 del RGPD, el cual obliga a facilitar información sobre el responsable del tratamiento.

La publicación de esta información resulta esencial para asegurar que las personas interesadas en procedimientos en los que intervengan sistemas algorítmicos pueden proteger sus derechos pues, de esta forma, sabrán a quién dirigirse para poder defenderse frente a las decisiones sugeridas o adoptadas por “la máquina”.

Por otra parte, también entendemos que resulta esencial incluir información acerca de quiénes son las organizaciones proveedoras de los sistemas de IA. Esto es algo que enlaza con la sexta sección del presente bloque del informe, pues hemos dedicado ésta a la información sobre transparencia en la contratación pública de IA. Es esencial que la contratación de sistemas algorítmicos se lleve a cabo con la mayor transparencia posible y asegurando en todo caso la adecuación del contratista y del producto contratado.

Finalmente, la sección séptima se refiere a los límites a la información que debe proporcionarse. Aquí inspiramos nuestra propuesta en gran medida en las normas fijadas por el ICO inglés que entiende que puede justificarse la limitación de la información publicada en relación con los sistemas algorítmicos siempre que existan riesgos de ciberseguridad, propiedad intelectual o de que la transparencia absoluta facilite que se “engañe” al sistema. En todo caso, se insiste en la necesidad de justificar, de manera exhaustiva, todas aquellas razones que sirvan de base a la limitación de la transparencia de los sistemas algorítmicos.

III. Gobernanza interna

El despliegue adecuado de las obligaciones de publicidad activa algorítmica previstas en la Ley 1/2022, de 13 de abril exige una estructura interna de

gobernanza que habrá de ser definida y desarrollada por la GVA para poder garantizar un correcto cumplimiento de las mencionadas obligaciones que, además, permitan un desarrollo garantista para los ciudadanos y a la vez eficaz para el despliegue administrativo, que tenga en cuenta todas las previsiones materiales y de contenido ya propuestas. Para ello, se proponen diferentes instrumentos de tipo organizativo y de gestión, que pasamos a detallar.

Registro de algoritmos públicos de la GVA

Las obligaciones de publicidad activa obligan a la GVA a elaborar un registro público de algoritmos empleados por la administración valenciana que permita centralizar en un solo instrumento la información debida en materia de transparencia respecto de todos aquellos algoritmos que presentan un nivel de riesgo lo suficiente alto. A través de este registro se recopilará toda la información legalmente obligatoria de los sistemas automatizados utilizados por la GVA y, además, se podrá mantener un seguimiento actualizado de estos.

Cauces para elaborar el Registro de algoritmos públicos.

Para la elaboración de este registro, se hace necesario estructurar dos procedimientos dependiendo de si la GVA cuenta ya con el algoritmo o pretende implementarlo en un futuro.

- Para los algoritmos que ya se están utilizando por parte de la GVA, se propone llevar a cabo toda una serie de actuaciones que faciliten la recopilación inicial del conjunto potencial de algoritmos presentes.
- Para los futuros algoritmos que se implementarán, consideramos que sería adecuado aprovechar el procedimiento que ya está definido en el Decreto 220/2014 sobre el Reglamento de Administración Electrónica de la GVA para integrar la recopilación de los algoritmos y su registro.

Será necesario obtener cierta información algorítmica tanto de aquellos sistemas que ya se están utilizando como de aquellos que se pretenda su despliegue en un futuro. Esta información ayudará a:

- Evaluar y establecer el riesgo asignado a cada sistema algorítmico.
- Definir el contenido de publicidad activa exigible.

Una vez se haya recopilado esa información algorítmica, se llevará a cabo la evaluación y asignación del riesgo. En función del riesgo asignado se establecerán unas u otras medidas de gobernanza algorítmica. Ello tiene en cuenta la necesidad de un despliegue que combine garantías con eficacia administrativa y que, además, tiene en cuenta los distintos niveles de riesgo que, a partir de los parámetros detallados en el informe, obligan a tratar de manera diferenciada algoritmos y sistemas de IA que puedan tener efectos diferentes sobre los ciudadanos y su esfera de derechos e intereses, en la línea de las distinciones que también han sido realizadas por los instrumentos normativos en preparación en el ámbito de la IA a nivel europeo. A modo de resumen:

- A los sistemas catalogados de alto riesgo se les exigirá las obligaciones de publicidad activa, además de su obligatoria inscripción en el registro de algoritmos públicos.
- Los sistemas catalogados de riesgo medio únicamente se

registrarán en el registro de algoritmos, sin que deban cumplir con todas las obligaciones de publicidad activa, manteniendo en todo caso respecto de ellos un seguimiento sobre los riesgos potenciales.

- Respecto de los sistemas catalogados de riesgo bajo no se propone el despliegue de ninguna medida de tipo administrativo, entendiendo que precisamente por su naturaleza y riesgo bajo no entrarían en el contenido de la obligación legal de transparencia por no depender de ellos, ni afectarlas de modo sustancial, la adopción de decisiones administrativas.

Órganos de Gobernanza interna

Se propone la creación de dos órganos/unidades administrativas para ejecutar las diferentes obligaciones en materia de publicidad activa algorítmica.

- Oficina de registros de algoritmos. (ORA): Sus funciones se centrarán en facilitar los trámites administrativos para la puesta en marcha del registro de algoritmos.
- Consejo del Registro de Algoritmos Públicos de la Generalitat Valenciana. (CRAP): órgano cuya función principal se focalizará en evaluar el riesgo de los algoritmos.

Tanto un órgano como otro se les asignan otra serie de funciones que se detallan más adelante.

Actuaciones de gobernanza interna

Corresponderá a los órganos previamente indicados en colaboración con las unidades/equipos/responsables de los diferentes algoritmos catalogados de alto o riesgo medio desarrollar diferentes actuaciones que permita: el seguimiento adecuado de las medidas de publicidad activa.; seguimiento de los riesgos potenciales que se puedan generar, resolución de consultas presentadas por los responsables de los sistemas, etc.

Primero. Cuándo debe facilitarse información de los sistemas

Lorenzo Cotino Hueso

I. Una aproximación normativa

1. Texto de la Ley 1/2022 valenciana de transparencia (art. 16.1.I)

La Ley 1/2022, de 13 de abril, de la Generalitat, de Transparencia y Buen Gobierno de la Comunitat Valenciana en su artículo 16, sobre “Información de relevancia jurídica”, dispone que:

1. Las administraciones públicas del artículo 3.2 deben publicar:

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.

Interesa ahora centrarse en la delimitación del ámbito de aplicación de este precepto, en particular, en:

- “sistemas algorítmicos o de inteligencia artificial”

- “que tengan impacto en los procedimientos administrativos o la prestación de los servicios públicos”.

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).

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 ley valenciana, así como el artículo 23 Ley 15/2022 y el Decreto-ley 2/2023, de 8 de marzo, de medidas urgentes de impulso a la inteligencia artificial en Extremadura, inspirados expresamente en el apartado XVIII de la Carta de Derechos Digitales.

En todo caso, aunque no está aprobado, cabe tener en cuenta la transparencia interna (entre proveedor y usuario del sistema) exigida en razón del el futuro Reglamento de IA de la UE.

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 inglés², desde Países Bajos (en especial Directrices, 2021³) y desde el Etalab francés⁴. Por su parte, los organismos de normalización técnica NIST (EEUU), ISO, CEN CENELEC (UE) por lo general abordan -e incipientemente- la transparencia interna, no la transparencia dirigida al público en general⁵.

La implantación de esta obligación no es sencilla⁶ y se pretende superar experiencias voluntarias como las del City of Helsinki AI Register⁷ o el de Ámsterdam⁸. En Francia destaca el inventario de algoritmos de la ciudad de Antibes Juan-les-Pins⁹ y las listas de algoritmos de Nantes-Métropole¹⁰ y de Pôle Emploi¹¹. Ciertamente en Reino Unido, Francia o Países Bajos se trata de

² ICO (Information Commissioner Office), *What do we need to do to ensure lawfulness, fairness, and transparency in AI systems?*, ICO, 2021, <https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/guidance-on-ai-and-data-protection/what-do-we-need-to-do-to-ensure-lawfulness-fairness-and-transparency-in-ai-systems/>

Y de especial utilidad los estándares ICO (Information Commissioner Office), *Algorithmic transparency data standard*, ICO, 2021 (última versión julio 2022), <https://www.gov.uk/government/collections/algorithmic-transparency-standard>

También, de especial interés, ICO y Alan Turing Institute, *Explaining decisions made with Artificial Intelligence*, 2020, <https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/explaining-decisions-made-with-ai/>

³ De particular interés los criterios generales afirmados en el documento Directrices para el sector público, Ministerie van Justitie en Veiligheid (Ministerie van Justitie en Veiligheid), Rijksoverheid (Gobierno central), *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 08-03-2021, <https://www.rijksoverheid.nl/documenten/richtlijnen/2021/09/24/richtlijnen-voor-het-toepassen-van-algoritmen-door-overheden-en-publieksvoorlichting-over-data-analyses#:~:text=Rijksoverheid-,Richtlijnen%20voor%20het%20toepassen%20van%20algoritmen,en%20publieksvoorlichting%20over%20data%20analyses.&text=Doel%20van%20de%20richtlijnen%20is,de%20publieksvoorlichting%20daarbij%20door%20overheden>

⁴ Etalab (2022): *Expliquer les algorithmes publics*.

⁵ Son variados los estándares existentes que hacen referencia a la transparencia algorítmica. Para EEUU cabe señalar los primeros trabajos NIST, *AI Risk Management Framework: Initial Draft*, 17 de marzo de 2022, <https://www.nist.gov/document/ai-risk-management-framework-initial-draft> Asimismo, con importante presencia de la transparencia, NIST; SCHWARTZ, R. y otros (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>

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 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.

⁷ <https://ai.hel.fi/en/ai-register/>

⁸ <https://algoritmeregister.amsterdam.nl/en/ai-register/>

⁹ <https://www.antibes-juanlespins.com/administration/acces-aux-documents-administratifs>

¹⁰ https://data.nantesmetropole.fr/pages/algorithmes_nantes_metropole/

¹¹ <https://www.pole-emploi.fr/candidat/algorithmes.html>

acciones de gran profundidad. Aún sin que sea obligatorio normativamente quizá el mejor ejemplo en España lo brinda el ámbito de Barcelona.

2. Carta de Derechos Digitales 2021

Pese a que no se trate de un texto jurídico normativo y, por tanto, vinculante, resulta importante tener en cuenta la Carta de Derechos Digitales, adoptada por el Gobierno español en 2021.

Por cuanto a la transparencia y explicabilidad, en particular cabe tener en cuenta los apartados XXV Derechos ante la inteligencia artificial, XVIII Derechos digitales de la ciudadanía en sus relaciones con las Administraciones públicas¹² y XXIII Derecho a la protección de la salud en el entorno digital.

XXV Derechos ante la inteligencia artificial

1. La inteligencia artificial deberá asegurar un enfoque centrado en la persona y su inalienable dignidad, perseguirá el bien común y asegurará cumplir con el principio de no maleficencia.

2. En el desarrollo y ciclo de vida de los sistemas de inteligencia artificial:

a) Se deberá garantizar el derecho a la no discriminación cualquiera que fuera su origen, causa o naturaleza, en relación con las decisiones, uso de datos y procesos basados en inteligencia artificial.

b) *Se establecerán condiciones de transparencia, auditabilidad, explicabilidad, trazabilidad, supervisión humana y gobernanza. En todo caso, la información facilitada deberá ser accesible y comprensible.*

c) Deberán garantizarse la accesibilidad, usabilidad y fiabilidad.

3. Las personas tienen derecho a solicitar una supervisión e intervención humana y a impugnar las decisiones automatizadas tomadas por sistemas de inteligencia artificial que produzcan efectos en su esfera personal y patrimonial.

XVIII Derechos digitales de la ciudadanía en sus relaciones con las Administraciones públicas

6. Se promoverán los derechos de la ciudadanía en relación con la inteligencia artificial reconocidos en esta Carta en el marco de la actuación administrativa reconociéndose en todo caso los derechos a:

a) Que las decisiones y actividades en el entorno digital respeten los principios de buen gobierno y el derecho a una buena Administración digital, así como los principios éticos que guían el diseño y los usos de la inteligencia artificial.

b) *La transparencia sobre el uso de instrumentos de inteligencia artificial y sobre su funcionamiento y alcance en cada procedimiento concreto y, en particular, acerca de los datos utilizados, su margen de error, su ámbito de aplicación y su carácter decisorio o no decisorio. La ley podrá regular las condiciones de transparencia y el acceso al código fuente, especialmente con objeto de verificar que no produce resultados discriminatorios.*

c) *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*

¹² “Derechos ante la administración digital y la inteligencia artificial (XVIII Y XV)”, en COTINO HUESO, Lorenzo (editor), *La Carta de Derechos Digitales*, Tirant Lo Blanch, Valencia, 2022, pp. 251-284. [acceso](#)

explique la decisión administrativa cuando esta se separe del criterio propuesto por un sistema automatizado o inteligente.

d) Que la adopción de decisiones discrecionales quede reservada a personas, salvo que normativamente se prevea la adopción de decisiones automatizadas con garantías adecuadas.

7. Será necesaria una evaluación de impacto en los derechos digitales en el diseño de los algoritmos en el caso de adopción de decisiones automatizadas o semiautomatizadas.

XXIII Derecho a la protección de la salud en el entorno digital

[...] 4. El empleo de sistemas digitales de asistencia al diagnóstico, y en particular de procesos basados en inteligencia artificial no limitará el derecho al libre criterio clínico del personal sanitario.

5. Los entornos digitales de salud garantizarán, conforme a la legislación sectorial, la autonomía del paciente, la seguridad de la información, la transparencia sobre el uso de algoritmos, la accesibilidad y el pleno respeto de los derechos fundamentales del paciente y en particular su derecho a ser informado o renunciar a la información y a consentir en el tratamiento de sus datos personales con fines de investigación y en la cesión a terceros de tales datos cuando tal consentimiento sea requerido.

3. Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación

Se trata de la primera norma estatal que expresa la proyección de los contenidos de la Carta de Derechos digitales. Pese a que no sean prescriptivas, incluye referencias a la transparencia algorítmica, de especial interés en el ámbito de los sesgos de los algoritmos.

Artículo 23. Inteligencia Artificial y mecanismos de toma de decisión automatizados.

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

2. Las administraciones públicas, en el marco de sus competencias en el ámbito de los algoritmos involucrados en procesos de toma de decisiones, priorizarán la transparencia en el diseño y la implementación y la capacidad de interpretación de las decisiones adoptadas por los mismos.

3. Las administraciones públicas y las empresas promoverán el uso de una Inteligencia Artificial ética, confiable y respetuosa con los derechos fundamentales, siguiendo especialmente las recomendaciones de la Unión Europea en este sentido.

4. Se promoverá un sello de calidad de los algoritmos.

4. Decreto 203/2021

El Decreto 203/2021 da tímidos avances respecto de las actuaciones automatizadas.

, las insuficiencias del artículo 41 Ley 40/2015 son palmarias y el artículo 13 Decreto 203/2021 en modo alguno las atempera. Así, para la AGE será precisa una resolución de autorización de las decisiones íntegramente automatizadas. Dicha resolución, expresando los recursos procedentes, se publicará en la sede de la AGE. El contenido de la misma puede adivinarse en razón del artículo 11. 1º 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.” Asimismo, el artículo 13. 2º menciona genéricamente que la “resolución” “establecerá medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos de las personas interesadas.”

Artículo 11. Contenido y servicios de las sedes electrónicas y sedes asociadas.

1. Toda sede electrónica o sede electrónica asociada dispondrá del siguiente contenido mínimo a disposición de las personas interesadas:

h) Relación actualizada de los servicios, procedimientos y trámites disponibles

i) Relación actualizada de las actuaciones administrativas automatizadas vinculadas a los servicios, procedimientos y trámites descritos en la letra anterior. 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.

Artículo 13. Actuación administrativa automatizada.

1. La tramitación electrónica de una actuación administrativa podrá llevarse a cabo, entre otras formas, de manera automatizada de acuerdo con lo previsto en el artículo 41 de la Ley 40/2015, de 1 de octubre.

2. En el ámbito estatal la determinación de una actuación administrativa como automatizada se autorizará por resolución del titular del órgano administrativo competente por razón de la materia o del órgano ejecutivo competente del organismo o entidad de derecho público, según corresponda, y se publicará en la sede electrónica o sede electrónica asociada. La resolución expresará los recursos que procedan contra la actuación, el órgano administrativo o judicial, en su caso, ante el que hubieran de presentarse y plazo para interponerlos, sin perjuicio de que las personas interesadas puedan ejercitar cualquier otro que estimen oportuno y establecerá medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos de las personas interesadas.

3. En el ámbito de las Entidades Locales, en caso de actuación administrativa automatizada se estará a lo dispuesto en la disposición adicional octava del Real Decreto 128/2018, de 16 de marzo, por el que se regula el régimen jurídico de los funcionarios de Administración Local con habilitación de carácter nacional.

5. Reglamento (UE) 2016/679, de 27 de abril de 2016 (de protección de datos)

Artículo 22

Decisiones individuales automatizadas, incluida la elaboración de perfiles

1. Todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar.

3. En los casos a que se refiere el apartado 2, letras a) y c), el responsable del tratamiento adoptará las medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, como mínimo el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión.

Artículo 13

Información que deberá facilitarse cuando los datos personales se obtengan del interesado

2 [...] f) 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.

Artículo 14

Información que deberá facilitarse cuando los datos personales no se hayan obtenido del interesado

2 [...] g) 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.

6. Decreto-ley 2/2023, de 8 de marzo, de medidas urgentes de impulso a la inteligencia artificial en Extremadura

En 2023 se adoptó el Decreto-ley 2/2023, de 8 de marzo, de medidas urgentes de impulso a la inteligencia artificial en Extremadura. Del mismo procede en su caso destacar sus artículos 11 (Sistemas de inteligencia artificial en la toma de decisiones)¹³ y, en particular, el artículo 12 (Garantías para la

¹³ El mismo no viene a aportar nada en concreto. Artículo 11. Sistemas de inteligencia artificial en la toma de decisiones. 1. La Administración pública autonómica podrá adoptar actos administrativos mediante sistemas de inteligencia artificial en el marco de un procedimiento administrativo, de acuerdo con los Derechos digitales de la ciudadanía en sus relaciones con las Administraciones Públicas, descritos en la Carta de Derechos Digitales del Gobierno de España

utilización de la inteligencia artificial en los procedimientos administrativos). Resulta especialmente interesante por cuanto parte de la premisa de la obligatoriedad de regular expresamente el impacto del uso de sistemas de IA con carácter general en la “prestación de los servicios públicos”. Asimismo, la obligatoriedad de dejar “constancia” de “validaciones” sobre riesgos y cuestiones relevantes sobre los derechos.¹⁴

7. Ley “Rider”, estatuto de los trabajadores

En 2021 se aprobó la Ley 12/2021, de 28 de septiembre, por la que se modifica el texto refundido de la Ley del Estatuto de los Trabajadores, aprobado por el Real Decreto Legislativo 2/2015, de 23 de octubre, para garantizar los derechos laborales de las personas dedicadas al reparto en el ámbito de plataformas digitales. En razón de la misma, el artículo 64. 4º .d) dispone:

Artículo 64. Derechos de información y consulta y competencias.

4. El comité de empresa, con la periodicidad que proceda en cada caso, tendrá derecho a:

d) *Ser informado por la empresa de los parámetros, reglas e instrucciones en los que se basan los algoritmos o sistemas de 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.*

8. Regulación municipal de Barcelona

La *Gaseta Municipal* de Barcelona, el 23 de diciembre de 2022 en “Otros anuncios – Normativa” publicó el “PROTOCOLO de Definición de metodologías de trabajo y protocolos para la implementación de sistemas algorítmicos de 15 de diciembre de Comisión”¹⁵. Este documento es una de las acciones previstas por la [“Medida de gobierno de la estrategia municipal de algoritmos y datos para el impulso ético de la inteligencia artificial”](#) publicada en el 2021 por el Ayuntamiento de Barcelona, y tiene como finalidad crear un protocolo y una metodología de trabajo común en el Ayuntamiento de Barcelona para la

y la Declaración Europea sobre los Derechos y Principios Digitales para la Década Digital (2023/C 23/01).

2. Para ello, además de los requisitos previstos en el artículo 41 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público, se dará la debida publicidad del mecanismo de decisión, de las prioridades asignadas en el procedimiento de evaluación y de la toma de decisiones, así como de todos los datos que puedan impactar en su contenido.

¹⁴ Artículo 12. Garantías para la utilización de la inteligencia artificial en los procedimientos administrativos.

1. Las normas que regulen los procedimientos administrativos harán referencia expresa al impacto del uso de sistemas de inteligencia artificial en la prestación de los servicios públicos que, en su caso, soporten la asistencia en la presentación de solicitudes, declaraciones responsables o comunicaciones, la comprobación o verificación de los requisitos de los interesados, así como la toma de decisiones.

2. En el Inventario de Información Administrativa, se dejará constancia sobre las validaciones realizadas por el órgano responsable de los procedimientos administrativos respecto del proceso lógico diseñado para la realización de los actos, los riesgos que implica y cualesquiera otros aspectos que garanticen los derechos de los interesados.

¹⁵ Original en catalán. <https://bcnroc.ajuntament.barcelona.cat/jspui/handle/11703/128042>

Versión texto url

https://bcnroc.ajuntament.barcelona.cat/jspui/bitstream/11703/128042/3/GM_2022-12-23_Protocol-IA.pdf.txt

contratación pública o la admisión de pruebas piloto. El mismo va acompañado del documento Protocolo “Definición de metodologías de trabajo y protocolos para la implementación de sistemas algorítmicos”¹⁶. Como ahí se afirma, la regulación de actuaciones automatizadas municipales el Ayuntamiento de Barcelona tiene como antecedente el decreto de Comisión de Gobierno por el que se regulan las actuaciones administrativas automatizadas municipales aprobado en sesión de 22 de junio de 2022 y publicado en el BOPB de 23 de agosto de 2022.¹⁷

II.- Cuando se empleen “sistemas algorítmicos o de inteligencia artificial”

El punto de partida de la ley valenciana, el ámbito sobre el que se proyecta la obligación de registros y publicación de información activa gira sobre el concepto de “sistemas algorítmicos o de inteligencia artificial”.

Resulta de utilidad acudir a las definiciones del JCR de la Comisión Europea¹⁸ de los siguientes conceptos: algoritmo, inteligencia artificial, automático/automatización/automatizado, autonomía/autónomo, sistema experto, aprendizaje automático, sistema de IA de autoaprendizaje/aprendizaje autosupervisado¹⁹. Asimismo, es de interés aproximarse a los conceptos manejados en el Protocolo de Barcelona (Anexo I), dada la funcionalidad que tiene.

Algoritmo: Un algoritmo consiste en un conjunto de instrucciones o pasos utilizados para resolver un problema (por ejemplo, no incluye los datos). El algoritmo puede ser abstracto e implementarse en diferentes lenguajes de programación y bibliotecas de software (JCR). Por su parte para el Ayuntamiento de Barcelona:

“Algoritmo, en un sentido muy amplio: desde modelos de regresión y árboles de decisión que pueden realizar predicciones y agilizan procesos, hasta sistemas más complejos, como las redes neuronales y los modelos bayesianos, que funcionan con aprendizaje automático a medida que van realizando cálculos y predicciones avanzadas.”

Inteligencia artificial: Un sistema de IA es un sistema basado en una máquina capaz de influir en el entorno produciendo un resultado (recomendaciones de predicción o decisiones) para un conjunto determinado de objetivos. Utiliza datos y entradas basados en máquinas y/o humanos para (i) percibir entornos reales y/o virtuales; (ii) abstraer estas percepciones en modelos mediante análisis de forma automática (por ejemplo, con aprendizaje automático) o manual; y (iii) utilizar la inferencia del modelo para formular

¹⁶ <https://ajuntament.barcelona.cat/digital/es/blog/protocolo-para-implantar-la-inteligencia-artificial-en-todos-los-servicios-municipales-con>

¹⁷ https://cido.diba.cat/normativa_local/13591771/regulacio-de-les-actuacions-administratives-automatitzades-municipals-ajuntament-de-barcelona

¹⁸ JRC-AI Watch, Estévez Almenzar, M. y otros (2022): *Glossary of human-centric artificial intelligence*, JRC-Unión Europea

¹⁹ algorithm, artificial intelligence, automatic/automation/automated, autonomy/autonomous, expert system, machine learning, self-learning ai system/ self-supervised learning.

opciones de resultados. Los sistemas IA están diseñados para funcionar con distintos niveles de autonomía.²⁰

Para el Ayuntamiento de Barcelona:

Inteligencia artificial: disciplina científica que incluye diversos enfoques y técnicas como el aprendizaje automático (del que derivan el aprendizaje profundo y el aprendizaje por refuerzo), el razonamiento automático (que incluye la planificación, la programación, la representación y el razonamiento de conocimientos, la búsqueda y la optimización), y la robótica (que incluye el control, la percepción, los sensores y accionadores), así como la integración de todas las demás técnicas en sistemas "ciberfísicos". Según la definición del Consejo de Europa: "La IA es entendida como un conjunto de ciencias, teorías y técnicas, cuyo propósito es reproducir mediante una máquina las capacidades cognitivas de un ser humano".

Automático/Automatización/Automatizado: Proceso o sistema que, en determinadas condiciones, funciona sin intervención humana.²¹

Autonomía/Autónomo: Autonomía: Capacidad de actuar en ausencia o con un bajo grado de influencia externa. e Autónomo: Agente dotado de autonomía

Para el Ayuntamiento de Barcelona, Sistemas inteligentes: algoritmos de decisión automatizada, capaces de tomar decisiones sin intervención humana. La única exigencia es que el algoritmo analice el escenario dado a partir de unos datos y tome una decisión en función de ese análisis.

Sistema experto: Programa informático interactivo que formula las mismas preguntas que un experto humano y, a partir de la información que le proporciona el usuario, (intenta) dar la misma respuesta que daría el experto.²²

Aprendizaje automático (*machine learning*): El aprendizaje automático es una rama de la inteligencia artificial (IA) y la informática que se centra en el desarrollo de sistemas capaces de aprender y adaptarse sin seguir instrucciones explícitas, imitando la forma en que aprenden los humanos y mejorando gradualmente su precisión mediante el uso de algoritmos y modelos estadísticos para analizar y extraer conclusiones a partir de patrones en los datos.

Sistema de IA de autoaprendizaje/Aprendizaje autosupervisado: Los sistemas de IA de autoaprendizaje (o aprendizaje autosupervisado) reconocen patrones en los datos de entrenamiento de forma autónoma, sin necesidad de supervisión²³.

²⁰ OECD AI Principles Overview <https://oecd.ai/en/a-principles>

Los sistemas de inteligencia artificial (IA) son sistemas de software (y posiblemente también de hardware) diseñados por humanos que, dado un objetivo complejo, actúan en la dimensión física o digital percibiendo su entorno mediante la adquisición de datos, interpretando los datos estructurados o no estructurados recopilados, razonando sobre el conocimiento o procesando la información, dverdi de estos datos y decidiendo la mejor acción o acciones a tomar para lograr el objetivo dado.

HLEG AI, Assessment List for Trustworthy AI: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=6834

²¹ ISO/IEC DIS 22989(en).T. erms related to Artificial Intelligence. <https://www.iso.org/obp/ui/fr/#iso:std:is-oiec:22989:dis:ed-1:v1:en>

²² Quinn, 1990 (Computational Disciplines) IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. https://standards.ieee.org/contnet/dam/ieee-standards/standards/web/documents/other/eadv2_glossary.p

²³ HLEG AI, Assessment List for Trustworthy AI (ALTAI) https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=6834

Y para el Ayuntamiento de Barcelona, Actuación Administrativa Automatizada: la actuación realizada íntegramente a través de medios electrónicos, en el marco de un procedimiento administrativo, en la que no haya intervenido de forma directa el personal al servicio del Ayuntamiento y que utiliza alguno de los sistemas de firma establecidos en la normativa administrativa de aplicación, sello electrónico o código seguro de verificación.

1. La definición de IA en el futuro Reglamento IA de la UE

No es una tarea sencilla definir lo que es inteligencia artificial. Se han llegado a señalar más de 55 definiciones²⁴, ello puede tener perspectivas muy diversas como a los efectos de investigación, política e institucional, económica y de mercado. Debe tenerse en cuenta, además, que la “inteligencia artificial” atrae mucha inversión y muchos llamados sistemas de IA ciertamente no lo serán más que en el nombre.

Las dificultades para una definición son mayores cuando se trata de proyectar un régimen jurídico a un sistema de IA. Ello tiene objetivos políticos e institucionales claros y, al mismo tiempo se requiere la mayor seguridad jurídica posible. Pero además, en todos los casos, buscar una definición tiene la dificultad de que debe ser adaptativa para los necesarios cambios que la tecnología depara en el futuro.

La Comisión y la Unión Europea han variado en sus conceptos de IA, así Comisión Europea (2018)²⁵, por el Alto Grupo de Expertos de la Comisión²⁶, por el Parlamento Europeo (2020)²⁷. Finalmente, la UE parece decantarse por concepto de la OCDE²⁸ buscando un mayor consenso internacional, sin perjuicio de diversas concreciones.

Así las cosas, la Unión Europea ha seguido una evolución de conceptos por parte de la Comisión Europea en 2018²⁹, por el Alto Grupo de Expertos de la

²⁴ Así, (Samoili et al. 2020) *análisis de 55 definiciones (1955-2019)*, cito por JRC

²⁵ Comunicación de la Comisión al Parlamento Europeo, al Consejo Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones sobre Inteligencia Artificial para Europa, Bruselas, 25.4.2018 COM(2018) 237 final, pág. 1.

²⁶ HLEG (2019): *A definition of AI: main capabilities and disciplines*, Independent High-Level Expert Group On Artificial Intelligence, European Commission, p. 6.

²⁷ Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL)). ANEXO B art. 4).

²⁸ *Recommendation of the Council on Artificial Intelligence*, de 22 de mayo de, OECD/LEGAL/0449, adopted by the OECD Council at Ministerial level.

²⁹ Comunicación de la Comisión al Parlamento Europeo, al Consejo Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones sobre Inteligencia Artificial para Europa, Bruselas, 25.4.2018 COM(2018) 237 final, pág. 1.

Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions on Artificial Intelligence for Europe, Brussels, 25.4.2018 COM(2018) 237 final

“La inteligencia artificial (IA) se refiere a los sistemas que muestran un comportamiento inteligente al analizar su entorno y tomar acciones, con cierto grado de autonomía, para lograr objetivos específicos.

Los sistemas basados en IA pueden basarse puramente en software, actuando en el mundo virtual (por ejemplo, asistentes de voz, software de análisis de imágenes, motores de búsqueda, sistemas de reconocimiento de voz y rostro) o la IA puede integrarse en dispositivos de hardware (por ejemplo, robots avanzados, automóviles autónomos, drones o aplicaciones de Internet de las Cosas).

En este documento ampliamos esta definición para aclarar ciertos aspectos de la IA como disciplina científica y como tecnología, con el objetivo de evitar malentendidos, lograr un conocimiento común compartido de la IA que pueda ser utilizado de manera fructífera también por no expertos en IA, y para proporcionar detalles útiles que se pueden usar en la discusión sobre las pautas de ética de IA y las recomendaciones de políticas de IA.” (traducción automatizada).

Comisión en 2019³⁰, por el Parlamento Europeo en 2020³¹. Finalmente, la UE parece decantarse por concepto de la OCDE en 2019 buscando un mayor consenso internacional³².

³⁰ HLEG, *A definition of AI: main capabilities and disciplines*, Definition developed for the purpose of the AI HLEG's deliverables, Independent High-Level Expert Group On Artificial Intelligence Set Up By The European Commission. Document made public on 8 April 2019, pág. 6

Proponemos utilizar la siguiente definición actualizada de IA:

“Los sistemas de inteligencia artificial (IA) son sistemas de software (y posiblemente también de hardware) diseñados por humanos (*) que, dado un objetivo complejo, actúan en la dimensión física o digital al percibir su entorno a través de la adquisición de datos, interpretando los datos recopilados estructurados o no estructurados, razonar sobre el conocimiento, o procesar la información, derivada de estos datos y decidir la(s) mejor(es) acción(es) a tomar para lograr el objetivo dado. Los sistemas de IA pueden usar reglas simbólicas o aprender un modelo numérico, y también pueden adaptar su comportamiento analizando cómo el entorno se ve afectado por sus acciones anteriores.

Como disciplina científica, la IA incluye varios enfoques y técnicas, como el aprendizaje automático (de los cuales el aprendizaje profundo y el aprendizaje por refuerzo son ejemplos específicos), el razonamiento automático (que incluye planificación, programación, representación y razonamiento del conocimiento, búsqueda y optimización), y robótica (que incluye control, percepción, sensores y actuadores, así como la integración de todas las demás técnicas en sistemas ciberfísicos)”.

y referirse a este documento como una fuente de información adicional para apoyar esta definición.

* Los humanos diseñan sistemas de IA directamente, pero también pueden usar técnicas de IA para optimizar su diseño.

We propose to use the following updated definition of AI:

“Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans (*) that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.

As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems).”

and to refer to this document as a source of additional information to support this definition.

*Humans design AI systems directly, but they may also use AI techniques to optimise their design.

³¹ Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL)). ANEXO B. que establece la

Propuesta de Reglamento Del Parlamento Europeo Y Del Consejo sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas. En concreto:

Artículo 4. Definiciones. A efectos del presente Reglamento, se entenderá por:

a) «inteligencia artificial», un sistema basado en programas informáticos o incorporado en dispositivos físicos que manifiesta un comportamiento inteligente al ser capaz, entre otras cosas, de recopilar y tratar datos, analizar e interpretar su entorno y pasar a la acción, con cierto grado de autonomía, con el fin de alcanzar objetivos específicos;”

Según se afirma, es una adaptación de la definición que figura en la Comunicación de la Comisión Europea COM(2018) 237 final, de 25.4.2018, p. 1.

En aquel caso, determinó en la letra b) «alto riesgo», riesgo significativo, derivado del desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas, de causar lesiones o daños a las personas o a la sociedad, vulnerando los derechos fundamentales y las normas de seguridad establecidas en el Derecho de la Unión, teniendo en

El futuro Reglamento busca una definición única de la IA con suficiente claridad y precisión, que dé seguridad jurídica y que pueda resistir el paso del tiempo. Asimismo, se pretende que sea funcional y lo más tecnológicamente neutra posible. De igual modo, se pretende una definición que resista al paso del tiempo lo mejor posible dada la dinamicidad tecnológica y del mercado. Para ello se faculta a la Comisión para que en el futuro adopte actos delegados.

La propuesta inicial de la Comisión en el RIA incluía una definición (art. 3. 1º) que, a su vez, remitía al anexo I. Sin embargo, en las últimas versiones se ha optado por incluir la definición sólo en el art. 3. 1º: “un sistema diseñado para funcionar con un cierto nivel de autonomía y que, basándose en datos e insumos proporcionados por máquinas y/o personas, infiere cómo lograr un conjunto determinado de objetivos definidos por el ser humano utilizando enfoques basados en el aprendizaje de máquinas y/o en la lógica y el conocimiento, y produce resultados generados por el sistema, como contenidos (sistemas de IA generativos), predicciones, recomendaciones o decisiones, que influyen en los entornos con los que interactúa el sistema de IA.” Así, los humanos definen el conjunto de objetivos, no la forma de alcanzarlos y el sistema de IA encuentra un conjunto adecuado de pasos para lograr estos objetivos.

El elemento distintivo de la IA respecto de otros desarrollos de software y programación es esencialmente la capacidad de “inferir la manera de lograr un conjunto determinado de objetivos definidos por el ser humano mediante el aprendizaje, el razonamiento o la modelización” para así generar contenidos (como texto, vídeo o imágenes), así como predicciones, recomendaciones o decisiones. Así, los humanos definen el conjunto de objetivos, no la forma de

cuenta su uso o finalidad específicos, el sector en el que se desarrollan, despliegan o usan y la gravedad de las lesiones o daños que cabe esperar que se produzcan;”

³² OECD, *Recommendation of the Council on Artificial Intelligence*, de 22 de mayo de 2019, OECD/LEGAL/0449, adopted by the OECD Council at Ministerial level on 22 May 2019, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>

ACUERDA que, a los efectos de la presente Recomendación, los siguientes términos deben entenderse como sigue:

–Sistema de IA : Un sistema de IA es un sistema basado en una máquina que puede, para un conjunto determinado de objetivos definidos por humanos, hacer predicciones, recomendaciones o decisiones que influyan en entornos reales o virtuales. Los sistemas de IA están diseñados para operar con diferentes niveles de autonomía.

–Ciclo de vida del sistema de IA : Las fases del ciclo de vida del sistema de IA incluyen: i) 'diseño, datos y modelos'; que es una secuencia dependiente del contexto que abarca la planificación y el diseño, la recopilación y el procesamiento de datos, así como la construcción de modelos; ii) 'verificación y validación'; iii) 'despliegue'; y iv) 'operación y seguimiento'. Estas fases a menudo tienen lugar de manera iterativa y no son necesariamente secuenciales. La decisión de retirar un sistema de IA de la operación puede ocurrir en cualquier momento durante la fase de operación y monitoreo.

–Conocimiento de IA : El conocimiento de IA se refiere a las habilidades y recursos, como datos, códigos, algoritmos, modelos, investigación, conocimientos, programas de capacitación, gobernanza, procesos y mejores prácticas, necesarios para comprender y participar en el ciclo de vida del sistema de IA.

–Actores de IA : los actores de IA son aquellos que desempeñan un papel activo en el ciclo de vida del sistema de IA, incluidas las organizaciones y las personas que implementan u operan IA.

–Partes interesadas : Las partes interesadas abarcan todas las organizaciones e individuos involucrados o afectados por los sistemas de IA, directa o indirectamente. Los actores de IA son un subconjunto de las partes interesadas.

alcanzarlos y el sistema de IA encuentra un conjunto adecuado de pasos para lograr estos objetivos³³

Hay que partir de que *no todo sistema informático que permiten procesos o decisiones automatizados son automáticamente IA, aunque puedan razonar (b) o modelar matemáticamente (c)*. Y si el sistema de decisiones automatizadas que no es un Sistema de IA se utiliza para un caso de uso y finalidades de alto riesgo (anexo III), quedará también fuera de la aplicación del AIA.

La Comisión Europea³⁴ estima que sólo el 10% de las unidades de IA estarán sujetas a los requisitos reglamentarios (es decir, las identificadas como de "alto riesgo"). Y que el coste total de cumplimiento para la industria mundial de la IA se estima que oscila entre 1.600 millones de euros y 3.300 millones de euros en 2025.

2. La ley valenciana opta por un concepto amplio e inclusivo de sistemas automatizados, algorítmicos y de inteligencia artificial para el sector público

No se pretende aquí definir lo que es inteligencia artificial sino, más bien, superar en cierto modo un concepto restrictivo para abordar un fenómeno que guarda muchos elementos comunes. Se han llegado a señalar más de 55 definiciones (JRC- Samoili y otros, 2021) de IA. Las dificultades para una definición son mayores cuando el concepto conlleva proyectar todo un régimen jurídico a un sistema de IA, como en el caso de los sistemas de IA -de alto riesgo- en el.

A los efectos de abordar con homogéneamente buena parte de los problemas que se generan, la ley no se ha vinculado a los sistemas públicos que utilicen IA sino que gire sobre el uso público de algoritmos, esto es, sistemas informáticos que integren fórmulas más o menos complejas y las apliquen a los datos. En consecuencia, en el ámbito de aplicación de la ley valenciana hay que incluir como regla general, a los sistemas automatizados, aunque no fueran de IA.

Y es que todos los sistemas automatizados o algorítmicos potencialmente pueden generar impactos y afecciones a los derechos de las personas u otros bienes e intereses protegibles. Ahora bien, los riesgos serán mayores conforme el sistema sea más complejo y especialmente en razón de la autonomía del sistema propiamente de IA. Y, en particular, los riesgos se dan con los sistemas predictivos, así como los sistemas de autoaprendizaje y aprendizaje profundo por falta de explicabilidad, que será difícil justificar su uso público.

Esta opción va en la misma que la de la aplicación de la *ley rider*, esto es, mencionada obligación del artículo 64.4.d) Estatuto de los trabajadores. Como su Guía de aplicación señala:

“la obligación de información a la representación legal de la plantilla no exige que la decisión sea íntegramente automatizada, incluyéndose también las decisiones semi-automatizadas con intervención humana.

³³ JRC.

³⁴ Comisión Europea, Andrea Renda (project leader), *Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe. Final Report (D5)*, abril 2021. <https://op.europa.eu/es/publication-detail/-/publication/55538b70-a638-11eb-9585-01aa75ed71a1>

El concepto de algoritmo usado por el artículo 64.4.d) ET no exige que la toma de decisión se realice sin intervención humana, sino que dichos algoritmos «afecten» a la toma de decisiones que «puedan incidir» sobre la persona trabajadora. Así, aunque se use el algoritmo como simple apoyo a la toma de decisiones por parte de la empresa, se aplicará el artículo 64.4.d) ET. Dicho en otras palabras, aunque el algoritmo no sea determinante para la decisión final tomada sobre la persona trabajadora, su mero uso, implica el nacimiento de los derechos de información de la representación legal de la plantilla.”³⁵

3. En “procedimientos administrativos o la prestación de los servicios públicos”. Los usos y procesos en los que se aplica la inteligencia artificial en la UE y EEUU

Como se ha señalado al inicio, el presupuesto de aplicación de la ley valenciana es respecto de los “sistemas algorítmicos o de inteligencia artificial que tengan impacto en los procedimientos administrativos o la prestación de los servicios públicos”.

La voluntad es extraordinariamente clara: la aplicación no se ciñe ni al artículo 41 Ley 40/2015, ni al artículo 22 RGPD, es decir, no se limita ni a las decisiones sólo automatizadas, realizadas íntegramente sin intervención humana, ni al criterio de particular impacto en la persona concreta.

Sin perjuicio de la precisión general que se ha hecho sobre el uso inclusivo que aquí se hace de la expresión “sistemas IA”, hay que destacar de inicio el mejor estudio existente sobre casos de uso IA públicos, si en 2020³⁶ fue de 240 casos de uso, son más de 600 en 2022³⁷ para la UE, entre los tipos de IA que se emplean en el sector público.

El uso de IA en la UE más habitual (JRC, Tangi, 2022, 39) es el de Servicios Públicos Generales (30%), después de Asuntos Económicos (18%), Salud (15%) y Orden Público y Seguridad (14%). Es similar al análisis difundido en 2020. Se echan de menos los usos en Protección Social (9%) y Protección Ambiental (4%). Hay ámbitos más habituales de subcontratación en los que no se detecta el uso de IA, como usos culturales o recreativos.

En el concepto de Servicios Públicos Generales se incluyen *chatbots* y asistentes virtuales tanto con el exterior como para acelerar procesos internos. También se dan usos para notificaciones, monitoreo, reconocimiento de varios tipos de espacios públicos por cámaras, micrófonos u otros sensores. Igualmente, gestión, detección y manejo de información errónea; clasificación, almacenamiento y búsqueda de documentos, videos o discursos grabados con extracción automática de metadatos e información. Igualmente aquí se incluyen varios tipos de detección de anomalías de datos o fraudes potenciales.

³⁵ Ministerio de Trabajo y Economía Social, *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 2022, https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/trabajo14/Documents/2022/1006_22-Guia_algoritmos.pdf

³⁶ JRC, 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.

³⁷ JRC, Tangi, L. y otros (2022): *AI Watch European landscape...cit.*

Desde el punto de vista del tipo de procesos o aplicaciones³⁸ y cómo se integra el uso de IA dentro del sector público, en el caso de la UE la mayor parte de casos se califican como Servicios Públicos y Compromiso con un 36%. Ahí se incluyen servicios o actividades de comunicación hacia ciudadanos y empresas. La mayoría están relacionados con la mejora del servicio, tanto a través de Personalización 13% como Integración 9%.

Por detrás con un 26% está el uso para el cumplimiento (*enforcement*): reconocimiento inteligente (por ejemplo, sistemas biométricos, videovigilancia y detección de objetos) con un 9%. También con un 9% servicios de y cumplimiento predictivo IA (por ejemplo, para identificar y clasificar cantidades sustanciales de datos históricos para determinar personas o lugares en riesgo). Ya por debajo, apoyos de procesos de inspección o auditorías de Apoyo (con 2%),

Con un 22% está el uso para la investigación de Análisis, Seguimiento e investigación normativa. Se incluyen técnicas para hacer predicciones (10%), análisis de información (7%) e uso para implementar políticas (6%).

El uso para la Gestión estrictamente Interna cuenta con un 16 %: Procesos de Apoyo financiero, recursos humanos, gestión de materiales, con un 9 %; Gestión Interna Procesos Primarios, como la evaluación de elegibilidad de un beneficio social, con un 8%. Sólo un 2% de usos son para la adjudicación de ayudas.

En algunos casos se ha cruzado el tipo de técnica de IA con las finalidades administrativas para la que se usa³⁹. A este respecto se deriva que los sistemas de reconocimiento inteligente son muy populares para fines de seguridad y orden público (para descubrir el uso de teléfonos móviles en vehículos en Bélgica; declaraciones de valores erróneos en el servicio de aduanas en Francia; la predicción de accidentes de tráfico en carreteras de los Países Bajos, etc.). Los sistemas de personalización de servicios abundan en Servicios Públicos Generales, Salud y Protección Social, en muchos casos *chatbots* y sistemas de recomendación para salud, viajes o transporte. Las aplicaciones de Predicción y planificación e Integración de sistemas se usan más para predicciones en sectores económicos (por ejemplo, sistema de predicción del suministro requerido de camiones en Francia y el monitoreo inteligente de puentes en los Países Bajos).

En la educación lo más utilizado es el análisis de información y la predicción y planificación (en Portugal para planificar el flujo de estudiantes del sistema educativo, o la detección de las causas del abandono escolar prematuro en Suecia). En el sector ambiental se acude a sistemas para la predicción y planificación, la implementación de políticas de monitoreo y el reconocimiento digital (como el sistema para identificar especies de árboles forestales, tala de bosques y altura de los bosques en Estonia, o el sistema italiano que predice lluvia y sus impacto en la tierra).

Para EEUU se apuntan en general diversos usos públicos⁴⁰

-hacer cumplir los mandatos regulatorios (*enforcement*) centrados en la eficiencia del mercado, la seguridad en el lugar de trabajo, el cuidado de la salud y la protección del medio ambiente.

³⁸ *Ibidem*, p. 42.

³⁹ *Ídem*.

⁴⁰ Freeman Engstrom, D. y otros (2020): "Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies. Report submitted to the Administrative conference of the United States", *NYU School of Law, Public Law Research Paper No. 20-54*, 122 págs, pp. 6-7.

-Adjudicar subvenciones, ayudas o privilegios del gobierno, desde beneficios por discapacidad hasta derechos de propiedad intelectual.

-Monitoreo y análisis de riesgos para la salud y seguridad pública.

-Extraer información útil de los flujos de datos masivos del gobierno, desde las quejas de los consumidores hasta los patrones climáticos, y

-comunicarse con el público sobre sus derechos y obligaciones como beneficiarios de asistencia social, contribuyentes, solicitantes de asilo y dueños de negocios.

Por cuanto a la a la tecnología IA empleada⁴¹ un 58% se categorizan como aprendizaje automático (*machine learning*), esto es, con la capacidad de aprender, decidir, predecir, adaptarse y reaccionar automáticamente a los cambios. Así sucede en el uso de IA para detección de fraude, mejora de la calidad de los documentos, predicciones basadas en los datos disponibles, automatización de tareas repetitivas con capacidad de adaptación. Se señala también que un 30% de uso son sistemas que infieren hechos a partir del conocimiento representado en varias formas de información y datos. Se trata de técnicas de razonamiento automatizado (enfoques lógicos/basados en el conocimiento, motores de inferencia y deducción, razonamiento simbólico, sistema experto, etc.). Estos sistemas se utilizan para describir el proceso de razonamiento sobre los datos y para brindar soluciones basadas en un conjunto de reglas simbólicas. Por ejemplo, el proyecto CityFlows (<https://cityflows-project.eu/>) , donde la IA se utiliza para automatizar el análisis de flujos de multitudes en grandes espacios públicos en ciudades como Ámsterdam, Milán y Barcelona. Estas dos tecnologías más utilizadas se combinan con procesamiento del lenguaje natural o la visión artificial.

Asímismo, un 26% son tecnologías de IA de Planificación y Programación, se trata del diseño y ejecución de un conjunto de acciones para llevar a cabo alguna actividad, realizadas por agentes inteligentes, robots autónomos y vehículos no tripulados. Los ejemplos son herramientas de planificación y gestión utilizadas en el sector público para impuestos, recursos, empleo, atención médica, energía, materiales y muchos más.

III. Criterios y variables para determinar el “impacto”, nivel de riesgo y relevancia jurídica de los sistemas algorítmicos públicos

Según se viene proponiendo desde 2022, el concepto de sistema algorítmico que maneja la ley, debe quedar vinculado a la evaluación del riesgo e impacto sobre los derechos, bienes e intereses en juego. Aunque en principio todo uso público que incorpore fórmulas, así como también cualquier actuación automatizada, la decisión sobre la aplicación de la ley y la intensidad de las garantías girará en razón del referido impacto.

En este punto hay que llamar la atención de que en Barcelona con claridad se excluye la aplicación de sus protocolos a cualquier sistema algorítmico, sino sólo a los de alto riesgo y en su caso se decidirá orgánicamente si se aplicaría a aplicaciones de menor riesgo.⁴²

⁴¹ Ibídem, pp. 36-37.

⁴² Así en “Protocolo”, *Gasetà*, cit. Así, se afirma: “En ningún caso se trata de un protocolo de aplicación general a cualquier sistema algorítmico, sino que se desplegará de forma

La propia “concepción del servicio” queda condicionada al riesgo. Así, se reserva a las aplicaciones con un riesgo más elevado toda la proyección de las garantías y la participación de expertos y afectados en el ciclo de aplicación de un sistema algorítmico en el Ayuntamiento de Barcelona⁴³.

1. El concreto al ámbito de actuación pública y los sistemas inteligencia artificial públicos de “alto riesgo”

Sin duda alguna, para considerar el impacto y la relevancia en los casos de personalización de servicios e información del sector público, habrá que analizar el concreto al ámbito de actuación, acceso a servicios o beneficios del que se trate. Para valorar el impacto hay que tener especialmente en cuenta la finalidad del uso o propósito del sistema IA.

Para ello son un buen punto de partida los usos de IA públicos de “alto riesgo” del futuro RIA, siguiendo su artículo 6 combinado con los anexos II y III. Como señala el futuro RIA (Considerando 27), “la calificación «de alto riesgo» debe limitarse a aquellos sistemas de IA que tengan consecuencias perjudiciales importantes para la salud, la seguridad y los derechos fundamentales de las personas de la Unión, y dicha limitación reduce al mínimo cualquier posible restricción del comercio internacional, si la hubiera.” La consideración de “alto riesgo” que ya se barajó en el Libro blanco IA de la UE 2020 y en la propuesta del Parlamento UE del mismo año, se ha concretado y mucho y ha pasado a ser el elemento esencial a determinar por cuanto implica la sujeción a todo un régimen de cumplimiento y conjunto de obligaciones.

El futuro RIA sigue un sistema dual de determinación de cuándo un sistema de IA es de alto riesgo. El primer conjunto de sistemas de alto riesgo es relativo a productos cubiertos por la legislación de armonización de la Unión enumerada en el anexo II o componentes de seguridad de éstos. Especialmente hay que tener en cuenta el uso de IA en dispositivos médicos, automóviles, maquinaria, así como el software independiente de dispositivos médicos. También los robots con fines de asistencia y cuidado personal o en el sector sanitario. De igual modo, cuando la IA se incorpora como componentes de seguridad de productos o sistemas en los sectores de aviación civil, vehículos agrícolas o forestales, equipos marinos, sistema ferroviario, vehículos de motor y sus remolques. Obviamente, todo lo anterior puede afectar al sector público.

En cualquier caso, el mayor interés para el sector público lo atrae el segundo bloque de cuáles son los sistemas IA de alto riesgo. El mismo viene definido por el artículo 6.3º RIA en combinación con el Anexo III. Se trata de sistemas IA utilizados no sólo de manera accesoria o meramente instrumental, sino de manera decisiva o sustantiva para las finalidades previstas del Anexo III.

sistemática en las aplicaciones de los sistemas algorítmicos de alto riesgo, y los órganos de gobernanza de la IA ética del Ayuntamiento podrán decidir si lo aplican también en aplicaciones de riesgo limitado y mínimo. Este riesgo será determinado por la Oficina Técnica de Inteligencia Artificial del IMI.”

⁴³ Concepción del servicio: El momento que transcurre entre que se detecta la necesidad de desarrollar un sistema algorítmico en el Ayuntamiento de Barcelona y el desarrollo de dicho sistema (ya sea a través de un proyecto piloto o de los procedimientos de contratación estándares) es clave para garantizar que su implementación se realiza de forma transparente, segura, legítima y proporcionada y que se incluye a las personas afectadas y las personas externas expertas en el proceso de deliberación, en aquellas aplicaciones que tienen un riesgo más elevado.

Nótese que no se requiere que la decisión sea exclusivamente o sólo automatizada, sino que no sea meramente instrumental. De este modo y sin duda considero que todos los sistemas que evalúan hechos y datos y sirven para iniciar investigaciones o actuaciones sí que quedarían incluidos, a diferencia de la situación actual.

Una parte muy importante del listado del Anexo III viene referido directamente a usos públicos de IA (especialmente policial, de seguridad y judicial) y en otros muchos casos, será muy habitual que los usuarios de estos sistemas de alto riesgo sean del sector público (administración prestacional, educación, empleo, emergencias, etc.). Así, se considera de alto riesgo el uso de sistemas de identificación biométrica en el contexto de la seguridad pública, el uso de sistemas de IA para la gestión y priorización de emergencias y fronteras, los usos relacionados con la aplicación policial y judicial de la ley: evaluación de riesgos individuales de delinquir y reincidir, evaluación de perfiles, pruebas y polígrafos, detección estado emocional, análisis de documentación. Asimismo, en el Anexo III del RIA se consideran de alto riesgo usos de IA pública para determinar la admisión y programas en el ámbito de la educación, evaluación y seguimiento de aprendizaje. También los sistemas relativos a los servicios de empleo y evaluación de rendimiento. Ya más próximo a la actividad administrativa, son de alto riesgo los sistemas de evaluación, concesión, revocación de prestaciones y servicios esenciales de asistencia pública. El presente estudio en buena medida deja fuera el ámbito de justicia y penitenciario que tantas particularidades -e impactos en derechos procesales- puede presentar, sin perjuicio de que se comparten muchas bases. Respecto de los casos de uso en 2020 eran más de 200 los relativos a la Administración de justicia en la UE (Comisión Europea- TRASYs International, 2020). Y respecto del reconocimiento facial y biométrico e IA, me remito a recientes estudios (Cotino 2022b).

En cualquier caso, se dice que un tercio de los sistemas de IA públicos serán de alto riesgo⁴⁴, mientras que sólo un 10% de los públicos y privados lo serían⁴⁵.

Sin duda, podemos coincidir en que son de alto riesgo todos los que están en el listado del RIA. De hecho, muchos de ellos para el CEPD y el Supervisor deberían estar directamente prohibidos (2021).

El Documento de Barcelona se vincula casi exclusivamente a la clasificación del futuro RIA:

Clasificación de riesgos propuesta por la Comisión Europea

La clasificación de riesgos de cualquier sistema de IA licitado o impulsado por el Ayuntamiento de Barcelona será la que determine el reglamento europeo de IA, en proceso de tramitación. La Oficina Técnica del IMI será la que adjudicará la categoría de riesgo a las solicitudes de sistemas.

Estos riesgos hacen referencia al contexto de aplicación del sistema y a la finalidad que se le da. Los aspectos relacionados con la protección de datos personales y la ciberseguridad se regulan de acuerdo con lo previsto en las leyes vigentes sobre la materia.

RIESGO INACEPTABLE

⁴⁴ JRC, Tangi, L. y otros (2022): *AI Watch European landscape on the use of Artificial Intelligence by the Public Sector*, JRC Science For Policy Report, Unión Europea, p. 58.

⁴⁵ Comisión Europea, Renda A. (2021): *Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe. Final Report (D5)*, abril 2021.

Se prohibirán todos los sistemas de IA que se consideren una clara amenaza para la seguridad y los derechos de las personas, desde el *social scoring* de los gobiernos hasta los juguetes que utilicen asistencia de voz y que fomenten comportamientos peligrosos. Se prohibirán también los sistemas de identificación biométrica remota en espacios públicamente accesibles.

RIESGO ALTO

Aquí se incluye la tecnología de IA utilizada en los siguientes casos:

- identificación biométrica remota en tiempo real o en diferido de personas físicas;
- infraestructuras críticas (por ejemplo, el transporte) que podrían poner en riesgo la vida y la salud de las personas;
- formación educativa o profesional, que puede determinar el acceso a la educación y el desarrollo profesional de alguien (por ejemplo, la evaluación de los exámenes);
- los componentes de seguridad de los productos (por ejemplo, la aplicación de la IA en la cirugía asistida por robots);
- la ocupación, la gestión de los trabajadores y el acceso al autoempleo (por ejemplo, software de clasificación de currículums para los procedimientos de contratación);
- servicios privados y públicos esenciales (por ejemplo, la cualificación de créditos que nieguen a la ciudadanía la oportunidad de obtener un préstamo);
- una aplicación de la ley que puede interferir en los derechos fundamentales de las personas (por ejemplo, evaluación de la fiabilidad de las pruebas);
- la gestión de la migración, el asilo y el control de fronteras (por ejemplo, verificación de la autenticidad de los documentos de viaje), y
- la administración de justicia y los procesos democráticos (por ejemplo, la aplicación de la ley a un conjunto concreto de hechos).

Todos los sistemas de identificación biométrica a distancia se consideran de alto riesgo y están sujetos a requisitos estrictos. El uso de la identificación biométrica a distancia en espacios de acceso público con finalidades policiales está, en principio, prohibido.

Se definen y regulan estrictamente algunas excepciones, como cuando es necesario buscar a un niño desaparecido, prevenir una amenaza terrorista específica e inminente o detectar, localizar, identificar o procesar a una persona autora o sospechosa de un delito grave.

Esta utilización está sujeta a la autorización de un órgano judicial u otro órgano independiente y a unos límites adecuados en cuanto a tiempo, alcance geográfico y bases de datos consultadas.”

2. Más allá del futuro RIA. El “impacto” o relevancia y los efectos significativos del sistema inteligencia artificial

Puede haber muchos otros usos y sistemas públicos de IA especialmente impactantes que no están en estos listados de alto riesgo del futuro RIA. Llama muy poderosamente la atención, por ejemplo, que no se hayan incluido los sistemas para la persecución del fraude, blanqueo, impuestos, seguridad social, inspección de trabajo, sanción de tráfico o similares que pasan por ser especialmente conflictivos. Se trata de una opción a mi juicio inadmisibles, dado

el enorme avance de estos usos de IA pública y que tanto impacto generan con total opacidad (entre otros, Olivares 2020 y 2021).

Para evaluar el impacto hay que ver cada caso específico de uso y los concretos derechos en juego. Así, por ejemplo, los errores, sesgos o posibles discriminaciones deben ser evaluados y determinarse hasta qué punto son asumibles por la organización. Como recuerda Zlotnik⁴⁶, hay que utilizar métricas de impacto y no sólo métricas de error. El impacto no es lo mismo que el error, dado que un 0,1% de errores de que una persona muera, se prorrogue su estancia en la cárcel, sea detenida o que se produzca un incidente en una central nuclear, nada tiene que ver con la posibilidad de que se le conceda erróneamente una ayuda escolar, se le priorice para una cita médica general o se le concedan entradas gratuitas para a una actividad cultural. De igual modo hay que tener en cuenta si los sesgos o errores se tienen “efectos sustancialmente más perjudiciales para un grupo protegido”⁴⁷. De ahí que en el terreno de la robustez, exactitud y sesgos, no es fácil establecer apriorísticamente los umbrales asumibles o no, pero sí es posible exigir que se hayan analizado, determinado, evaluado y adoptado medidas al respecto.

Algunos criterios básicos sobre el impacto del uso de IA están establecidos normativamente y cabe proyectarlos también para los sistemas IA público. Así, el impacto o relevancia del uso puede ser muy variable. Un criterio consolidado legislativamente es el que la decisión automatizada o el sistema IA en nuestro caso público “produzca efectos jurídicos en él o le afecte significativamente de modo similar” al individuo (art. 22 RGPD). Y acierta también la Ley 1/2022, de 13 de abril valenciana cuando un elemento esencial (para que deba o no hacerse publicidad activa de un sistema) es “que tengan impacto en los procedimientos administrativos o la prestación de los servicios públicos” (art. 16.1 I). Este concepto de “impacto”, tan habitual en las evaluaciones de riesgos y en el contexto de la protección de datos, pasa a ser clave y, de hecho, del ámbito de la protección de datos hay que abogar a la generalización de este instrumento en formato multirriesgo no limitado a un derecho (Mantelero, 2022; Cotino, 2022a).

Por cuanto a los “efectos significativos” del artículo 22 RGPD, para el sector privado se consideran ejemplos típicos “la denegación automática de una solicitud de crédito en línea” o “los servicios de contratación en red en los que no medie intervención humana alguna” (considerando 71 RGPD). El Grupo del artículo 29, ahora CEPD, (G29 2018: 24) incluye como “significativas” decisiones automatizadas de crédito, servicios sanitarios, oportunidades laborales o de acceso a la educación y, así lo será si se trata del sector público.

Podríamos pensar que el uso de la IA para la personalización y servicios 360° no tienen impacto o efectos jurídicos. No obstante, el artículo 20 de la Ley nº 13.709, de 14 de agosto de 2018 de protección de datos de Brasil incluye entre decisiones automatizadas quedan “incluidas las decisiones destinadas a definir su perfil personal, profesional, de consumo y de crédito o los aspectos de

⁴⁶ Zlotnik, Alexander (2019): “Inteligencia Artificial en las Administraciones Públicas: definiciones, evaluación de viabilidad de proyectos y áreas de aplicación”, *Boletic*, núm. 84, 27-28.

⁴⁷ STC 145/1991, FJ 2º; STC 128/1987, FJ 6º, ver voto particular a la STC 240/1999.

su personalidad”. Es más, el G29⁴⁸ considera que el perfilado en la publicidad en línea sí que puede ser “significativo” y por tanto, las garantías del artículo 22 sí que serían aplicables en razón del “intrusismo” en el perfilado, las expectativas y deseos de las personas afectadas; la forma en que se presenta el anuncio; o el uso de conocimientos sobre las vulnerabilidades de los interesados.” Estos criterios también podrían aplicarse para el sector público en los casos de personalización de información y servicios.

3. Usos públicos de IA para tomar decisiones individualizadas respecto de personas, decisiones internas administrativas o para la elaboración de políticas y su impacto colectivo.

Un elemento también decisivo para valorar el impacto y la capacidad de producir efectos jurídicos es el grado en el que la salida o decisión del sistema de IA público se individualice y aplique a una persona. En esta dirección el artículo L311-3-1 de la ley francesa⁴⁹ hace referencia a “una decisión individual” y en la guía del Etalab para su cumplimiento⁵⁰, se tiene en cuenta que “con ayuda de este tratamiento se adopten decisiones administrativas individuales con respecto a personas físicas o jurídicas, de derecho público o privado, designadas nominativamente”.

Así, la incisión o impacto sobre personas será muy diferente si el sistema IA está enfocado a tomar decisiones sobre individuos o predecir su comportamiento o, por ejemplo, si en el ámbito G2G, está dedicado a tomar decisiones internas administrativas o a proponer políticas y regulaciones, con difícil individualización de las decisiones. Es claro que tiene menor impacto -y por su puesto la admisibilidad- el uso de IA para la mejora generalizada de sistemas Administrativos o configuración de servicios públicos.

En cualquier caso, en principio tiene mayor impacto una decisión IA que se individualice en una persona y más si se adopta la decisión directamente por el sistema. También impacta indudablemente el sistema IA que fundamenta la decisión de intervenir médicamente o no a una persona, o sirve para aprobar o no su examen. Ello tiene claramente más impacto que si el sistema propone la regulación y funcionamiento de los semáforos o la carencia de los autobuses de una línea de transporte.

Dicho lo anterior, la salida más fácil sería sostener que hay menor impacto si los sistemas IA se utilizan como base para elaborar políticas o configuración de servicios públicos. Sin embargo, (Cotino, 2022a) , no se puede ignorar el impacto estructural y masivo del uso de estos sistemas en decisiones generales políticas o normativas. También la Recomendación Unesco es llamativamente insistente en la necesidad de un enfoque colectivo. En esta dirección, respecto de la transparencia en Francia se hace referencia a que hay que explicar sus efectos, “explicando un resultado individual, pero también especificando los impactos generales y particulares” (Etalab, 2022: 9).

⁴⁸ *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*. 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, pp. 24-25.

⁴⁹ Código de relaciones entre el público y la administración https://www.legifrance.gouv.fr/codes/article_lc/LEGinteligencia_artificialRTI000033205535

⁵⁰ Etalab (2022): *Expliquer les algorithmes publics*, p. 8.

4. Usos masivos o perpetuadores de ilegalidad y la necesidad de “recalibrar” las garantías exigibles

A la hora de valorar el impacto, hay que tener muy en cuenta la potencialidad de que el uso de IA sea más o menos generalizable y masivo. De igual modo, al valorar el impacto, hay que tener en cuenta que un error o sesgo del sistema puede reforzar o perpetuar aplicaciones y resultados discriminatorios o sesgados en el futuro, pues los sistemas muy posiblemente acentuarán sus decisiones al nutrirse de nuevos datos cada vez más negativos para los sectores perjudicados.

Es por ello que para sistemas de IA públicos que en principio se aplican masivamente hay que “recalibrar” estos umbrales aceptables y las garantías aplicables. Y considero que en general hay que ser mucho menos tolerantes (Citron, 2008: 1286). Para establecer el umbral de afectación del sesgo o diferencia algorítmica en un caso concreto, no sólo hay que tener en cuenta sólo la afectación de un sistema IA al concreto sujeto que se le ha aplicado, sino que habrá de ponderarse el peligro que supone un error o sesgo masivo en innumerables casos futuros, así como el beneficio significativo de evitar que se replique en miles o millones de decisiones. Además, si el error no se controla, analiza y en su caso se corrige, las decisiones erróneas pasarán a ser *big data* que alimentará a los futuros algoritmos, haciendo que el sesgo o la discriminación se cronifiquen. En consecuencia, en vez de considerar como relevantes diferencias de trato “muy elevadas”, “considerablemente inferiores” o “muy superiores”, habrán de bastar tasas de afectación mucho menores. Se seguiría así para el ámbito algorítmico la doctrina Sentencia TJUE 9 de febrero de 1997 C-167/97 (Soriano, 2021; sobre el tema Cotino, 2023a).

5. La mayor o menor automatización de la actuación administrativa

Niveles de automatización de la actuación administrativa

Para determinar el mayor o menor riesgo del sistema de IA público y la posibilidad de que impacte en derechos, bienes o intereses jurídicos protegibles, son importantes los niveles de mayor o menor automatización e intervención o supervisión humana. Para ello, resultan útiles los estándares que se generaron en 2016 para el vehículo autónomo⁵¹: Nivel 0: Sin automatización de conducción; Nivel 1: Asistencia al conductor; Nivel 2: Automatización de conducción parcial; Nivel 3: Automatización de conducción condicional; Nivel 4: Alta automatización de conducción y Nivel 5: Automatización de conducción completa.

Sobre esta base, recientemente Roehl⁵² ha afirmado seis niveles de automatización de la actuación administrativa: “automatización mínima”; “recuperación y tratamiento de datos”; cuando el sistema sugiere los “pasos procedimentales a seguir”; “decisiones asistidas”; “decisiones automatizadas”,

⁵¹ <https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic> y https://www.sae.org/standards/content/j3016_202104/

⁵² Roehl, U. (2022): “Understanding Automated Decision-Making in the Public Sector: A Classification of Automated, Administrative Decision-Making”, en Juell-Skielse, G. y otros (eds.), *Service Automation in the Public Sector. Progress in IS*. Springer, Cham, 35-63. Sigo directamente de Moral Soriano, L. (2022): “Inteligencia artificial y filosofía del derecho decisiones automatizadas, derecho administrativo y argumentación jurídica”, en Llano Alonso F. H. (dir.), *Inteligencia artificial y filosofía del derecho*, Laborun, Madrid, 475-498.

hasta el grado “decisiones autónomas” con sistemas dinámicos de machine learning no supervisado.

Tipo A: automatización mínima. El operador jurídico decide sobre todos los aspectos de expediente administrativo y recibe la asistencia de tecnologías como un procesador de texto. Utilizará una *check-list*, instrucciones, y otro tipo de estándares decisorios que no están volcados en algoritmos.

Tipo B: recuperación y tratamiento de datos. La decisión es compartida entre el operador jurídico y la tecnología. Ésta recaba, graba y presenta los datos relevantes para resolver el expediente. Por ejemplo, la concesión de becas al estudio requiere una tecnología que examine las solicitudes y extraiga los datos relevantes de las bases de datos de la Administración Pública.

Tipo C: pasos procedimentales a seguir. Igualmente se produce una decisión compartida entre el operador y la tecnología. En este caso, la tecnología además de recuperar y seleccionar los datos relevantes, sugiere los siguientes pasos en el procedimiento. Por ejemplo, la tecnología utilizada en Estados Unidos para decidir las ayudas a los niños con discapacidad pertenece a esta categoría ya que el sistema evalúa las solicitudes: para los casos más sencillos se hace una recomendación automática de decisión, mientras que, para los casos más complejos, la tecnología sugiere que se evalúen directamente por el operador jurídico.

Tipo D: decisiones asistidas. La decisión es compartida entre el operador jurídico y la tecnología. Ésta recaba, graba y presenta algunos o todos los datos relevantes de un expediente y además sugiere un número limitado de soluciones o incluso una decisión específica. El ejemplo anterior sirve aquí también en tanto que la máquina propone o recomienda las decisiones posibles que puede adoptar el operador jurídico.

Tipo E: decisiones automatizadas. La tecnología, no el operador jurídico, es el autor principal de la decisión. Todos los aspectos se confían a la tecnología que opera automáticamente a partir de estadísticas y correlaciones, sin la asistencia del funcionario en el proceso de toma de decisión. Siguiendo el ejemplo de la concesión de becas, tras la recuperación y cruce de datos, el algoritmo decide la cuantía de la beca sin intervención del operador jurídico. Otro ejemplo lo ofrece la tecnología que identifica y notifica a los ciudadanos la deuda adquirida por haber recibido beneficios sociales indebidos; si el ciudadano no impugna la notificación en un determinado plazo, la tecnología comienza el procedimiento para la recuperación de la deuda. Algunos aspectos de estas decisiones automatizadas podrían incluso considerarse propias del siguiente tipo de tecnología.

Tipo F: Decisiones autónomas. De nuevo aquí el autor principal de la decisión es la tecnología. Todos los aspectos de la decisión administrativa se confían a la tecnología basadas en sistemas dinámicos de *machine learning* no supervisado, en los que el operador jurídico no interviene en el proceso de toma de decisión.

Ahora bien, pese a que esta escala pueda resultar útil, las posibles discusiones se centran en el grado en el que la decisión totalmente automatizada luego se integre en una decisión, actuación, proceso o procedimiento Administrativo. Así, puede haber decisiones totalmente autónomas que se integren más o menos en un proceso decisonal Administrativo. También puede discutirse el hecho de que el sistema IA cuente con supervisión humana para verificar su correcto funcionamiento -como las obligaciones de supervisión del

artículo 13 y otros RIA-. El funcionamiento es supervisado por humanos, pero puede ser para comprobar que en general es correcto, pero pese a que se dé dicha supervisión, se puede considerar que el sistema y sus decisiones siguen siendo completamente automatizadas. Habrá que determinar la supervisión real de la que se trate puese ello puede, por ejemplo, a determinar si es aplicable o no el actual artículo 41 Ley 40/2015. Asimismo, si se considera que el sistema IA público sí que está supervisado humanamente no es fácil considerar si son aplicables las especiales garantías del artículo 22 RGPD. Es decir, la supervisión que impone el RIA podría llevar a privar de las garantías por ejemplo de la explicación de la decisión “sólo” automatizada”. Considero que en modo alguno ello debería ser así y en cualquier caso aplicarían las garantías exigibles de buena Administración y debido proceso para el sector público.

Mayor riesgo en decisiones “únicamente”, “íntegramente” automatizadas, sin intervención directa humana o que no sean “accesorias” para la decisión humana

Cabe recordar que las especiales garantías del artículo 22 RGPD sólo se dan respecto de las decisiones “únicamente” automatizadas (por todos Palma, 2022), en nuestro caso, del sector público. Igualmente, las escasas garantías del artículo 41. 1º LRJSP se reservan a “cualquier acto o actuación realizada íntegramente a través de medios electrónicos” y que “no haya intervenido de forma directa un empleado público”. Hoy por hoy estos requisitos sirven para una peligrosa huida del Derecho, esto es, en tanto en cuanto se considera que hay intervención humana en la decisión administrativa que se adopte, no se aplicarían las especiales garantías que en su caso suponen estos preceptos. No obstante, el Grupo del artículo 29 (2018: 23) ha puesto límites, consciente del seguimiento rutinario por los humanos del sistema automatizado. Para que no rijan estas especiales garantías la intervención humana ha de ser “significativa, en vez de ser únicamente un gesto simbólico” y llevada a cabo por “persona autorizada y competente”. El estudio de impacto que debe hacerse ha de registrar el grado de intervención humana. La Carta de Derechos Digitales exige el estudio de impacto también en las decisiones “semiautomatizadas” públicas (apartado XVIII. 4º).

El futuro RIA en sus versiones de 2022 de la Presidencia Francesa y especialmente la de la Presidencia Checa, para considerar si el sistema es de alto riesgo ha introducido en el artículo 6.3º que “la salida del sistema no es puramente accesoria con respecto a la acción o decisión pertinente que debe tomarse”. Así pues, hay que valorar si la salida del sistema de IA público, por ejemplo, una propuesta de resolución, puede pasar a ser elemento sustancial para la decisión final que adopte el usuario de su sistema.

Acierta Huergo (2020, 65 y ss., 66 y 75 y ss.) cuando señala que los mayores riesgos se dan respecto de los sistemas predictivos por cuanto incluyen en la acción administrativa añadiendo un “contenido de elaboración propia”. Por el contrario, otros sistemas tienen “un valor meramente auxiliar en la aplicación de la norma” y cabe ubicarlos en el ámbito de la actuación reglada. Valero (2007: 74-75) y Martín Delgado (2009: 368) hablaban de los sistemas automatizados de “baja intensidad” frente a los sistemas de “discrecionalidad política o discrecionalidad técnica no parametrizable”. Afirmaban estos autores que estos últimos no eran posibles en la legislación española.

Sobre estas bases, considero que el peligro del uso de sistemas de IA públicos de menor a mayor pasaría por:

- los sistemas de uso meramente instrumental (procesadores, traductores, etc.);

-también meramente sería instrumental el uso de sistemas automatizados para actos reglados, siempre que el sistema se limite a seguir la decisión normativa.

-Más riesgos en el uso de IA pública se dan en ámbitos de discrecionalidad técnica,

- y más si cabe en el uso para ámbitos de discrecionalidad política o discrecionalidad técnica no parametrizable.

-También podría añadirse a esta escala entre los niveles de mayor riesgo, las decisiones IA administrativas para la aplicación de conceptos jurídicos indeterminados.

Ahora bien, pese a la utilidad de estas escalas para valorar el riesgo, hay que insistir que desde el nivel más básico de sistemas automatizados y algorítmicos es más que posible que el sistema genere lesiones de derechos o intereses y vulneraciones de la legalidad. Basta una mala selección o calidad de datos o un mero error en un simple algoritmo o fórmula para que quede viciado e impacte en la decisión final que se adopte, pese a que sea formalmente humana.

6. Una escala de usos públicos de inteligencia artificial más o menos impactantes

Dejando ahora al margen los usos vinculados a defensa y seguridad nacional, me permito escalar de mayor a menor impacto y riesgo, así como de atracción jurídica diversas categorizaciones de usos de IA públicos, a saber:

-Sistemas calificables como de alto riesgo por el futuro RIA.

- Sistemas que identifican o priorizan objetivos para la aplicación de la ley o para realizar inspecciones en el ámbito de infracciones penales, administrativas y persecución de ilicitudes y fraude, cada vez más habituales en ámbitos de seguridad, mercado y competencia, trabajo, salud, cuidado del medio ambiente, etc.

- Sistemas para la adjudicación de contratos, subvenciones, privilegios del gobierno.

- Sistemas IA para personalizar, priorizar o apoyar la prestación de servicios a los ciudadanos en salud, educación, empleo, servicios sociales, etc.

- Sistemas de extracción de información, investigación, recopilación, supervisión y el análisis de datos, para la elaboración de políticas, la toma de decisiones, monitoreo general y análisis de riesgos.

- Sistemas para la gestión de la organización interna, recursos humanos y las adquisiciones y la gestión de los recursos tecnológicos.

- Sistemas para interactuar y comunicarse con el público sobre sus derechos y obligaciones y su participación.

Señala Huergo (2020: 77) que los sistemas de selección de sujetos para ser inspeccionados “quedarían en un segundo plano y con una “irrelevancia jurídica casi total”. Sin embargo, pese a que las normas y jurisprudencia actuales en España parecen darle la razón, es momento de evitar esta huida del Derecho (Cotino, 2021) y aplicar las necesarias garantías a estos sistemas automatizados que encienden las luces rojas determinando o guiando la actuación humana. Los falsos positivos y negativos pueden generar importantes lesiones de derechos y

posibles discriminaciones. El caso judicializado más importante hasta la fecha de Países Bajos de SyRI (Cotino, 2021), con mucho acierto determinó la nulidad del sistema que se utilizaba sólo en la fase de selección de alertas.

Segundo. Qué información hay que facilitar de los sistemas algorítmicos y de inteligencia artificial en razón de la Ley valenciana

Lorenzo Cotino Hueso
Alba Soriano Arnanz

I. A modo de introducción

La Ley 1/2022, de 13 de abril, de la Generalitat, de Transparencia y Buen Gobierno de la Comunitat Valenciana en su artículo 16, sobre “Información de relevancia jurídica”, dispone que:

1. Las administraciones públicas del artículo 3.2 deben publicar:

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.*

Interesa ahora centrarse en la delimitación del ámbito de aplicación de este precepto, en particular, en qué información habrá de contener el registro de algoritmos o publicidad activa obligatoria en el portal de transparencia.

A continuación, se lleva a cabo una interpretación del precepto y una propuesta de los elementos a tener en cuenta para la implantación del mismo.

Siendo que la transparencia es esencialmente un instrumento para lograr unas finalidades, el estudio se cuestiona el ¿para qué? de la transparencia y los muy variados motivos o finalidades.

Hay varias formas en que la transparencia hace avanzar la democratización de las sociedades. En primer lugar, la transparencia empodera a la ciudadanía en la medida en que proporciona información sobre las acciones y procedimientos realizados por los poderes públicos, permitiendo a ciudadanos y ciudadanas la posibilidad de participar en los procesos de toma de decisiones que puedan interesarles. En segundo lugar, y en estrecha relación con el punto anterior, para que las instituciones públicas rindan cuentas no solo es necesario establecer procedimientos que permitan a los individuos defender sus derechos e intereses, sino también proporcionarles los conocimientos necesarios sobre la forma en que se ha adoptado una decisión que les afecta y que pueden desear impugnar. En tercer lugar, la transparencia también es necesaria para proporcionar a las personas y organizaciones los conocimientos necesarios para impugnar e iniciar debates públicos sobre cualquier acción o decisión adoptada por los poderes públicos y, por lo tanto, para ejercer una supervisión efectiva sobre dichas acciones. Por último, como consecuencia de todo lo anterior, la transparencia ayuda a prevenir que se lleven a cabo conductas ilícitas y

proporciona legitimidad a las instituciones públicas. Si la ciudadanía conoce las acciones y decisiones tomadas por los poderes públicos, cómo impugnarlas en caso de que les afecten y se les permite participar en los diferentes procesos de toma de decisiones, aumentará su confianza en el sistema.

Con carácter general, esencialmente la transparencia algorítmica es un elemento instrumental que favorece la necesitada confiabilidad en la IA. Y sobre todo puede discernirse, por un lado, que la transparencia algorítmica es la herramienta básica de la garantía de toda una serie de principios democráticos y derechos fundamentales al posibilitar su control.

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. Asimismo, 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.

Mucha transparencia puede resultar excesiva e incluso contraproducente para conocer el sistema de IA. Y, sobre todo, mucha transparencia puede generar problemas de privacidad, seguridad, confidencialidad y propiedad intelectual. Encontrar el balance adecuado para las finalidades que se persigan es la clave.

Según las Directrices para el sector público de Países Bajos⁵³, el grado de exigencia de la explicabilidad y la transparencia depende de (1) el impacto del algoritmo en la decisión, el resultado y el ciudadano; (2) el grado de autonomía en la toma de decisiones (es decir, hasta qué punto se garantiza la participación humana); y (3) el tipo y la complejidad del algoritmo.

En el caso del Reino Unido⁵⁴, se establece que la transparencia es particularmente relevante para aquellos sistemas que:

- Tienen una influencia significativa en un proceso de toma de decisiones con efecto público directo o indirecto, o
- Interactúan directamente con el público en general.

Asimismo, estos criterios, referentes al impacto, grado de autonomía y tipo y complejidad del sistema son también útiles a la hora de decidir cuáles deben ser los sistemas algorítmicos que deben dotarse de transparencia con mayor urgencia y, por tanto, deben ser los primeros en ser incorporados al registro de algoritmos. En este sentido, cabe que nos refiramos a las directrices

⁵³ Directrices para el sector público, Ministerie van Justitie en Veiligheid (Ministerie van Justitie en Veiligheid), Rijksoverheid (Gobierno central), *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 08-03-2021, <https://www.rijksoverheid.nl/documenten/richtlijnen/2021/09/24/richtlijnen-voor-het-toepassen-van-algoritmen-door-overheden-en-publieksvoorlichting-over-data-analyses#:~:text=Rijksoverheid-.Richtlijnen%20voor%20het%20toepassen%20van%20algoritmen,en%20publieksvoorlichting%20over%20data%20analyses.&text=Doel%20van%20de%20richtlijnen%20is,de%20publieksvoorlichting%20daarbij%20door%20overheden>

⁵⁴ Gov.uk, "Algorithmic Transparency Recording Standard - Guidance for Public Sector Bodies". Disponible en: <https://www.gov.uk/government/publications/guidance-for-organisations-using-the-algorithmic-transparency-recording-standard/algorithmic-transparency-recording-standard-guidance-for-public-sector-bodies>

desarrolladas por el *Digital Forum Lab* de *Eurocities*. Dicha organización ha elaborado una propuesta de Registro de algoritmos para su implementación por parte de diferentes ciudades europeas y, entre la documentación que proporciona, establece los pasos a seguir en la creación del registro de algoritmos indicando, como segundo paso, la necesidad de fijar un orden a la hora de registrar sistemas algorítmicos. En este caso, los criterios señalados por Eurocities son, principalmente, el impacto del sistema sobre las personas, su riesgo y su nivel de explicabilidad (a menor explicabilidad, mayor transparencia)⁵⁵.

Se pueden afirmar como principios o planteamientos generales que, a más impacto o potencialidad de daño de un sistema de IA en derechos e intereses, es precisa más transparencia.

“Cuanto mayor es el impacto, más importante es la transparencia”, “Cuanto mayor sea el impacto de los análisis de datos en los ciudadanos, más importante será la transparencia. Esto se aplica ciertamente en los casos en que tales análisis (pueden) conducir a conclusiones o decisiones.”

Una buena muestra es la propia regulación de protección de datos respecto de las decisiones automatizadas (además de las garantías generales de protección de datos, el artículo 22 RGPD consagra garantías añadidas respecto de decisiones solo automatizadas y relevantes).

Es necesaria más transparencia a menor intervención, control o supervisión humana del sistema de IA.

Si un servicio público recurre a la transparencia en términos generales con respecto a sus análisis de datos, pero no divulga ciertos aspectos más detallados con vistas a los intereses de la investigación, por ejemplo, debe al menos compensar esa conducta con una supervisión interna y externa suficiente de esos aspectos⁵⁶.

También cabe aceptar como principio que, a mayor opacidad o complejidad del sistema de IA, mayor transparencia y, como se verá, mayor explicabilidad y otros contenidos del genérico de la transparencia serán requeridas.

No hay que perder el rumbo de la exigencia de transparencia. En esta dirección, la Recomendación sobre la ética de la IA UNESCO⁵⁷ (nº 37): “transparencia y la explicabilidad de los sistemas de IA suelen ser condiciones previas fundamentales para garantizar el respeto, la protección y la promoción de los derechos humanos, las libertades fundamentales y los principios éticos.” Esencialmente permiten “conocer los motivos por los que se ha tomado una decisión [...] y tener la posibilidad de presentar alegaciones. [...] La falta de transparencia también podría mermar la posibilidad de impugnar eficazmente las decisiones.

Recuerda también Mantelero que, pese a los muchos obstáculos de la transparencia de los algoritmos, la transparencia y la inteligibilidad de la IA son

⁵⁵ Algorithm Register, “Guidance for European Cities: Getting started with an Algorithm Register”. Disponible en: <https://www.algorithmregister.org/guidance>

⁵⁶ *Ibidem*, pág. 29.

⁵⁷ Conferencia General 41ª reunión - París, 41 C/73, 22 de noviembre de 2021. Anexo. https://unesdoc.unesco.org/ark:/48223/pf0000379920_spa

elementos esenciales para salvaguardar la autodeterminación individual y colectiva⁵⁸.

Se siguen a continuación los materiales de referencia de diversas autoridades especialmente extranjeras, así como las declaraciones y manifiestos que hemos hecho desde la Red DAIA⁵⁹ o los mejores trabajos en nuestro país como los del monográfico que tuve ocasión de coordinar con Boix⁶⁰, destacando muy especialmente respecto de la transparencia los trabajos de Gutiérrez⁶¹, Vestri⁶² y Cerrillo⁶³.

II. Cómo ha de ser la información: “la descripción de manera comprensible”. El nivel de profundidad y comprensibilidad de la información que se ofrece. La información por capas

La norma valenciana claramente expresa que la información que se facilite vaya dirigida al público en general. No obstante, también respecto de estas finalidades de la transparencia algorítmica debe tenerse en cuenta a la comunidad de especialistas o científicos que sí que cuentan con alto conocimiento. Las Directrices o “Pautas” para el uso de IA en Países Bajos por el sector público (en adelante Directrices para el sector público)⁶⁴ o la guía de transparencia algorítmica pública del ICO para el Reino Unido⁶⁵ sí que están

⁵⁸ Mantelero, *Beyond Data. Human Rights, Ethical and Social Impact Assessment in AI*, Information Technology and Law Series, IT&LAW 36, 2022, pág. 98.pág. 31.

⁵⁹ Desde la Red Derecho Administrativo de la IA (DAIA) desde 2019 hemos subrayado elementos básicos jurídicos respecto del uso de la IA en el sector público, entre ellos las cuestiones de transparencia, que sin duda constituyen un referente. Así en las [Conclusiones](#) de Toledo de 1 de abril de 2019 y la [Declaración final de Valencia](#) de 24 de octubre de 2019.

⁶⁰ A. Boix Palop, y L. Cotino Hueso, (coords.) *Monográfico Derecho Público, derechos y transparencia ante el uso de algoritmos, inteligencia artificial y big data RGDA Iustel*, nº 50, febrero 2019. Acceso en https://www.iustel.com/v2/revistas/detalle_revista.asp?id=1

⁶¹ Posiblemente el trabajo más riguroso, aunando además los elementos técnicos esenciales de la cuestión, Gutiérrez David M.E., Gutiérrez David, “Administraciones inteligentes y acceso al código fuente y los algoritmos públicos. Conjurando riesgos de cajas negras decisionales”, en *Derecom*, nº 31, págs. 19-105, ver págs. 55-56, 2021, <http://www.derecom.com/derecom/>

⁶² G. Vestri, “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, págs. 368-398.

⁶³ A. Cerrillo i Martínez, “La transparencia de los algoritmos que utilizan las administraciones públicas”, *Anuario de Transparencia Local*, nº. 3, 2020, págs. 41-78.

⁶⁴ Ministerie van Justitie en Veiligheid (Ministerie van Justitie en Veiligheid), Rijksoverheid (Gobierno central), *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 08-03-2021, <https://www.rijksoverheid.nl/documenten/richtlijnen/2021/09/24/richtlijnen-voor-het-toepassen-van-algoritmen-door-overheden-en-publieksvoorlichting-over-data-analyses#:~:text=Rijksoverheid-,Richtlijnen%20voor%20het%20toepassen%20van%20algoritmen, en%20publieksvoorlichting%20over%20data%20analyses.&text=Doel%20van%20de%20richtlijnen%20is,de%20publieksvoorlichting%20daarbij%20door%20overheden>

⁶⁵ ICO (Information Commissioner Office), *What do we need to do to ensure lawfulness, fairness, and transparency in AI systems?*, ICO, 2021, <https://ico.org.uk/for-organisations/guide->

concebidas para la obligación general de informar a todo el público. En el caso de autoridades de protección de datos, se trata bien de la información general al público obligatoria, o bien de las obligaciones más concretas respecto de los afectados, especialmente por decisiones algorítmicas. El HLEG aunque hace referencia a “usuarios”⁶⁶ no parece discernir el tipo de sujeto del que se trata por lo general, tampoco la Recomendación UNESCO 2021.

En estos casos los datos, información o contenidos en que consiste la transparencia y explicabilidad no necesariamente han de ser de un carácter básico o elemental dirigidos al conocimiento medio de la ciudadanía. De hecho, en muchos supuestos, habrá de facilitarse información de alto contenido técnico similar a la dirigida a usuarios (entendidos como las personas físicas o jurídicas que empleen sistemas de IA bajo su autoridad) o entidades supervisoras. Ello es así por cuanto la información de profundidad será la necesaria para poder ejercer derechos por parte del interesado en cuestión, así como por los colectivos o la sociedad civil. De igual modo, la comunidad científica juega un papel muy relevante en el acceso a la información algorítmica.

Que la información deba ir dirigida a diferentes niveles de conocimiento de usuario no es problema. Como es conocido especialmente en el ámbito de la protección de datos, la información debe estar estructurada en capas, de un nivel general más accesible y simple y una capa que permite profundizar en el conocimiento⁶⁷. 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 IA⁶⁸.

Resultan de especial interés las Directrices para el sector público neerlandés 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

[to-data-protection/key-dp-themes/guidance-on-ai-and-data-protection/what-do-we-need-to-do-to-ensure-lawfulness-fairness-and-transparency-in-ai-systems/](https://www.gov.uk/government/collections/algorithmic-transparency-standard)

Y de especial utilidad los estándares ICO (Information Commissioner Office), *Algorithmic transparency data standard*, ICO, 2021 (última versión julio 2022), <https://www.gov.uk/government/collections/algorithmic-transparency-standard>

También, de especial interés, ICO y Alan Turing Institute, *Explaining decisions made with Artificial Intelligence*, 2020, <https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/explaining-decisions-made-with-ai/>

⁶⁶ HLEG (Comisión Europea - Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial), *Directrices éticas para una IA fiable*, 2019, pág. 10, <https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>

⁶⁷ En general puede seguirse, AEPD, *Guía para el cumplimiento del deber de Informar*, AEPD, 2018, págs. 5 y ss. <https://www.aepd.es/es/media/guias/guia-modelo-clausula-informativa.pdf>

⁶⁸ Así, las Directrices para el sector público de Países bajos para el caso de información sobre decisiones automatizadas se afirma que la información “está “estratificada” (Rijksoverheid, *Richtlijnen voor het toepassen van algoritmen...* cit. pág. 32. Y especialmente, los estándares del ICO para el sector público están claramente estructurados en estos dos niveles (ICO (Information Commissioner Office), *Algorithmic transparency data standard*, ICO, 2021 (última versión julio 2022), <https://www.gov.uk/government/collections/algorithmic-transparency-standard>.)

entonces en riesgo.”⁶⁹ 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, la propuesta de Reglamento de Inteligencia Artificial de la UE (Ley de inteligencia artificial o AIA por sus siglas en inglés) 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. 2º 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 conocimiento⁷⁰. 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 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.”⁷¹

También en el terreno concreto de la IA, los estándares del ICO para el sector público están claramente estructurados en estos dos niveles.⁷² 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,

⁶⁹ Rijksoverheid, *Richtlijnen voor het toepassen van algoritmen...* cit.

⁷⁰ En general puede seguirse, AEPD, *Guía para el cumplimiento del deber de informar, ...* cit. págs. 5 y ss.

⁷¹ Así, las Directrices para el sector público de Países Bajos para el caso de información sobre decisiones automatizadas se afirma que la información “está "estratificada" (Rijksoverheid, *Richtlijnen voor het toepassen van algoritmen...* cit. pág. 32.

⁷² ICO, *Algorithmic transparency data standard...* cit.

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 para el sector público de Holanda 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.”⁷³

Por su parte, en el caso de Francia⁷⁴, las obligaciones de transparencia se definen en tres niveles:

- A. “Información general: publicar en línea las normas que definen las principales operaciones de tratamiento utilizadas en el ejercicio de sus funciones cuando constituyan la base de decisiones individuales,
- B. Información explícita: incluya una declaración en línea y en los documentos (avisos, notificaciones) en la que se especifiquen los fines del tratamiento, un recordatorio del derecho de acceso y los procedimientos para ejercer este derecho,
- C. Información individual: a petición de la persona interesada, facilitar la siguiente información: la medida y el modo en que el tratamiento algorítmico contribuye a la toma de decisiones, los datos tratados y sus fuentes, los parámetros del tratamiento y su ponderación, aplicados a la situación del interesado, las operaciones realizadas por el tratamiento.”

III. Normativa de referencia sobre la información a facilitar

1. La información básica que exige el Decreto 203/2021

Este precepto está claramente inspirado en el Decreto 203/2021 estatal, en su artículo 11. 1º i) que impone la difusión en la sede electrónica: “descripción de su diseño y funcionamiento, los mecanismos de rendición de cuentas y transparencia, así como los datos utilizados en su configuración y aprendizaje.”⁷⁵

⁷³ “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.”

⁷⁴ Etalab, “Expliquer les algorithmes publiques”, 11/08/2022. Disponible en: https://guides.etalab.gouv.fr/algorithmes/guide/#_3-le-cadre-juridique-applicable

⁷⁵ Para el ámbito tributario, cabe remitir a Bernardo D. Olivares Olivares, en *Transparencia y aplicaciones informáticas en la Administración tributaria*, *Crónica Tributaria*, n.º 174, 2020, pp. 89-113. Disponible en: <https://www.ief.es/docs/destacados/publicaciones/revistas/ct/174.pdf>.
https://drive.google.com/file/d/1p1_fzP0e2ACIfVKdxce_DawVbLAKe4bt/view

2. La información a detallar según el Estatuto de los trabajadores

En el caso del Estatuto de los Trabajadores la ley simplemente prescribe que se ha de informar de “los parámetros, reglas e instrucciones en los que se basan los algoritmos o sistemas de inteligencia artificial”.

Y la Guía de implantación ha desarrollado esta información en cuatro grupos⁷⁶:

- A. Información general sobre el uso de algoritmos o sistemas de inteligencia artificial para tomar decisiones automatizadas
- B. Información sobre la lógica y funcionamiento de cada algoritmo usado
- C. Información sobre las consecuencias que pueden derivarse de la decisión automatizada o del uso del algoritmo
- D. Otra información relevante para la representación legal de la plantilla

A. Información general sobre el uso de algoritmos o sistemas de inteligencia artificial para tomar decisiones automatizadas

- 1. ¿La empresa utiliza o tiene previsto utilizar algoritmos y sistemas de decisión automatizada para la gestión de personas trabajadoras o candidatas? Sí
- 2. ¿La empresa utiliza algoritmos o sistemas de decisión automatizada para la elaboración de perfiles de las personas? Sí
- 3. ¿Qué decisiones de gestión de personas son automatizadas mediante algoritmos o sistemas de inteligencia artificial? Selección de personas
- 4. ¿Qué tipo de tecnología utiliza el algoritmo?
 - a) ¿Genera un modelo de “caja negra”? Sí
 - b) ¿Es un algoritmo de aprendizaje continuo? Sí
 - c) ¿Quién ha desarrollado el modelo?
- 5. ¿Cuál es el software o producto utilizado por el algoritmo?
 - a) ¿Quién es la empresa suministradora?
 - b) ¿Se han realizado modificaciones o alteraciones del software en su instalación? Sí
 - c) En su caso, ¿de qué tipo y con qué finalidad?
 - d) ¿El software cuenta con algún tipo de certificación? Sí
- 6. ¿Existe intervención humana cualificada en el proceso de decisión? Sí
 - a) En su caso, ¿cuál es el grado de intervención de la persona en la adopción de la decisión final?
 - b) ¿Cuál es la competencia y autorización de la persona que interviene en el proceso de decisión?

⁷⁶ Ministerio de Trabajo y Economía Social, *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 2022, https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/trabajo14/Documents/2022/100622-Guia_algoritmos.pdf

c) ¿Cuál es su capacidad para modificar la decisión adoptada por el algoritmo?

B. Información sobre la lógica y funcionamiento de cada algoritmo usado

7. En caso de elaboración de perfiles, ¿qué tipología de perfiles elabora el algoritmo?

a) Específicamente para la persona trabajadora o candidata, ¿a qué perfil ha sido asignada?

b) El perfil ¿ha sido usado o está previsto usarlo para una finalidad distinta?

8. ¿Cuáles son las variables que utiliza el algoritmo?

a) a) ¿Se han incluido como variables datos de carácter personal? Sí

b) ¿Qué tipología de datos de carácter personal son utilizados?

c) ¿Se han cumplido las obligaciones de información respecto del uso o, en su caso, reutilización de datos personales?

d) ¿Qué tipología de datos de carácter no personal son utilizados?

9. ¿Qué parámetros utiliza el algoritmo?

10. ¿Cuáles son las reglas e instrucciones que utiliza el algoritmo en el proceso de toma de decisión?

11. ¿Qué base de datos de entrenamiento y, en su caso, de validación se ha utilizado?

a) ¿Los datos de entrenamiento son adecuados, pertinentes y no excesivos? Sí

b) ¿Se han verificado estos extremos? En caso afirmativo, ¿quién ha realizado la verificación?

c) ¿Su utilización como datos de entrenamiento se corresponde con la finalidad para la que fueron obtenidos? Sí

d) En su caso, ¿qué tipo de patrones se han identificado en los datos de entrenamiento?

12. ¿Se han detectado errores o imprecisiones del algoritmo en la clasificación de las personas en los diferentes perfiles? Sí

a) ¿Cuál es el porcentaje aproximado de error?

b) ¿Cuáles son las métricas de rendimiento por perfil?

13. ¿Se han realizado o se prevé realizar auditorías o evaluaciones de impacto?

a) ¿Se realizan con medios propios o externos?

b) ¿Cuál ha sido el resultado de la auditoría y/o evaluación de impacto?

C. Información sobre las consecuencias que pueden derivarse de la decisión automatizada o del uso del algoritmo

14. ¿Qué consecuencias pueden derivarse para las personas trabajadoras o candidatas de la decisión adoptada por el algoritmo o del perfil creado?

15. A la representación legal de la plantilla, ¿cuál es el impacto el algoritmo

en materia de igualdad y no discriminación entre mujeres y hombres?

- a) ¿Existe constancia de posibles problemas de sesgo?
- b) ¿Se ha incluido el impacto del algoritmo en el diagnóstico previo a la elaboración de un Plan de Igualdad?

D. Otra información relevante para la representación legal de la plantilla

16. ¿Se ha informado a las personas trabajadoras o candidatas respecto del uso de algoritmos para la toma de decisiones automatizadas?

3. Información a proporcionar en el caso del artículo 22 RGPD

Destaca la concreción de la información a proporcionar en el caso del artículo 22 RGPD elaborada por la AEPD⁷⁷.

En el caso de que el interesado esté sometido a decisiones automatizadas o en la elaboración de perfiles a los que hace referencia el artículo 22 del RGPD, un aspecto importante que se establece en el artículo 13.2.f del RGPD es que éste ha de “disponer de información significativa sobre la lógica aplicada” y “la importancia y las consecuencias previstas”.

Aunque dependerá del tipo de componente IA utilizado, un ejemplo de información que podría tener relevancia de cara al interesado, sería:

- 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.

4. Información básica de la base de datos del futuro Reglamento de IA

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

⁷⁷ Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción, de febrero 2020, pág 24. <https://www.aepd.es/media/guias/adecuacion-rgpd-ia.pdf>

tienen que registrarse algunos sistemas de alto riesgo⁷⁸ (art. 51 AIA) y todo el mundo puede acceder a información bastante básica que se detalla en el Anexo VIII.

Como elemento positivo a destacar de esta base de datos, cabe subrayar que pretende ser de aplicación tanto al sector público como al sector privado. Es decir, la obligación de publicar los datos del sistema depende del uso al que esté destinado y no de las características de la entidad que lo utiliza.

Así, el artículo 60, relativo a “Base de datos de la UE para los sistemas de IA de alto riesgo enumerados en el anexo III” dispone también quién debe introducir la información⁷⁹.

ANEXO VIII

INFORMACIÓN QUE DEBE PRESENTARSE EN EL MOMENTO DEL REGISTRO DE LOS OPERADORES Y DE LOS SISTEMAS DE IA DE ALTO RIESGO, DE CONFORMIDAD CON EL ARTÍCULO 51

Los proveedores, los representantes autorizados y los usuarios que sean autoridades, agencias u organismos públicos presentarán la información a que se refiere la Parte I. Los proveedores o, en su caso, los representantes autorizados velarán por que la información sobre sus sistemas de IA de alto riesgo a que se refieren los puntos 1 a 11 de la Parte II sea completa, correcta y se mantenga actualizada. La información establecida en la parte II, 12 será generada automáticamente por la base de datos.

Parte I. Información relativa a los operadores (en el momento del registro de los operadores)

-1. Tipo de operador (proveedor, representante autorizado o usuario);

1. Nombre, dirección y datos de contacto del proveedor;

2. Cuando la presentación de la información sea realizada por otra persona en nombre del del operador, el nombre, la dirección y los datos de contacto de esa persona;

Parte II. Información relacionada con el sistema de IA de alto riesgo

⁷⁸ 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. Los sistemas de alto riesgo cuya información debe ser publicada son, a rasgos generales, aquellos que se emplean en los siguientes contextos: Identificación biométrica y categorización de personas físicas; gestión de infraestructuras esenciales; educación; empleo, en el acceso y disfrute a servicios públicos y privados esenciales; aplicación de la ley; en control de fronteras, migración y asilo y administración de justicia y procesos democráticos.

⁷⁹ “2. Los datos enumerados en el anexo VIII, parte I, serán introducidos en la base de datos de la UE por los proveedores, los representantes autorizados y los usuarios pertinentes, según proceda, en el momento de su registro. Los datos enumerados en el anexo VIII, parte II, 1 a 11, serán introducidos en la base de datos de la UE por los proveedores o, en su caso, por el representante autorizado, de conformidad con el artículo 51. Los datos enumerados en el Anexo VIII, Parte II, 12 serán generados automáticamente por la base de datos a partir de la información proporcionada por los usuarios pertinentes de conformidad con el apartado 2 del artículo 51. Los datos enumerados en el Anexo VIII bis serán introducidos en la base de datos por los posibles proveedores o prestadores, de conformidad con el artículo 54 bis.”

1. Nombre, dirección y datos de contacto del proveedor
2. Nombre, dirección y datos de contacto del representante autorizado, si procede;
3. El nombre comercial del sistema de IA y cualquier otra referencia inequívoca que permita la identificación y la trazabilidad del sistema de IA;
Descripción de la finalidad prevista del sistema de IA;
4. Estado del sistema de IA (en el mercado, o en servicio; ya no está en el mercado/en servicio, retirado);
5. Tipo, número y fecha de caducidad del certificado emitido por el organismo notificado y el nombre o número de identificación de dicho organismo notificado, cuando proceda;
6. Una copia escaneada del certificado mencionado en el punto 76, cuando proceda;
7. Estados miembros en los que el sistema de IA es o ha sido comercializado, puesto en servicio o disponible en la Unión;
8. Una copia de la declaración de conformidad de la UE mencionada en el artículo 48;
9. Instrucciones electrónicas de uso; URL para información adicional (opcional).
10. Nombre, dirección y datos de contacto de los usuarios

5. Ayuntamiento de Barcelona

En el caso del Ayuntamiento de Barcelona, se distingue entre “Transparencia del procedimiento”⁸⁰, “Transparencia técnica”⁸¹ y “Explicabilidad”⁸².

⁸⁰ Transparencia del procedimiento: Se trata del suministro de información pública relativa al sistema algorítmico o al proceso seguido en el desarrollo y aplicación del mismo y de los datos utilizados en este contexto. En cualquier caso, la información suministrada debe ser suficiente para garantizar que, si surgen daños, se puede rastrear su causa, así como para garantizar la comprensión de las opciones y supuestos asumidos y las categorías de datos utilizadas en el desarrollo del sistema algorítmico. Debe incluir también la forma en que se da la intervención humana en el sistema algorítmico, la metodología utilizada para identificar los riesgos, los riesgos identificados y las medidas adoptadas para mitigarlos, así como las partes que han participado en el desarrollo del sistema algorítmico y sus funciones.

⁸¹ Transparencia técnica: Se trata del suministro de información por parte del proveedor al Ayuntamiento de Barcelona —respetando la propiedad intelectual— que permita comprender el funcionamiento técnico del sistema algorítmico. Esta información debe incluir la divulgación del código fuente del sistema, las especificaciones técnicas y los datos utilizados en el desarrollo del sistema, la información técnica sobre cómo se han obtenido los datos, el proceso de desarrollo utilizado, la justificación de la elección del modelo concreto y sus parámetros, así como la información sobre el rendimiento del sistema.

⁸² Explicabilidad: Información sobre por qué un sistema algorítmico lleva a una decisión o resultado concreto. Esto incluirá una indicación clara de los factores clave que han llevado a un sistema algorítmico a un resultado concreto y los cambios que deben realizarse en los datos de entrada de cara a obtener un resultado diferente. Hacer que un sistema algorítmico sea explicable incluye proporcionar toda la información técnica y otra tipología necesaria para explicar a los espacios de participación o en procedimientos cómo los estudios de impacto, las auditorías algorítmicas o en procedimientos de impugnación, recurso u otros procedimientos judiciales,

La transparencia en el procedimiento y la transparencia técnica la consideran necesaria para la evaluación del sistema algorítmico por los propios servicios municipales.

Ya de cara a la ciudadanía, “la transparencia en el procedimiento se centra también en proporcionar a la ciudadanía y otras partes interesadas información general sobre la utilización del sistema algorítmico”. El establecimiento de un registro de algoritmos del Ayuntamiento de Barcelona⁸³, previsto para el presente año, se basa en las directrices y principios del *Algorithm Transparency Standard* y la propuesta de registro de algoritmos elaborados por el *Eurocities Digital Forum Lab*⁸⁴.

Para los sistemas de “alto riesgo” según su propia clasificación, y “en la medida del posible” para el resto de sistemas, se afirma que “el Ayuntamiento de Barcelona será libre de divulgar la información facilitada por el proveedor relacionada con la transparencia del procedimiento. En cambio, la información relacionada con la transparencia técnica que facilitarán los proveedores no será de carácter público”.

Así las cosas, en principio de cara a la ciudadanía se facilitaría la información sobre, según su definición:

- sistema algorítmico
- proceso seguido en el desarrollo y aplicación del mismo
- los datos utilizados en este contexto.
- la forma en que se da la intervención humana en el sistema algorítmico,
- la metodología utilizada para identificar los riesgos,
- los riesgos identificados y las medidas adoptadas para mitigarlos,
- las partes que han participado en el desarrollo del sistema algorítmico y sus funciones.

Se afirma asimismo que “la información suministrada debe ser suficiente para garantizar que, si surgen daños, se puede rastrear su causa, así como para garantizar la comprensión de las opciones y supuestos asumidos y las categorías de datos utilizadas en el desarrollo del sistema algorítmico.”

En otro apartado relativo a los “Espacios de transparencia”, para todos los tipos de sistemas, se “pondrá a disposición de la ciudadanía y de todas las partes interesadas la información pública relacionada con los sistemas algorítmicos que licite en el Registro Público de Algoritmos que se creará a estos efectos.”

Y se afirma que “de forma preliminar y no restrictiva”²³, se prevé que se publiquen:

- Información básica sobre el sistema algorítmico:

Finalidad del sistema algorítmico

como se ha llegado a una decisión y ofrecer a la otra parte o cualquier parte interesada el acceso a la información y una protección jurídica satisfactoria.

⁸³ Ajuntament de Barcelona, “Un registre d’algoritmes municipal transparent i ètic per millorar els serveis públics”. Disponible en: https://www.barcelona.cat/infobarcelona/ca/tema/drets-digitalis/un-registre-dalgoritmes-municipal-transparent-i-etic-per-millorar-els-serveis-publics_1245666.html

⁸⁴ Algorithm Register, “Guidance for European Cities...”, *cit.*

Área municipal que lo impulsa

Actores implicados en las decisiones del sistema

Alcance del sistema algorítmico

Datos que se harán servir

Incorporación de aprendizaje automático

Política de utilización

Calendario de aplicación

Riesgos que podrían aparecer y estrategia de gestión de estos

Legitimación de la finalidad en el caso de tratamiento de datos personales, previamente alineado en el RAT

- Los estudios de impacto algorítmico llevados a término por el Consejo Asesor (sólo de los sistemas que acaben siendo licitados)

- El modelo de Cláusulas Tipo por la Contratación de Sistemas de IA Fiables

- Los informes de auditoría algorítmica

- Informe de balance final

- Información clara y detallada de los mecanismos de comunicación ciudadana disponibles para tramitar peticiones de información o consultas, incidencias, quejas y sugerencias en relación al funcionamiento de los sistemas algorítmicos

Además, se “facilitará la participación de la sociedad civil, empresas y entidades [...] dando la oportunidad a todas las partes implicadas de debatir la aplicación de los sistemas algorítmicos y de conocer sus resultados” y “el IMI tendrán que informar sobre el desarrollo del sistema algorítmico y los miembros de los órganos podrán hacer preguntas sobre éste y debatirlos.”

Respecto de la “Auditoría y otros tipologías de inspección” y sólo para los sistemas de “riesgo alto”, se afirma que “El Ayuntamiento de Barcelona tendrá derecho a difundir públicamente las conclusiones del informe en el Registro de Algoritmos.”

El observatorio de inteligencia artificial en salud de Catalunya

En el marco de la estrategia de IA catalana, se ha creado el programa Salud/IA, insertándose dentro de este el Observatorio de IA en salud de Catalunya⁸⁵ cuyo objetivo es:

- investigar y difundir los avances en IA en el ámbito de la salud
- analizar los riesgos y oportunidades que estos avances generan

En este contexto se está creando actualmente un registro de algoritmos empleados en el ámbito de la salud. De manera muy similar al sistema seguido

⁸⁵ Salut/IA, “Observatorio de la IA en el ámbito de la salud”. Disponible en: <https://iasalut.cat/es/observatori/>

en otros registros, las personas responsables de los sistemas de IA deben rellenar un formulario que contiene información acerca del sistema.

Esta información se refiere, entre otras cuestiones:

1. Al tipo de tecnología que emplea el sistema (si es una IA más o menos avanzada),
2. Al nivel de automatización
3. Al ámbito de la salud en que se va a emplear: atención primaria (50%), hospitalaria (20,91%), socio-sanitaria (6,36%), salud mental (8,18%) u otros (14,55%)
4. A la especialidad médica (en la que más se aplican es en oncología)
5. Explicación del funcionamiento
6. Nivel de desarrollo (un 40% de los algoritmos utilizados se encuentran en fases avanzadas de aplicación, un 25% en fases intermedias de validación y un 35% están en fases iniciales de laboratorio). Hay 9 niveles.

Hasta la fecha se han registrado 107 algoritmos, pero la información detallada todavía no se ha hecho pública.

Uno de los elementos a destacar del sistema de transparencia de algoritmos que está implantando la Generalitat Catalana en el ámbito de la salud es que se inserta en una estrategia general de IA que está dotada de una cantidad significativa de recursos y de planificación.

Es decir, la intención de dotar de transparencia a los sistemas de IA no es solo puramente estética, sino que verdaderamente se trata de realizar un proceso continuo de control y evaluación de los sistemas algorítmicos, asegurando que se extraen las ventajas de la IA pero siempre de forma bastante controlada. Por esta razón, también se dispone de una cantidad significativa de documentos y guías que tratan de orientar la implantación de la IA desde una perspectiva ética.

IV. Qué información hay que facilitar: sobre “su diseño y funcionamiento”

1. Transparencia sobre los datos del sistema de IA (de entrenamiento, 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. La relevancia de los datos empleados en el proceso de entrenamiento, validación y prueba del sistema se refleja en el artículo 10 del AIA, pues este establece una serie de directrices generales referidas al tratamiento de los datos, haciendo especial hincapié, entre otras cuestiones, en la necesidad de atender a posibles

sesgos, la necesidad de formular supuestos pertinentes y la posibilidad de tratar categorías especiales de datos siempre que dicho tratamiento se haga con la finalidad de proteger los derechos fundamentales de las personas físicas.

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 utilizados⁸⁶.

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)⁸⁷.

Las Directrices para el sector público de Países Bajos entran si cabe en mayor detalle. Así, requiere que se documente y registre datos de entrada y su uso efectivo y relevancia⁸⁸, se describa la calidad de los datos y cómo se

⁸⁶ Ministerie van Binnenlandse, *Impact Assessment...* cit.

⁸⁷ ICO, *Algorithmic transparency data standard...* cit.

⁸⁸ Directrices para el sector público, Rijksoverheid, *Richtlijnen voor het toepassen van algoritmen...* cit. Se 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

examinó⁸⁹. De igual modo que se describa, registre y documenten las decisiones sobre la inclusión o no de determinados datos por su calidad o suficiencia⁹⁰.

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. I).

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”⁹¹. También la AEPD en su Guía de Auditoría de IA 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.”⁹²

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

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

⁸⁹ *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.”

⁹⁰ *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.”

⁹¹ AEPD, *Guía para el cumplimiento del deber de informar*, ... cit. pág. 15.

⁹² AEPD, *Requisitos para Auditorías* ... cit. págs. 15-16.

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)⁹³: 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)⁹⁴. 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)⁹⁵. 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 quién suscribe y Palma hemos prestado especial atención⁹⁶. 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 que estos datos generados por la IA que se vinculen a una persona particular son datos personales, aunque no hayan sido facilitados el interesado⁹⁷. 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 rectificación⁹⁸. Es por ello que es necesaria una regulación o criterios interpretativos de las autoridades regulatorias o judiciales.

2. *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,

⁹³ AEPD, *Adecuación al RGPD ... cit.* pág. 23.

⁹⁴ Grupo del Artículo 29, *Directrices sobre decisiones automatizadas ... cit.*

⁹⁵ AEPD, *Adecuación al RGPD ... cit.* pág. 23.

⁹⁶ Ver apartado 9, “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, Lorenzo (editor), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Thompson-Reuters Aranzadi, FIADI, Cizur, 2022. También Palma Ortigosa, Adrián, *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, 2021, págs. 185 y ss.

⁹⁷ Grupo del Artículo 29, *Directrices sobre decisiones... cit.* p.10.

⁹⁸ 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.

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.

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. l) 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 para el sector público⁹⁹ 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 Bajos¹⁰⁰ 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

⁹⁹ ICO, *Algorithmic transparency data standard...* cit.

¹⁰⁰ Ministerie van Binnenlandse, *Impact Assessment...* cit.

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 alternitud en la elección del sistema y de los algoritmos concretos.

En particular, con respecto a la información sobre alternativas tecnológicas, conviene destacar que sería conveniente ir elaborando un listado de mejores técnicas disponibles para los diferentes posibles usos de la IA en el ámbito público de manera paralela a la incorporación de sistemas algorítmicos al registro¹⁰¹.

Los criterios para el Sector público de Países Bajos, respecto de la arquitectura del sistema (2.8) señalan que se explique el método de análisis utilizado y se mida su precisión.¹⁰² También en el documento del estudio de impacto de aquel país se afirma que hay que dar información sobre cómo se entrena el algoritmo¹⁰³. El ICO 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 actualización¹⁰⁴.

La Directrices para el sector público de Países Bajos 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.”¹⁰⁵ Asimismo, que se “compruebe que estos módulos funcionan correctamente tanto por separado como en combinación”.¹⁰⁶

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”.

¹⁰¹ Soriano Aranz, A., *Data protection of the prevention of algorithmic discrimination. Protecting from discrimination and other harms caused by algorithms through privacy in the EU: possibilities, shortcomings and proposals*, Thomson Reuters – Aranzadi, 2021, pp. 283-284.

¹⁰² La documentación técnica describe el método o métodos utilizados

Se explicó el método de análisis utilizado.

Se midió y describió la precisión del método analítico. Directrices para el sector público, Rijksoverheid, *Richtlijnen voor het toepassen van algoritmen...* cit.

¹⁰³ Ministerie van Binnenlandse, *Impact Assessment...* cit.

¹⁰⁴ ICO, *Algorithmic transparency data standard...* cit.

¹⁰⁵ Hay un script "principal" en el que se llaman todos los códigos. Por ejemplo, para cada fuente de datos hay un código en el que se preparan los datos.

El código está organizado en módulos.

Cada módulo puede evaluarse por separado y combinarse. Directrices para el sector público, Rijksoverheid, *Richtlijnen voor het toepassen van algoritmen...* cit.

¹⁰⁶ Todos los módulos se han probado por separado, siempre que ha sido posible,

Todos los módulos se han probado en combinación con otros. Se trata de una parte estándar del desarrollo del modelo

Los módulos pueden probarse para comprobar su correcto funcionamiento tanto individualmente como en combinación. *Ibidem*.

El acceso al “código fuente”

El acceso al “código fuente” del sistema IA puede ser especialmente conflictivo. Cerrillo¹⁰⁷ 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 Bajos¹⁰⁸. La Recomendación UNESCO lo reserva para supuestos especiales en los que se “puede requerir también que se compartan códigos o conjuntos de datos.” Unesco (nº 39). La propuesta de normativa de la ELI regula las reservas al acceso al código fuente¹⁰⁹. Su artículo 8.3º propuesto, 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.”¹¹⁰ 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.¹¹¹ 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.

¹⁰⁷ A. Cerrillo i Martínez, “La transparencia de los algoritmos que utilizan las administraciones públicas”, *Anuario de Transparencia Local*, nº. 3, 2020, págs. 41-78, pág. 68.

¹⁰⁸ Ministerie van Binnenlandse, *Impact Assessment... cit.*

¹⁰⁹ European Law Institute (ELI), *Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration*, European Law Institute, Universidad de Viena, 2022, https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Model_Rules_on_Impact_Assessment_of_ADMSs_Used_by_Public_Administration.pdf

¹¹⁰ 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.

¹¹¹ M.E., Gutiérrez David, “Acceso al código fuente y a los algoritmos de las Administraciones inteligentes. Lecciones a partir de experiencias comparadas”, en L. Cotino Hueso y J. Castellanos (coords.), *Transparencia y explicabilidad de la inteligencia artificial*, Tirant lo Blanch, Valencia, 2022, ppxxxx.

V. Información sobre “el nivel de riesgo que implican”: la existencia, finalidad, incidencia en las decisiones públicas, la lógica y motivación comprensible

1. Información de que existe el sistema IA y de que no había mejores alternativas

La premisa es la “comunicación” de la existencia del sistema IA público, esencialmente como “derecho a saber que [las personas] están interactuando con un sistema de IA” y para ello los sistemas de IA se deben identificar como tales a los humanos.¹¹² La comunicación se desdobra en primer lugar, como un principio de “notificación cuando se interactúa con un sistema de IA” especialmente importante “cuando los sistemas de aprendizaje automático se utilizan en la esfera pública”.¹¹³

Es muy preocupante que tan siquiera sepamos qué sistemas de IA existen e impactan en la actuación y prestación de servicios así como en nuestros derechos¹¹⁴. 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 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. 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 alguna¹¹⁵.

A partir de la existencia del sistema, también cabe 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

¹¹² HLEG, *Directrices éticas...* cit. N° 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.”

¹¹³ J. Fjeld, et. al., “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. January, 2020, pág. 4, <https://dash.harvard.edu/handle/1/42160420>

¹¹⁴ Sobre el tema dedico parte del estudio en L. Cotino Hueso, “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 Carles Ramió (coord.), *Repensando la administración digital y la innovación pública*, Instituto Nacional de Administración Pública (INAP), Madrid, 2021. <https://links.uv.es/FUW2pz6>

¹¹⁵ M.E., Gutiérrez David, “Administraciones inteligentes cit. ver págs. 29-30, 2021, <http://www.derecom.com/derecom/>

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)".¹¹⁶

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"¹¹⁷. En esta línea, las Directrices para el sector público de Países Bajos¹¹⁸ 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".

2. *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 mismos¹¹⁹. Un elemento esencial es la finalidad prevista del sistema de IA (artículo 3. 12^o¹²⁰), 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

¹¹⁶ ICO, *Algorithmic transparency data standard...* cit. detallando la información de segundo nivel.

¹¹⁷ AEPD, *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, págs. 15-16 <https://www.aepd.es/es/media/guias/requisitos-auditorias-tratamientos-incluyan-ia.pdf>

¹¹⁸ Directrices para el sector público, Rijksoverheid, *Richtlijnen voor het toepassen van algoritmen...* cit.

¹¹⁹ Así M.E., Gutiérrez David, "Administraciones inteligentes y acceso..." cit. págs. 55-56

¹²⁰ 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."

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. 3º 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 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 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”¹²¹.

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 tratamiento de datos¹²². 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.¹²³ 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, AEPD)¹²⁴, así como su impacto consecuencias del análisis para los ciudadanos afectados¹²⁵. 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).

También, en el marco francés, el artículo 6 de la Ley n ° 2016-1321 (*sur la république numérique*) reformó el Código de relaciones entre el público y la

¹²¹ “Por ejemplo, la cantidad de decisiones tomadas por mes o la cantidad de ciudadanos que interactúan con la herramienta (2.5)”. ICO, *Algorithmic transparency data standard...* cit.)

¹²² AEPD *Guía para el cumplimiento del deber de informar*, 7.2 Epígrafe “Finalidad”, pp. 10 y ss. <https://www.aepd.es/es/media/guias/guia-modelo-clausula-informativa.pdf>

¹²³ Ministerie van Binnenlandse, *Impact Assessment...* cit.

¹²⁴ En la concreción de la información y garantías específicas en razón del artículo 22 RGPD, Grupo del Artículo 29, *Directrices sobre decisiones automatizadas de 6 de febrero de 2018*, <https://www.aepd.es/sites/default/files/2019-12/wp251rev01-es.pdf>. Destaca la concreción de la AEPD, *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, de febrero 2020, pág 22. <https://www.aepd.es/media/guias/adecuacion-rgpd-ia.pdf>

¹²⁵ Directrices para el sector público, Rijksoverheid, *Richtlijnen voor het toepassen van algoritmen...* cit.

Administración, de manera que, el art. 312-1-3 de esta norma fija, para la mayor parte de Administraciones, la obligación de “publicar en línea las reglas que definan los principales procesos algorítmicos utilizados en el ejercicio de sus funciones cuando constituyan la base de decisiones individuales”.

En este sentido, 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?” (2022: 18).

3. 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

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”.¹²⁶

Concreta Gutiérrez 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 técnica¹²⁷.

El ICO 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”¹²⁸. 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”¹²⁹. 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

¹²⁶ J. Fjeld, et. al., “Principled Artificial Intelligence... cit. “Notification when AI Makes a Decision about an Individual Regular Reporting Requirement”.

¹²⁷ M.E., Gutiérrez David, “Administraciones inteligentes y acceso...cit. pág. 69.

¹²⁸ ICO, *Algorithmic transparency data standard...* cit.

¹²⁹ “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).” Nº 77, p. 22.

los parámetros, reglas e instrucciones en los que se basan los algoritmos o sistemas de 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.”¹³⁰

Estas referencias están claramente vinculadas al artículo 22 RGPD¹³¹ 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. ¹³², 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 2021¹³³ .

Como en otros lugares he tenido ocasión de examinar¹³⁴, 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 transparencia¹³⁵. Así, se da

¹³⁰ A este respecto, Ministerio de Trabajo y Economía Social, *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 2022, https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/trabajo14/Documents/2022/100622-Guia_algoritmos.pdf

¹³¹ De especial referencia al respecto son las “[Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679](#)” del ya extinto Grupo del Artículo 29, de 3 de octubre de 2017, revisadas el 6 de febrero de 2018.

¹³² 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>

¹³³ 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 o 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”.

¹³⁴ Sobre este “derecho”, me remito especialmente a mi trabajo “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, 2019, págs. 917-952, <http://links.uv.es/BmO8AU7>, así como a los trabajos de A. Palma Ortigosa, su mencionada tesis doctoral y la obra *Decisiones automatizadas y protección de datos personales. Especial atención a los sistemas de inteligencia artificial*, Dykinson, 2022

¹³⁵ En particular por cuanto a la transparencia en razón de este derecho, cabe remitir al exhaustivo trabajo M. Medina Guerrero, “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, 2022, págs. 141-171,

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).

En 2018 G29, luego revisado por el Comité Europeo de Protección de datos, detalló algo que la información que debe facilitarse¹³⁶. En 2020 la AEPD lo perfiló más para la IA en España¹³⁷, matizando en todo caso que “dependerá del tipo de componente IA utilizado”.

Para el sector público, las Directrices para el sector público de Países Bajos también han detallado la información a facilitar¹³⁸ así como en Alemania las del Comisionado Federal para la Protección de Datos¹³⁹. Más adelante se conectan

<https://dialnet.unirioja.es/descarga/articulo/8450038.pdf> También, A. Cerrillo i Martínez, “La transparencia de los algoritmos... cit. págs. 62 y ss.

¹³⁶ Grupo del Artículo 29, *Directrices sobre decisiones automatizadas* ... cit. pág. 35 detalla que la información que debe facilitarse es sobre:

-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.

¹³⁷ AEPD, *Adecuación al RGPD* ... cit. pág. 23 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.”

¹³⁸ Rijksoverheid, *Richtlijnen voor het toepassen van algoritmen*... cit. 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.”

¹³⁹ Se sigue el documento sobre la transparencia de la administración a la hora de utilizar algoritmos del Comisionado Federal para la Protección de Datos y Libertad de Información, “Transparenz der Verwaltung beim Einsatz von Algorithmen für gelebten Grundrechtsschutz unabdingbar”, en *36 Konferenz der Informationsfreiheitsbeauftragten in Deutschland*, págs. 3-4, https://www.datenschutzzentrum.de/uploads/informationsfreiheit/2018_Positionspapier-Transparenz-von-Algorithmen.pdf

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 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 perfiles realizados y sus implicaciones (AEPD)¹⁴⁰; 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)¹⁴¹. En el estudio de impacto IA de Países Bajos¹⁴² 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. La autoridad alemana 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.”¹⁴³

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.”¹⁴⁴

¹⁴⁰ AEPD, *Adecuación al RGPD ... cit.* pág. 23.

¹⁴¹ Grupo del Artículo 29, *Directrices sobre decisiones automatizadas ... cit.*

¹⁴² Ministerie van Binnenlandse, *Impact Assessment... cit.*

¹⁴³ Comisionado Federal para la Protección de Datos y Libertad de Información, “Transparenz der Verwaltung ... cit. págs. 3-4.

¹⁴⁴ AEPD, *Requisitos para Auditorías... cit.* págs.14-15.

4. Grado de influencia del sistema automatizado en la decisión que se adopte

Desde DAIA quien suscribe propuso que este apartado 5º se completara¹⁴⁵. 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 OdiselA¹⁴⁶, 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).

5. Transparencia y motivación comprensible de las decisiones públicas con IA

A lo largo del estudio se han 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

¹⁴⁵ Red DAIA, *Aportación a consulta de derechos digitales de julio, 2020* <https://www.dropbox.com/s/w9xlb64o1zemaca/DAIACONSULTAveniada.pdf?dl=0>

¹⁴⁶ OdiselA (Cotino Hueso, Lorenzo, coord.), *Documento propuestas mejora Carta Derechos Digitales*, diciembre 2021, <https://links.uv.es/Twqnt2U>

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.”¹⁴⁷

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 Gamero¹⁴⁸ 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 todos¹⁴⁹. 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 ubica la obligación de motivación algorítmica en el derecho a la buena Administración¹⁵⁰

Como Martín Delgado apunta también con DAIA 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”¹⁵¹. Se trata del

¹⁴⁷ Ver mi estudio sobre dicho apartado XVIII “Derechos ante la Administración digital y la inteligencia artificial” en L. Cotino Hueso, (editor), *La Carta de Derechos Digitales*, Tirant Lo Blanch, Valencia, 2022.

¹⁴⁸ E. Gamero Casado, “Necesidad de motivación e invalidez de los actos administrativos sustentados en inteligencia artificial o en algoritmos”, *Almacén de Derecho*, 4 febrero 2021.

¹⁴⁹ Se sigue en general, E. Gamero Casado y S. Fernández Ramos, *Manual básico de Derecho Administrativo*, (13ª ed.), Tecnos, 2016. págs. 526-527.

¹⁵⁰ FRA, *Artificial Intelligence and Fundamental Rights*, Agencia de los Derechos Fundamentales de la Unión Europea, 2020, p.81, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-artificial-intelligence_en.pdf

¹⁵¹ A. Cerrillo y otros, *Carta de Derechos digitales y sector público: propuestas de mejora*. Aportaciones de la Red DAIA a la Carta en el marco de la Consulta pública Subdirección General para la Sociedad Digital, diciembre 2020, <https://www.dropbox.com/s/gsbirh4kcn6vIk0/DAIAcartadic20.pdf?dl=0>

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 incurren 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. Gutiérrez¹⁵² o Ponce¹⁵³ detallan estas exigencias vinculadas a garantías de defensa o buena Administración.

Añade Gamero 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

¹⁵² Y desde OdiselA, E. 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.” En L. Cotino Hueso (coord.), *Comentarios y propuestas de modificación a la Carta de Derechos Digitales*, OdiselA, Consulta pública, diciembre 2020 <https://www.dropbox.com/s/6kydkk18eb950k7/consultaOdiselAv6.pdf?dl=0>

¹⁵³ En la consulta para la Carta, desde DAIA 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. A. Cerrillo y otros, *Carta de Derechos digitales y sector público...cit.*

las multas de velocidad¹⁵⁴. 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.

También en este sentido, debemos referirnos nuevamente a las inferencias resultantes de los sistemas de IA. Wachter y Mittelstadt consideran que resulta necesario el reconocimiento de un derecho a inferencias razonables de forma que se asegure la relevancia de los datos empleados para analizar y predecir un aspecto de la realidad social o comportamiento humano en el contexto de procesos automatizados o semi-automatizados de toma de decisiones¹⁵⁵. Es decir, las predicciones que sirven de base a la decisión que se toma deben partir de información cuya relación con el comportamiento analizado sea verificable.

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 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.

Por otra parte, con respecto al sesgo de automatización, son varios los estudios que apuntan a que la supervisión humana de procedimientos automatizados por especialistas en la materia en relación con la cual se emplea el sistema algorítmico o de IA, puede prevenir de manera significativa la aparición de esa tendencia a dar por válidos los resultados o predicciones propuestos por la máquina¹⁵⁶.

¹⁵⁴ 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.

¹⁵⁵ Wachter, S. y Mittelstadt, B., “A Right to Reasonable Inferences Re-Thinking Data Protection Law in the Age of Big Data and AI”, *Columbia Business Law Review*, vol. 2019, núm. 2., pp. 494-620.

¹⁵⁶ De-Arteaga, M. Flogliato, R. & Chouldechova, A., “A case for humans-in-the-loop: decisions in the presence of erroneous algorithmic scores”, in CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, New York, Association for computing machinery, 2020, pp. 1-12.

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.¹⁵⁷ 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”¹⁵⁸. 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 2017.¹⁵⁹ 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 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

¹⁵⁷ Por todos, M.E., Gutiérrez David, “Acceso al código... cit. También, A. Cerrillo i Martínez, “La transparencia de los algoritmos... cit. En el mismo libro, S. Palumbo, “Algoritmo y transparencia de la administración pública en el ordenamiento jurídico italiano. Algunas reflexiones críticas a partir del caso ‘Buona Scuola’”. Diana-Urania Galetta, “Derechos y garantías concretas respecto del uso por los poderes públicos de decisiones automatizadas e inteligencia artificial en el procedimiento administrativo”, en Cotino Hueso, Lorenzo (editor), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Thompson-Reuters Aranzadi, Cizur, 2022, págs. 79-93.

¹⁵⁸ A. Cerrillo Martínez, ¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?, *European Review of Digital Administration & Law* Volume 1 – Issue 1-2 – 2020, DOI: 10.4399/97888255389603, págs., 18-36, ver 31 y ss.

¹⁵⁹ 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

<https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000034194929#:~:text=Le%20d%C3%A9cret%20pr%C3%A9cise%20les%20modalit%C3%A9s,en%20application%20des%20articles%20R.>

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 decisiones¹⁶⁰. 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 *[mencione 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, 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.

VI. Información 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° 39) apunta la transparencia

¹⁶⁰ Las otras dos condiciones son que (2°) la decisión debe ser recurrible administrativamente. Y (3°) el Consejo Constitucional confirma la prohibición específica de decisiones exclusivamente automatizadas si incluyen el uso de datos especialmente protegidos. Es más, el nuevo artículo 10.1° Loi n° 78-17 de 6 de enero de 1978 prohíbe las decisiones semiautomatizadas en el ámbito de la justicia.

“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.

En su caso, habrá de informarse de la existencia de instrumentos de análisis de riesgos realizados, como puedan ser los relativos a la protección de datos, análisis de riesgos, estudios de impacto.

De igual modo, en razón de la normativa interna aplicable a la Generalitat Valenciana, cabría hacer referencia a la categorización o no del sistema como un sistema automatizado y, en su caso, a la existencia de informes o instrumentos exigibles respecto de estos sistemas en razón de las medidas de seguridad, gobernanza o regulación interna.

Según se ha visto, será de especial interés informar, tal y como se afirma para el caso de Barcelona de:

- la metodología utilizada para identificar los riesgos,
- los riesgos identificados y las medidas adoptadas para mitigarlos.

De igual modo, en la “información básica” se afirma que incluirá la relativa a “Riesgos que podrían aparecer y estrategia de gestión de estos”.

Y especialmente hay que tener en cuenta que en el caso de Barcelona que exige un específico estudio de impacto (no el de protección de datos regulado en general) se obliga a difundir a la ciudadanía “Los estudios de impacto algorítmico llevados a término por el Consejo Asesor (sólo de los sistemas que acaben siendo licitados)”. También los “informes de auditoría algorítmica”.

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 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 disponible¹⁶¹. 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 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

¹⁶¹ El nombre de la evaluación (5.1); una descripción de la evaluación de impacto realizada (5.2); la fecha en que completó la evaluación (5.3); un enlace a la evaluación o un resumen de la evaluación, si está disponible (5.4). ICO (Information Commissioner Office), *Algorithmic transparency data standard*... cit.

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).¹⁶² 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)¹⁶³. 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”.

Transparencia de las medidas y mitigaciones frente a riesgos

Respecto de las medidas y mitigaciones frente a riesgos, también el ICO 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)¹⁶⁴. Y que se “proporcione una descripción detallada de las acciones que ha tomado para mitigar los riesgos” (5.7)¹⁶⁵. En esta línea, la AEPD 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.”¹⁶⁶

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).

También el registro holandés de algoritmos incluye, para cada sistema cuya información se publica, apartados referentes al sistema de control, el nivel de

¹⁶² ELI, *Model Rules ... cit.*

¹⁶³ 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.

¹⁶⁴ 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”.

¹⁶⁵ ICO, *Algorithmic transparency data standard... cit.*

¹⁶⁶ AEPD, *Requisitos para Auditorías... cit.* págs.14-15.

intervención humana y el análisis de riesgos y evaluaciones de impacto realizadas.

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. (3.5)”.¹⁶⁷

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”.¹⁶⁸ 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.”

VII. Información sobre “el punto de contacto al que poder dirigirse en cada caso” y otra información sobre responsables, proveedores y usuarios, contratación del sistema

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 I).

Una interpretación estricta de este precepto llevaría a la mínima información de un medio de contacto a quien dirigirse por la ciudadanía. No obstante, esta exigencia debe interpretarse y puede completarse a la luz de otras regulaciones concurrentes.

Así, hay que ponerla en relación con:

- la obligatoriedad de facilitar información de responsable de tratamiento de datos y del delegado de protección de datos.

-Divulgación del responsable del sistema y aplicación y en su caso, del informe preceptivo realizado en razón del Decreto 220/2014, de 12 de diciembre de Administración Electrónica valenciana

-En su caso, identificación de los responsables en razón del Decreto 130/2012, de 24 de agosto, del Consell, por el que se establece la organización de la seguridad de la información de la Generalitat

-Asimismo, debería incluirse información sobre quiénes son los proveedores de sistemas de IA y quiénes los utilizan en la Administración.

¹⁶⁷ ICO, *Algorithmic transparency data standard...* cit.

¹⁶⁸ AEPD, *Adecuación al RGPD ... cit.* pág. 23.

1. Información de responsable de tratamiento de datos y del delegado de protección de datos

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 público 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 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.

Por otra parte, el RGPD obliga a facilitar “los datos de contacto del delegado de protección de datos, en su caso”. Se trata únicamente de los datos de contacto, no necesariamente la identidad de la persona física o jurídica que ocupe estos datos, sin perjuicio de que se pueda facilitar en razón de tratarse de información difundible.

2. Divulgación del responsable del sistema y aplicación y en su caso, del informe preceptivo realizado en razón del Decreto 220/2014, de 12 de diciembre de Administración Electrónica valenciana

Decreto 220/2014, de 12 de diciembre, del Consell, por el que se aprueba el Reglamento de Administración Electrónica de la Comunitat Valenciana

El artículo 6. 2º reconoce el derecho “a conocer el órgano o unidad que tenga encomendada la responsabilidad del correcto funcionamiento de las aplicaciones, programas y comunicaciones realizadas por medios electrónicos”¹⁶⁹.

Y con relación a ello, cabe recordar que “La gestión, administración y modificación de las aplicaciones y sistemas de información que se empleen

¹⁶⁹ “2. Además, los ciudadanos tienen, en relación con la utilización de los medios electrónicos en la actividad administrativa, y en los términos previstos en el presente decreto, los siguientes derechos: [...]”

d) Derecho a conocer el órgano o unidad que tenga encomendada la responsabilidad del correcto funcionamiento de las aplicaciones, programas y comunicaciones realizadas por medios electrónicos.”

corresponde al órgano directivo con competencias horizontales en materia de administración electrónica.” (artículo 25).

Asimismo, procede al aprobación previa de los “Los sistemas o aplicaciones informáticas” a través de “orden o resolución del órgano administrativo”, además previo el informe regulado en el artículo 94 (art. 78. 1º). La resolución debe “Establecer el órgano u órganos competentes” y fijar el “órgano que debe ser considerado responsable a los efectos de impugnación.” (art. 78. 2º)¹⁷⁰

Asimismo está regulado el “catálogo de sistemas de información y aplicaciones aprobadas de la Generalitat”, que es un instrumento de información “facilitará información tanto funcional como técnica de sus elementos catalogados”, especialmente con relación a la posible “reutilización del software de la Generalitat”, lo cual puede ser relevante.¹⁷¹

Asimismo, hay que tener en cuenta la referida “Aprobación de aplicaciones y sistemas de información” (Artículo 96) con informe preceptivo relativo a la oportunidad, adecuación a estándares, funcionalidad, órgano competente, seguridad, normalización de los medios de acceso, conservación de los soportes utilizados, calidad¹⁷².

¹⁷⁰ Artículo 78. Sistemas de información para la gestión de actuaciones administrativas automatizadas

1. Los sistemas o aplicaciones informáticas que vayan a ser utilizados para la gestión de actuaciones administrativas automatizadas deberán ser previamente aprobados mediante orden o resolución del órgano administrativo que tenga atribuida la competencia para la resolución del procedimiento, previo informe de administración electrónica dictado según el artículo 94 del presente decreto y cuyo contenido será vinculante.

2. Además de lo dispuesto en el artículo 94 del presente decreto, la resolución u orden aprobatoria del sistema o aplicaciones informática deberá:

a) Establecer el órgano u órganos competentes, según los casos, para la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente.

b) Indicar el órgano que debe ser considerado responsable a los efectos de impugnación.

¹⁷¹ Artículo 95. El catálogo de sistemas de información y aplicaciones aprobadas de la Generalitat

1. El órgano directivo con competencias horizontales en materia de tecnologías de la información y las comunicaciones de la Generalitat definirá, creará y gestionará el catálogo de sistemas de información y aplicaciones de la Generalitat, incluyendo las conocidas como apps. Dicho catálogo facilitará información tanto funcional como técnica de sus elementos catalogados.

2. Una utilidad fundamental del catálogo será facilitar la reutilización del software de la Generalitat. A tal efecto, el catálogo nacerá como un instrumento de uso interno, a disposición de los departamentos y organismos de la Generalitat y las administraciones locales valencianas. En su momento, el órgano competente en materia de tecnologías de la información y las comunicaciones establecerá una política de publicación externa de la totalidad o de un subconjunto del catálogo a los efectos de reutilización por otras administraciones.

¹⁷² Capítulo II. Aprobación y homologación de aplicaciones

Artículo 96. Aprobación de aplicaciones y sistemas de información

1. En el ámbito de los sujetos referidos en el número 1 del artículo 2 del presente decreto, los sistemas de información y las aplicaciones que se utilicen para el ejercicio de potestades administrativas habrán de ser previamente aprobados por el titular de la conselleria o el director o presidente del organismo autónomo que sea competente en el ámbito material al que se refiera la actuación. Para dicha aprobación se requerirá el informe preceptivo y favorable regulado en este artículo y emitido por el órgano competente. En caso de informe negativo, este será fundamentado respecto a los criterios establecidos en el apartado siguiente.

2. El informe preceptivo analizará, valorará y se pronunciará sobre los siguientes aspectos:

Pues bien, no cabe duda de que es preceptiva la fijación del responsable y el derecho a conocer el mismo por la ciudadanía. Por cuanto a la aprobación de la aplicación, en el caso de ser una Orden, sin duda será publicada en el DOCV. Puede dudarse en el caso de que la aprobación se realice por resolución, en este supuesto puede no publicarse la misma. El informe preceptivo no tiene por qué ser publicado al no ser componente de la orden. No obstante, sería de interés la posibilidad de acceder por publicidad activa al informe, a salvo de motivos de limitación de acceso al mismo.

Se recomienda la publicidad activa del órgano o unidad competente, esto es que tenga encomendada la responsabilidad del correcto funcionamiento de las aplicaciones. Asimismo, que se difunda igualmente el “órgano que debe ser considerado responsable a los efectos de impugnación.” (art. 78. 2º).

Sería especialmente interesante divulgar activamente el informe preceptivo de la aplicación o sistema utilizado, en su caso limitando la información que se considere justificadamente que no debe divulgarse.

3. En su caso, identificación de los responsables en razón del Decreto 130/2012, de 24 de agosto, del Consell, por el que se establece la organización de la seguridad de la información de la Generalitat

Cabe tener en cuenta la normativa específica de la Generalitat Valenciana. De la misma puede tener relevancia las diversas figuras de “agentes de la organización de la seguridad” que ahí se recogen, como (art. 6), la organización estará compuesta por:

- Responsable de la Información.
- Comité de Seguridad de la Información.
- Responsable de los Ficheros de Datos de Carácter Personal.

a) Oportunidad. Analizará si el sistema de información o aplicación cuyo desarrollo se solicita puede ver sus funcionalidades previstas satisfechas por un sistema de información unificado ya existente o en proyecto. En caso contrario, se valorará si el desarrollo solicitado es consistente con la estrategia y la planificación de la Generalitat. Si el resultado de ambos análisis es positivo, se considerará cumplido el requisito.

b) Adecuación a estándares. Verificará que sus especificaciones técnicas se sujeten a los estándares y criterios establecidos en el artículo 93, apartado 7, del presente decreto.

c) Funcionalidad. Valorará la adecuación de las funcionalidades requeridas a los requisitos funcionales establecidos previamente por el órgano que la utilizará para el ejercicio de sus competencias.

d) Órgano competente. Procederá a la identificación y comprobación del ejercicio de la competencia por el órgano correspondiente.

e) Seguridad. Verificará la aplicación de las medidas de seguridad contempladas en el Esquema Nacional de Seguridad en el ámbito de la administración electrónica, en los decretos 66/2012, de 27 de abril, y 130/2012, de 24 de agosto, ambos del Consell, así como las medidas establecidas en la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal, y su normativa de desarrollo.

f) Normalización de los medios de acceso. Verificará las especificaciones técnicas sobre los medios, códigos y formatos de acceso.

g) Conservación de los soportes utilizados. Valorará la proporción entre la durabilidad de los soportes y el tiempo en que deben mantenerse los datos en ellos incluidos.

h) Calidad. Valorará el proceso de desarrollo e implantación de la aplicación y su adecuación a los estándares establecidos.

- Responsable del Servicio.
- Responsable de Seguridad de la Información.
- Responsable de Seguridad de los Ficheros de Datos de Carácter Personal.
- Responsable del Sistema.
- Administradores de la Seguridad del Sistema.
- Administradores de la Seguridad de los Ficheros de Datos de Carácter Personal.

En su caso, habrá que valorarse la posible difusión activa de la información, bien su contacto, en su caso y también, la identidad.

4 .Debería incluirse información sobre quiénes son los proveedores de sistemas de IA y quiénes los utilizan en la Administración

No es momento ahora de detallar la compleja cadena de valor en la IA. Entre los distintos “operadores” de los sistemas IA¹⁷³, 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 distribuidor¹⁷⁴, 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 a especificar la información sobre “a) la identidad y los datos de contacto del proveedor y, en su caso, de su representante autorizado”.

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

¹⁷³ 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.”

¹⁷⁴ 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.”

ahora interesa, se incluye información del proveedor del sistema IA, en su caso del representante¹⁷⁵.

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).

Consideramos plenamente oportuno que se detalle la información sobre el proveedor del sistema y los instrumentos de contratación que han llevado a dicha relación. Si es un desarrollo propio, debe también detallarse.

5. Además, información sobre 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”¹⁷⁶. Se sigue en este sentido a *Access Now* 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.”¹⁷⁷ El tema de la contratación por el sector público de sistemas de IA bien merece un análisis jurídico detallado¹⁷⁸ 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, cabe ahora referir que procedería dar información sobre cómo se han adquirido e implementado esos sistemas de

¹⁷⁵ 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.”

¹⁷⁶ J. Fjeld, et. al., “Principled Artificial Intelligence... cit.

¹⁷⁷ Access Now, “Human Rights in the Age of Artificial Intelligence”, Access Now, 2018, nº 9, pág. <https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf>

¹⁷⁸ Para ello, entre otros G. Vestri, “Sistemas de inteligencia artificial en la contratación pública: entre códigos fuente y datos abiertos”, *Actualidad administrativa*, Nº 12, 2021. También, J. Miranzo Díaz, “Inteligencia artificial y contratación pública”, en I. Martín (Dir.), *Administración electrónica, transparencia y contratación pública*, INAP, 2020, págs. 105-142.

decisión algorítmica (mediante desarrollos *in house* o mediante licitación)¹⁷⁹. 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 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.”

Ya se ha hecho referencia en la conveniencia de informar sobre los proveedores de sistemas de IA y quiénes los utilizan en la Administración. Además, no puede obviarse que hay no pocas obligaciones de publicidad activa en el ámbito de la contratación.

VIII. Límites

Se hará referencia a los límites de la transparencia artículo 14 (propiedad intelectual, seguridad, etc.) u otros que puedan aplicar.

En la guía del gobierno del Reino Unido sobre transparencia algorítmica¹⁸⁰ se detallan las tres razones principales que justifican la no publicación de parte de la información referente a los sistemas de inteligencia artificial:

- Eficacia operativa y evitación de “trampas”: puede justificarse la limitación de la información que se publica sobre el sistema en aquellos casos en que proporcionar demasiada información sobre cómo funciona una herramienta algorítmica, o los detalles de los conjuntos de datos en los que se basa, podría comprometer la eficacia operativa de la herramienta. Por ejemplo, un usuario malintencionado podría modificar su comportamiento para evitar activar una alerta de fraude.
- Riesgos de ciberseguridad.
- Protección de los derechos de propiedad intelectual.

En todo caso el informe insiste en que resulta necesario publicar toda la información que sea posible y limitar la información no publicada a aquellos elementos cuya ocultación sea necesaria para evitar los riesgos indicados. Asimismo, cuando se omita la publicación de parte de la información, debe indicarse las razones por las que la transparencia se encuentra limitada.

¹⁷⁹ M.E., Gutiérrez David, “Administraciones inteligentes y acceso al código fuente... cit. pág. 48.

¹⁸⁰ Gov.uk, “Algorithmic Transparency Recording Standard - Guidance for Public Sector Bodies”, *cit.*

No es posible en este estudio adentrarse en el terreno de los límites a la transparencia algorítmica. Como punto de partida, potencialmente de forma voluntaria las administraciones pueden dar difusión activa a toda la información algorítmica “cuyo conocimiento sea relevante para garantizar la transparencia de su actividad relacionada con el funcionamiento y control de la actuación pública.” (art. 5. 1º Ley 19/2013). En todo caso, “serán de aplicación, en su caso, los límites al derecho de acceso a la información pública previstos en el artículo 14 y, especialmente, el derivado de la protección de datos de carácter personal, regulado en el artículo 15.” (art. 5. 3º Ley 19/2013). En principio, en todos los casos de uso público de IA, ya sean desarrollos propios o de desarrollos para el sector público, 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”). No obstante, puedan concurrir causas de inadmisión (art. 18) o límites aplicables (arts. 14 y 15). Es por ello aconsejable que la legislación regule las obligaciones de publicidad activa y habilite o esclarezca qué información debe difundirse proactivamente, aclarando posibles dudas que puedan surgir. Resultará especialmente aconsejable que la transparencia incluya los motivos por los que se considera que una información concreta de por qué no debe ser difundida. Sobre el tema cabe seguir, Boix (2022), Gutiérrez (2021), Cerrillo (2020 a), Vestri (2021 b y 2022). Las autoridades en su caso judiciales también podrán requerir información cuando se considere preciso, sin perjuicio de las fórmulas de secreto y reserva que puedan concurrir. Cabe señalar que - en ocasiones lamentablemente- no es extraño que las autoridades públicas consideren esta información algorítmica materia clasificada, por lo que concurrirá la normativa de secretos oficiales y se hará mucho más complejo el acceso a esta información.¹⁸¹

Resumen comparado de la información proporcionada en los registros algorítmicos ya establecidos o propuestos

ALGORITHMIC TRANSPARENCY STANDARD	ICO	ÁMSTERDAM Y HELSINKI	PAÍSES BAJOS	OBSERVATORIO IA EN SALUD (CATALUÑA)	GUÍA ETALAB FRANCIA
INFORMACIÓN BÁSICA	INFORMACIÓN GENERAL	VISIÓN GENERAL	INFORMACIÓN GENERAL	DATOS BÁSICOS DEL ALGORITMO	SOBRE LA ADMINISTRACIÓN
Nombre del algoritmo	Nombre del algoritmo	Objetivo principal al que contribuye el algoritmo	Nombre el algoritmo	Nombre del algoritmo	Nombre de la administración responsable
Descripción corta	Descripción breve	¿Cómo se está utilizando?	Organización responsable	Procedimiento en el que se enfoca el algoritmo	Dirección/servicio

¹⁸¹ Me remito a la investigación coordinada por el equipo liderado por Lucía Martínez Garay Universidad de Valencia-Amnistía Internacional sobre sistemas IA públicos del ámbito policial y penitenciario. En concreto, derechos de acceso ejercidos por uno de los miembros del equipo frente a la Guardia Civil (Exp. Transp. nº 062893) y los Mossos d'Esquadra (19 de noviembre de 2021, respuesta 18 de enero de 2022). En el caso de la Guardia Civil, se afirma que estos sistemas de identificación biométrica quedan afectados por Acuerdo del Consejo de Ministros de 28 de noviembre de 1986, ampliado por Acuerdos de 17 de marzo y 29 de julio de 1994, otorga el carácter de "Reservado" a aquella información relativa a "las plantillas de personal y medios y de equipo de las Unidades". En el caso de solicitud de información sobre Sistema Automático de Identificación Biométrica (ABIS) (nº expediente: 001-062892, la denegación es en razón del artículo 14 Ley 19/2013.

Organización responsable	Página web	¿En qué casos se utiliza?	Departamento	Qué tecnologías utiliza el algoritmo	Contacto
Departamento	E-mail de contacto	¿Sobre quién tiene impacto?	Descripción corta	Responsables del algoritmo	Fecha de actualización de la información
E-mail de contacto	PROPIEDAD Y RESPONSABILIDAD	¿Cuáles son los impactos esperados?	Tipo de algoritmo	Correo electrónico de los responsables	ALGORITMO Y DECISIÓN TOMADA
Número de teléfono de contacto	Administración	¿Cómo funciona el sistema a un alto nivel?	Dominio en el que se despliega el algoritmo	Teléfono de contacto de los responsables	Nombre del algoritmo
Nombre de contacto	Órgano y, en su caso, servicio o equipo concreto responsable	RENDICIÓN DE CUENTAS	Enlace a web de publicación	¿Algoritmo asociado a algún proyecto?	Contexto global
Área geográfica	Persona responsable dentro del órgano	Organizaciones responsables	Estado	Desarrolladores del algoritmo	Finalidad algoritmo
Dominio	Participación de proveedores externos e identificación	Departamentos responsables	APLICACIÓN	Correo electrónico de los desarrolladores	Cómo interviene el algoritmo en la decisión
Página web	Nivel de intervención del proveedor externo	Persona, email y teléfono de contacto	Objetivo	Teléfono de contacto de los desarrolladores	Cuando interviene el algoritmo en la decisión
Estado	Tipo de procedimiento de contratación pública	Socios y proveedores externos que intervienen	Impacto	¿Otra entidad involucrada?	Nivel de automatización de la decisión
Observaciones	Régimen de acceso a datos de los proveedores externos	CONJUNTO DE DATOS	Proporcionalidad	¿Algún centro de investigación involucrado?	Fundamentos jurídicos
ÁMBITO DE APLICACIÓN Y CONTROL	DESCRIPCIÓN Y JUSTIFICACIÓN	Nombre del conjunto de datos	Proceso de toma de decisiones	Propiedad del algoritmo	Recursos relacionados
Proceso de toma de decisiones	Descripción detallada	Descripción de la información representada en el conjunto de datos	LEGAL	Propiedad intelectual del algoritmo	IMPACTO DE LA DECISIÓN
Objetivos	Ámbito de aplicación	PROCESAMIENTO DE DATOS	Organismo autorizado para el despliegue del algoritmo	Derechos por el uso del algoritmo	Número de decisiones administrativas tomadas por año
Impacto	Beneficios de la utilización del sistema	Arquitectura modelo del sistema	Base legal	¿Algún proveedor externo involucrado?	Alcance de la decisión
Riesgos y medidas	Explicación de cómo se llevaba el proceso a cabo anteriormente	Comportamiento de sistema	Evaluación de impacto en los derechos humanos (IAMA)	ÁMBITO DE APLICACIÓN, OBJETIVOS E IMPACTO	Público afectado por la decisión
Categoría de riesgo (nivel de riesgo)	Alternativas consideradas	NO DISCRIMINACIÓN	Evaluación de impacto de protección de datos (DPIA)	Ámbito de actuación del algoritmo	Recursos relacionados
Proporcionalidad	PROCEDIMIENTO DE TOMA DE DECISIONES	¿Cómo se interpreta el sesgo injusto en el contexto del sistema?	Procedimiento de reclamación	Objetivo principal	FUNCIONAMIENTO INTERNO DEL ALGORITMO
Base legal	Forma en la que el algoritmo se integra en el procedimiento	¿Qué medidas existen para comprobar el sesgo?	METADATOS	Objetivos específicos	Datos tratados
Más información	Información proporcionada por el algoritmo en el marco del procedimiento	¿Es el sistema diseñado e implantado teniendo en cuenta la accesibilidad del servicio para personas con discapacidades?	Enlace a la web del registro de origen	Nivel de automatización	Fuente de los datos
DATOS	Nivel de intervención y supervisión humana	Implicación de las personas afectadas por el sistema en el diseño y desarrollo	E-mail de contacto	ÁMBITO DE APLICACIÓN	Modo de recolección de los datos
Datos de entrenamiento	Formación específica proporcionada las personas que emplean el sistema	SUPERVISIÓN HUMANA	Área geográfica	Descripción del uso del algoritmo	Tipo de algoritmo

Fuente de los datos	Sistema de recursos, revisión y quejas	Capacidad y soporte para la intervención humana en el diseño y desarrollo del sistema, ciclos de decisión, y en el seguimiento del funcionamiento del sistema	Idioma	Explicación a alto nivel del funcionamiento tecnológico del sistema	Operaciones efectuadas por el algoritmo
Sesgo de los datos	ESPECIFICACIONES TÉCNICAS Y DATOS	Competencias necesarias para la supervisión	ÁMBITO DE APLICACIÓN	Especialidad médica a la que se aplica el algoritmo	Recursos relacionados
Conexión con otras fuentes	Método usado por el sistema	Entrenamiento disponible para adquirir las competencias necesarias	Descripción larga	¿Algoritmo asociado a algún diagnóstico?	
Descripción de la evaluación de impacto de la protección de datos (DPIA)	Frecuencia de uso	RIESGOS	Uso de los registros base	Colectivo beneficiado	
Enlace al DPIA	Fase de desarrollo	Compensaciones clave entre riesgos y beneficios	Fuentes de datos	NIVEL DE MADUREZ DEL ALGORITMO	
ALGORITMO	Actuaciones necesarias para su mantenimiento	Nivel de riesgo	Métodos y modelos utilizados por el algoritmo	Fase del algoritmo	
Tipo de algoritmo	Rendimiento del sistema: incluyendo precisión, número de veces que una persona se ha apartado de la decisión recomendada y recursos contra las decisiones del sistema que han sido estimados	Métodos de gestión de los riesgos	SUPERVISIÓN	Nivel de madurez del algoritmo	
Explicación del funcionamiento	Arquitectura del sistema	Riesgos identificados	Monitorización	¿Está implementado el algoritmo?	
Métodos y modelos	Fuente de los datos de entrenamiento y funcionamiento	Medidas implementadas para mitigar los riesgos	Intervención humana	¿Es un proyecto en activo?	
Comportamiento del algoritmo	Descripción de los datos de entrenamiento y funcionamiento	Evaluaciones de impacto en la privacidad	Riesgos	¿Cuándo se ha iniciado?	
Alternativas consideradas	En su caso, página web donde se encuentran los datos	Evaluaciones de impacto en derechos humanos			
Enlace a la información central del producto (el algoritmo)	Forma en que se recogieron los datos	Otra evaluación de riesgos asociados a los métodos aplicados			
Enlace al código fuente	Procesos de limpieza de los datos	EXPLICABILIDAD			
Enlace a la página del producto	Análisis del nivel de integridad y representatividad de los datos	Razones que llevaron al resultado concreto con la aplicación del algoritmo			
Evaluación del impacto del algoritmo	Acceso y almacenamiento de datos por parte del proveedor del sistema	REFERENCIAS			
Enlace a la evaluación del impacto del algoritmo	RIESGOS MITIGACIÓN I EVALUACIONES DE IMPACTO	Referencias a documentación adicional : políticas de privacidad, contratos, informes de auditoría...			
SUPERVISIÓN	Evaluación de impacto: resumen y enlace de acceso a la evaluación completa				
Monitorización del comportamiento	Explicación de los riesgos y estrategias de mitigación				

Tipo de uso del algoritmo					
Intervención humana					
Procedimientos de reclamación					
Retroceso (posibilidad de anular los efectos de la aplicación del algoritmo)					
Metadatos					
Versión del estándar utilizado					
Identificador único					
Idioma					
Fecha de publicación					
Fecha de la última actualización					
Fecha de revisión del registro					
Nota sobre la revisión (revisión ordinaria, cambios en la ley...)					
Palabras clave					

Tercero. Cómo gestionar la gobernanza interna del registro de algoritmos

Adrián Palma y Jorge Castellanos

Las obligaciones de publicidad activa y el ejercicio del derecho de acceso que se desprenden de la implementación de la Ley 1/2022, de 13 de abril, obligan a la Generalitat Valenciana a elaborar un registro de los diferentes algoritmos que, en su caso, puedan quedar afectados por el ámbito de aplicación de la mencionada norma.

En este apartado se indican aquellas recomendaciones que consideramos necesarias para implementar un registro de algoritmos y las medidas de gobernanza de éste.

Teniendo en cuenta que actualmente las Administraciones Públicas valencianas ya están utilizando algoritmos a los que potencialmente les puede afectar esta ley, se indicará un procedimiento distinto para registrar los algoritmos que se utilizan y otro procedimiento para aquellos algoritmos respecto de los que aún no se dispone de ello, pero que en un futuro deberán estar integrados en tal registro.

I. Órganos de Gobernanza interna del Registro de Algoritmos

Para facilitar adecuadamente el despliegue del Registro de Algoritmos Públicos se proponen dos órganos administrativos que, en su caso, revestirán la forma jurídica que se estime oportuna.

Estos órganos/unidades son:

Oficina/Unidad	Oficina de registro de algoritmos (ORA)
Funciones principales	Esta oficina se encargaría entre otras funciones de: realizar la búsqueda de algoritmos, registrarlos, solicitar la información a las unidades y órganos administrativos titulares de los algoritmos, ayudar al Consejo del Registro de Algoritmos Públicos en la implementación de las medidas asignadas a éste, etc.
Miembros integrantes	Personal de la Conselleria con competencias en la materia (de Transparencia),

Oficina/Unidad	Consejo del Registro de Algoritmos Públicos de la Generalitat Valenciana. (CRAP)
Funciones principales	Entre otras funciones, el CRAP se encargaría de realizar las evaluaciones del riesgo, definir la información algorítmica que se ha de publicar, resolver las consultas propuestas por la ORA, etc.
Miembros integrantes	Este órgano estará presidido por el Conseller/a con competencias en la materia (de participación, transparencia, cooperación y calidad democrática); y el Director/a general con competencias en la materia (de transparencia, atención

	<p>a la ciudadanía y buen gobierno). Sus vocales se clasificarían de la siguiente manera:</p> <p>En representación de las universidades:</p> <ul style="list-style-type: none"> – Cinco representantes, uno por cada una de las universidades públicas valencianas, expertos en materia de derecho e inteligencia artificial. – Cinco representantes, uno por cada una de las universidades públicas valencianas, expertos en ciencia de datos. – Cinco representantes, uno por cada una de las universidades públicas valencianas, expertos en matemáticas. – Tres representantes, uno por provincia, pertenecientes a organizaciones de la sociedad civil. – Tres representantes, uno por provincia, pertenecientes a empresas tecnológicas. <p>El Delegado de protección de datos de la GVA, que actuaría, a su vez, como secretario</p>
--	--

II. Registro de algoritmos que ya se están utilizando

Actualmente, las AAPP valencianas están utilizando toda una serie de algoritmos en sus quehaceres cotidianos. La primera medida que se ha de adoptar por parte de las autoridades es tratar de recopilar y obtener información de dichos algoritmos.

Debido a la innumerable tipología y variedad de potenciales algoritmos que en su caso deberían incorporarse al mentado registro, exponemos a continuación toda una serie de medidas para facilitar esta labor.

Estas son:

- 1) Búsqueda de potenciales algoritmos;
- 2) Solicitud de información sobre los algoritmos;
- 3) Evaluación y asignación del Riesgo algorítmico y,
- 4) Registro de algoritmos.

1. Búsqueda de potenciales algoritmos

Corresponderá a la ORA realizar la búsqueda inicial de los potenciales algoritmos que actualmente se están utilizando por parte de las AAPP valencianas y que en su caso les puede resultar de aplicación las obligaciones de publicidad activa.

Proponemos dos métodos para facilitar esas tareas de búsqueda:

Auto búsqueda: se puede adoptar diferentes actuaciones:

- **Comprobar los Registros de Actividades de Tratamientos (RAT)**¹⁸² de datos de carácter personal para verificar si estos hacen mención a algún tipo de algoritmo o sistema de IA. Posiblemente esta información no se indicará de forma expresa en los RAT pero la información que aparece al describir el tratamiento de datos puede arrojar sospechas de que puede estar utilizándose algún algoritmo.

¹⁸² Véase el Registro de Actividades de Tratamiento de la Generalitat Valenciana.
<https://gvaoberta.gva.es/es/registre-activitats-de-tractament>

- **Conocimientos propios:** La ORA ya puede disponer de un catálogo interno o información sobre algunos algoritmos que se están utilizando.
- **Revisión de normativa, circulares, informes:** Se puede hacer un rastreo de los diferentes instrumentos jurídicos que se utilizan en el seno de la GVA para buscar potenciales menciones al uso de sistemas automatizados, algoritmos, etc.
- **Revisar el catálogo de sistemas de información y aplicaciones aprobadas de la Generalitat**¹⁸³.

Solicitud de ayuda/información al resto de órganos y unidades administrativas: Se pueden llevar a cabo diferentes actuaciones, entre otras:

- Envío de correos electrónicos avisando a las diferentes unidades.
- Orden directa a través de las distintas Consellerias informando de que se pongan en contacto con la ORA a efectos de que les comuniquen la existencia de sistemas algorítmicos.
- Uso de canales internos de información entre Consellerias.
- Colaboración con el órgano directivo con competencias horizontales en materia de administración electrónica al que hace referencia el artículo 25 del Decreto sobre Decreto 220/2014 a efectos de centralizar todas las peticiones de solicitud o facilitar los datos que consten sobre uso de sistemas automatizados¹⁸⁴.
- El subsecretario de cada una de las Consellerias centralizará las peticiones de información y coordinará el envío de la información pertinente para poder ordenar y trasladar la información relativa al registro

2. Solicitud de información algorítmica

Una vez se hayan recopilado los diferentes algoritmos y sistemas de IA que potencialmente pueden resultar afectados por las obligaciones de publicidad activa, el siguiente paso es que **la ORA envíe** a cada una de las unidades/órganos administrativos/personal responsable de los algoritmos una solicitud con un contenido mínimo.

Entre otros elementos de información algorítmica que debe integrar dicha solicitud consideramos los siguientes:

- Finalidad del sistema algorítmico:
- Alcance del sistema algorítmico:
- Datos que se utilizarán:
- Política de utilización:
- Principales riesgos de inexactitudes, sesgos o daños a comunidades específicas podrían aparecer.
- Medidas de mitigación de riesgos se utilizarán.
- Posibles intereses/bienes jurídicos afectados si se publica cierta

¹⁸³ Véase el artículo 95 del Decreto 220/2014, de 12 de diciembre, del Consell, por el que se aprueba el Reglamento de Administración Electrónica de la Comunitat Valenciana.

¹⁸⁴ Decreto 220/2014, de 12 de diciembre de Administración Electrónica valenciana.

información sobre el sistema algorítmico.

Dicha solicitud de información algorítmica debe ser lo suficientemente completa **para que el CRAP pueda** posteriormente evaluar el riesgo de los algoritmos y valorar si en su caso el nivel de transparencia algorítmica a desplegar resulta adecuado o puede perjudicar a otros bienes/intereses jurídicos en juego.

Corresponde a la **ORA** realizar todos los trámites administrativos posibles para conseguir que las unidades/órganos/responsables de los sistemas algorítmicos que se recopilaron inicialmente respondan a la solicitud de la información algorítmica requerida.

3. Evaluación y asignación del riesgo algorítmico

Tras la recopilación de la información algorítmica de los diferentes sistemas automatizados por parte de la ORA, el siguiente paso es que las diferentes solicitudes con la información ya contenida se envíe al CRAP.

A medida que el CRAP vaya recibiendo la información algorítmica de cada sistema, éste deberá proceder a evaluar su riesgo¹⁸⁵.

Para evaluar el riesgo, el CRAP deberá analizar los distintos riesgos que para los derechos y libertades de las personas y colectivos pueden estar presentes en el despliegue de la herramienta algorítmica objeto de evaluación.

Corresponde al CRAP asignar a cada sistema algorítmico un riesgo determinado, concretamente, **alto, medio y bajo**. A la hora de evaluar y asignar el riesgo, el CRAP puede requerir más información a los responsables de estos algoritmos a efectos de realizar un análisis certero de los riesgos. Para estas labores, el CRAP podrá requerir la ayuda de la ORA, la cual podrá gestionar el canal de comunicaciones.

3.1) Riesgo algorítmico alto

Para dicha evaluación remitimos al apartado de este informe que explica de forma pormenorizada los sistemas que a nuestro entender se consideran de alto riesgo¹⁸⁶. Resumidamente, se consideran de alto riesgo los siguientes sistemas:

- Sistemas calificables como de alto riesgo por el futuro RIA.
- Sistemas que identifican o priorizan objetivos para la aplicación de la ley o para realizar inspecciones en el ámbito de infracciones penales, administrativas y persecución de ilicitudes y fraude, cada vez más habituales en ámbitos de seguridad, mercado y competencia, trabajo, salud, cuidado del medio ambiente, etc.
- Sistemas para la adjudicación de contratos, subvenciones, privilegios del gobierno.
- Sistemas IA para personalizar, priorizar o apoyar la prestación de servicios a los ciudadanos en salud, educación, empleo, servicios

¹⁸⁵ Es importante destacar que la evaluación del riesgo también podría realizarla la ORA, sin embargo, las competencias asignadas a ésta deberían ser mayores que las encomendadas a través del modelo de gobernanza propuesto.

¹⁸⁶ Véase apartado “Criterios y variables para determinar el impacto, nivel de riesgo y relevancia jurídica de los sistemas algorítmicos públicos” de este informe.

sociales, etc.

- Sistemas de extracción de información, investigación, recopilación, supervisión y el análisis de datos, para la elaboración de políticas, la toma de decisiones, monitoreo general y análisis de riesgos.
- Sistemas para la gestión de la organización interna, recursos humanos y las adquisiciones y la gestión de los recursos tecnológicos.
- Sistemas para interactuar y comunicarse con el público sobre sus derechos y obligaciones y su participación.

3.2) Riesgo algorítmico medio

Quedaría integrado por todos aquellos algoritmos que existen dudas de si estos son considerados de alto riesgo o no. Para establecer tal calificación será preceptivo un **informe técnico** independiente que detalle aquellos elementos del algoritmo que hagan concluir que de su implementación no se deriva un riesgo potencial para los derechos de la ciudadanía. Si con la información relativa al algoritmo se puede concluir, previo informe técnico, que no hay una catalogación clara de alto riesgo, para mayor seguridad jurídica se integrará en esta categoría, sin perjuicio de una revisión posterior que permita su degradación a una categoría de riesgo algorítmico bajo.

La elaboración del informe deberá ser encomendada a personal técnico especializado y supervisado y aprobado por el CRAP.

3. 3) Riesgo algorítmico bajo

Estaría integrado por el resto de algoritmos que por los riesgos que genera a los derechos y libertades de los ciudadanos y ciudadanas, no existen dudas de su impacto reducido o nulo.

En función del riesgo determinado por el CRAP, se adoptarán entre otras las siguientes actuaciones:

Riesgo Alto	<ul style="list-style-type: none">• El sistema se incorporará al registro de algoritmos públicos dentro del listado de sistemas automatizados de riesgo alto.• Se exigirá el cumplimiento de las obligaciones de publicidad activa derivadas de la Ley.• Se deberá remitir un informe por parte del DPD de la evaluación de impacto en materia de protección de datos personales realizada cuando resulte aplicable esta normativa.• Se deberá elaborar un informe que contenga al menos los siguientes elementos:<ul style="list-style-type: none">○ Informe sobre los potenciales riesgos detectados○ Indicación de los posibles intereses/bienes jurídicos en juego legítimos que limiten la publicidad activa.○ Justificación de la limitación de las obligaciones de publicidad activa atendiendo a los límites previstos en la normativa autonómica.
--------------------	---

Riesgo Medio	<ul style="list-style-type: none"> • El sistema se incorporará al registro de algoritmos públicos dentro del listado de sistemas automatizados de riesgo medio. • No se exigirá el cumplimiento de las obligaciones de publicidad activa derivadas de la Ley. • Seguimiento de los potenciales riesgos detectados (véase apartado de gobernanza).
---------------------	--

Riesgo Bajo	<ul style="list-style-type: none"> • Ninguna actuación.
--------------------	--

4. Registro de algoritmos

Asignado el riesgo por parte del CRAP, la ORA llevará a cabo las actuaciones necesarias para registrar dichos algoritmos en el registro oportuno y llevará a cabo las tareas de publicidad activa derivadas de esta norma. Dependiendo de si el nivel de riesgo es medio o alto, se llevarán a cabo las obligaciones de publicidad activa.

III. Registros de algoritmos que se incorporarán en el futuro

Además de los algoritmos que se estén utilizando por parte de la GVA, también se debe contemplar un procedimiento específico para aquellos algoritmos que se pretendan incorporar al seno de las AAPP.

Para estos supuestos, las actuaciones básicas que se proponen son las siguientes:

Principio de coordinación entre órganos administrativos a la hora de implementar actuaciones administrativas automatizadas.

Solicitud de información algorítmica.

Evaluación del riesgo y asignación del riesgo.

Registro de algoritmos.

1. Principio de coordinación entre órganos administrativos.

El Decreto 220/2014, de 12 de diciembre, del Consell, por el que se aprueba el Reglamento de Administración Electrónica de la Comunitat Valenciana establece varias obligaciones que han de llevar a cabo la GVA cuando quiera implementar sistemas o aplicaciones informáticas que vayan a ser utilizados para la gestión de actuaciones administrativas automatizadas.

Entre estas obligaciones (ya comentadas)¹⁸⁷, podemos destacar:

- Aprobación previa del sistema a través de diferentes instrumentos normativos antes de su uso (Artículo 78.1).

¹⁸⁷ Véase el apartado específico dedicado a estas materias en el presente informe. “Divulgación del responsable del sistema y aplicación y en su caso, del informe preceptivo realizado en razón del Decreto 220/2014, de 12 de diciembre de Administración Electrónica valenciana”.

- Designación del órgano competente responsable a efectos de impugnaciones (Artículo 78.2).
- Reconocimiento del derecho de los ciudadanos a conocer el órgano responsable del sistema automatizado (Artículo 6.2º).
- Aprobación del catálogo de sistemas de información y aplicaciones aprobadas de la Generalitat (Artículo 95).

La estructura descrita y diseñada por el Decreto de 2014 es ideal para integrar algunas de las actuaciones que se deben desplegar a la hora de desarrollar el Registro de Algoritmos. Ello exige como es lógico la coordinación de diferentes órganos administrativo.

En este sentido, sería necesario que todos los órganos administrativos que pretendan implementar sistemas automatizados conozcan de la existencia de la obligación de la publicidad activa que se desprende de la Ley de Transparencia a efectos de su cumplimiento.

Para lograr este objetivo proponemos distintas actuaciones:

- Elaboración de un instrumento normativo que obligue a todos los órganos de la administración de la GVA a informar a la ORA del despliegue de los sistemas.
- Modificación del actual Decreto de 2014 para establecer dicha obligación de comunicación a la ORA.
- Desarrollar un reglamento específico que contemple esta nueva obligación, así como otros elementos que ya se vienen comentando.
- La Conselleria con competencias horizontales en materia de tecnologías de la información y las comunicaciones de la Generalitat se encargará de informar a la ORA de la incorporación de nuevos sistemas automatizados.

2. Solicitud de información algorítmica.

Una vez la ORA tenga constancia de la implementación de un nuevo sistema automatizado a través de alguno de los canales de comunicación propuestos, ésta deberá facilitar al órgano responsable de dicho sistema la solicitud de información algorítmica previamente explicada.

3. Evaluación y asignación del riesgo

A la hora de evaluar y analizar el riesgo de algoritmos de nuevo ingreso, se seguirá el mismo procedimiento y los mismos niveles descritos previamente. Recuérdese que se proponía riesgo alto, medio y bajo.

4. Registro de algoritmos públicos.

El registro de algoritmos públicos de nueva implementación se ejecutará en los términos anteriormente descritos.

IV. Gobernanza del Registro de Algoritmos Públicos

Para ejercer una gobernanza interna adecuada de los mecanismos y actuaciones descritas hasta ahora, proponemos otra serie de medidas

necesarias que se han de desplegar y que han de estar presentes durante el despliegue y seguimiento del Registro de Algoritmos Públicos.

1. Medidas de gobernanza general aplicable a los sistemas automatizados de riesgo alto

Revisión de la información algorítmica publicada

Anualmente el CRAP deberá emitir un informe en el que proponga posibles cambios relativos al tipo de información y transparencia que se ha de publicar respecto de los sistemas de alto riesgo.

Para ello, el CRAP puede valerse entre otras herramientas de:

- Reformas legislativas que impongan nuevas medidas de transparencia algorítmica.
- Resoluciones judiciales sobre la materia.
- Resoluciones de los Consejos de Transparencia autonómicos y nacional.
- Despliegue de desarrollos tecnológicos que mejoren la transparencia algorítmica.

Seguimiento del correcto funcionamiento del mecanismo de publicidad activa de los algoritmos de alto riesgo

Anualmente la ORA deberá emitir un informe sobre el seguimiento de las obligaciones derivadas de publicidad activa.

Para ello, la ORA puede valerse entre otras herramientas de:

- Encuestas a los responsables de los sistemas algorítmicos sobre posibles nuevos riesgos que se han podido detectar.
- Recopilación de las quejas y sugerencias de la ciudadanía referidas a la publicidad activa algorítmica.

2.) Medidas de gobernanza general aplicable a los sistemas automatizados de riesgo medio

Revisión del nivel de riesgo de los sistemas de riesgo medio.

Se debería establecer un procedimiento que permita a la CRAP revisar el riesgo de los sistemas considerados inicialmente de riesgo medio.

Este procedimiento podría iniciarse a través de diferentes formas:

- De oficio por el propio CRAP respecto de uno o varios sistemas automatizados.
- A solicitud de la ORA.
- A solicitud de los responsables de los distintos sistemas automatizados.

A la hora de valorar la elevación del riesgo medio al alto, la CRAP puede tener en cuenta entre otros elementos:

- Quejas y sugerencias de la ciudadanía referidas a la a los sistemas analizados.
- Indicaciones planteadas por los propios responsables de los

sistemas.

- Conocimiento de informes, documentos u otra información que ponga de manifiesto el riesgo de estos sistemas catalogados inicialmente de riesgo medio.

Conclusiones y recomendaciones para la puesta en marcha del sistema de publicidad activa de algoritmos empleados por la Generalitat Valenciana

Primera.- El creciente empleo por parte de la Administración pública valenciana de sistemas algorítmicos o de inteligencia artificial para la toma de decisiones o que puedan ayudar a ésta o condicionarla de manera sustancial, que se incrementará además notablemente en un futuro próximo, requiere en la actualidad del cumplimiento de una serie de medidas de transparencia y publicidad activa, de acuerdo con el art. 16.1 I) de la Ley 1/2022, de 13 de abril, de la Generalitat, de Transparencia y Buen Gobierno de la Comunitat Valenciana (LTBGCV), que por primera vez en España obliga a una administración pública a incorporar en su portal de transparencia la información relativa a qué sistemas de esta naturaleza está empleando, así como a indicar también los datos y elementos básicos que permitan dar a los ciudadanos cumplida información básica sobre el funcionamiento de estos sistemas. Se trata de una garantía indudable, muy necesaria, y de una mejora regulatoria que pone a la Comunitat Valenciana y a la Generalitat Valenciana en la vanguardia de los regímenes jurídicos de transparencia para el uso de IA por los poderes públicos a nivel no sólo español sino comparado.

Segunda.- Para el correcto desarrollo y despliegue de la referida obligación, y más allá de dar cumplimiento al art. 16.1 I) LTBGCV, es precisa la adopción de una serie de medidas de tipo ejecutivo y organizativo, no concretadas en la ley, por ser propias, en su caso, de su desarrollo reglamentario o, más sencillamente, por tratarse de medidas de tipo organizativo interno convenientes para dar mejor y cumplido desarrollo a la norma, o, sencillamente, medidas de tipo administrativo y ejecutivo que se refieran a cómo y cuándo publicar la información sobre los algoritmos y sistemas de IA empleados, posibles límites, forma de proporcionar la información, así como decisiones respecto a cómo realizar todas estas funciones de acuerdo con los marcos autonómico, nacional y europeo en protección de datos, procedimiento administrativo, transparencia e inteligencia artificial. Para ello, es conveniente, como se ha acometido a lo largo de este informe, analizar tanto las posibilidades de la norma como los mejores ejemplos comparados, de los que tenemos ya algunos en el marco europeo, y analizar la manera en que se han resuelto en esos casos algunos de los problemas que el desarrollo del art. 16.1 I) LTBGCV puede plantear. Se ha de señalar que, en los ejemplos descritos, la mayor parte de las decisiones son de tipo ejecutivo y práctico, exista o no cobertura legal, con la intención de garantizar una mejor transparencia y equilibrarla correctamente con las necesidades de eficiencia y eficacia de la actuación administrativa. A partir de estos postulados, se realizan las siguientes recomendaciones para el caso valenciano.

Tercera.- Para el óptimo desarrollo y aplicación de las obligaciones de transparencia que ha de cumplir la administración autonómica valenciana en materia de transparencia algorítmica de acuerdo con la actual legislación

vigente, se hacen necesarias ciertas medidas de tipo estrictamente organizativo para ayudar a la gestión en el cumplimiento de las mismas, reduciendo las cargas burocráticas y garantizando un mejor cumplimiento de la obligación para la defensa de los derechos y garantías de los ciudadanos. En este sentido, es conveniente que el cumplimiento de la obligación del art. 16.1 I) LTBGCV pase por la existencia de un registro único centralizado de algoritmos de la Generalitat Valenciana (del mismo modo que lo han creado otras administraciones públicas locales a nivel europeo), en el que de manera organizada y sistemática se incluyan todos los algoritmos y sistemas de IA que en el seno de la Generalitat valenciana adopten decisiones administrativas o influyan o puedan determinar de modo sustancial éstas, recogiendo el sentir de la obligación legal, y que sea la base de la información que a su vez aparecerá también accesible desde el portal de transparencia.

Cuarta.- La creación de este Registro de Algoritmos Públicos de la Generalitat obliga a una serie de procedimientos y cambios organizativos, y ha de ser acompañada desde su inicio de las correspondientes directrices, pero también de ciertas novedades institucionales y organizativas. Atendiendo a la legislación vigente y a la mejor manera de dar cumplimiento a las obligaciones legales, así como a los ejemplos comparados más exitosos, se ha considerado conveniente la creación, a estos efectos, y con las funciones propuestas en este informe, de una Oficina de Registro de Algoritmos (ORA) de la Generalitat, que será la responsable de la llevanza del mencionado registro y de la centralización y ejecución de todos los trámites para la recopilación de la información sobre los sistemas efectivamente empleados por la Generalitat, recabando la debida información, y tratándola de acuerdo a las previsiones y recomendaciones de este informe para su publicación en el Registro y así dar cumplimiento a las obligaciones en materia de transparencia. En este Registro habrán de incluirse únicamente aquellos algoritmos y sistemas de inteligencia artificial que, de acuerdo con el art. 16.1 I) LTBGCV hayan de estarlo, esto es, los que tengan una efectiva incidencia, que se corresponde con la clasificación de riesgo medio/alto que nos viene de las propuestas europeas sobre uso de algoritmos e IA en el sector público.

Quinta.- A efectos de la evaluación del riesgo de los algoritmos empleados por la Generalitat y el sector público valencianos se recomienda la creación del Consejo del Registro de Algoritmos Públicos de la Generalitat, órgano especializado de consulta y asesoramiento, con una composición a partir de especialistas en la materia suficientemente transversal de la propia administración autonómica y expertos externos, cuya función ha de ser la de determinar los criterios que, de acuerdo con el marco legal y la evaluación de los riesgos bajo, medio o alto hacen necesaria o no tanto la inclusión de un algoritmo o sistema en el Registro como la colaboración con la ORA y las unidades/equipos/responsables de los diferentes algoritmos catalogados respecto de la profundidad de la información proporcionada, en su caso; su actualización; límites; seguimiento... en el marco de lo propuesto en este informe.

Sexta.- Con carácter general, la puesta en marcha de la ORA requerirá de la recopilación de todos los algoritmos y sistemas de inteligencia artificial que ya

se estén empleando por parte de la Generalitat y que han de ser incluidos en el Registro de Algoritmos Públicos de la Generalitat. Esta labor de recopilación, con ayuda de todos los equipos/unidades/responsables ha de ser liderada por la Dirección General de Transparencia de la Generalitat mientras no exista la ORA, para pasar a ser responsabilidad de ésta una vez creada. Junto a esta labor de mapeado y puesta al día del conjunto de herramientas existentes, se ha de establecer un sistema, como el propuesto en este informe, de remisión automática de toda la información al órgano responsable (Dirección General o bien ORA, una vez creada) cuando se pretenda implantar un nuevo procedimiento o sistema que pueda ser afectado por las exigencias del art. 16.1) LTBGCV. A estos efectos, este informe recomienda, en este punto, seguir el procedimiento ya definido en el Decreto 220/2014, de Reglamento de Administración Electrónica de la Generalitat, con los debidos ajustes, para integrar la obligación de recopilación de algoritmos y su registro.

Séptima.- Absolutamente esencial en todo este proceso es la determinación de qué sistemas y algoritmos se consideran afectados por la obligación de transparencia del art. 16.1 I) LTBGCV al ser decisivos o suficientemente relevantes en la adopción de la decisión administrativa o despliegue de la política o prestación del servicio público por parte de la administración valenciana. En este sentido, se recomienda integrar este análisis con la debida evaluación del riesgo de cada sistema, integrando en este punto las diferentes categorías y diferenciaciones, trasladadas al sector público, de la regulación europea en elaboración sobre Inteligencia Artificial, que recoge el estado de maduración sobre evaluación comparada de riesgos algorítmicos. Así, y siguiendo los sistemas de evaluación de riesgo indicados, se propone que aquellos sistemas de IA o algoritmos empleados por el sector público valenciano sean incluidos en el Registro de Algoritmos Públicos cuando de la evaluación del riesgo, siguiendo los estándares europeos, se concluya que estamos ante un riesgo alto o medio. Sin embargo, se entiende que las obligaciones de publicidad activa en los términos de la LTBGCV son predicables sólo de los que puedan ser catalogados como de riesgo alto, por ser los que afectan a las decisiones que puedan tomar los poderes públicos que afectan a los derechos de los ciudadanos. Para los de riesgo medio, y a efectos de valorar que puedan pasar a ser de riesgo alto y a entrar en la órbita de lo previsto en el art. 16.1 I) LTBGCV, se propone además de su inclusión en el registro el seguimiento de sus riesgos potenciales. Por el contrario, y para los algoritmos o sistemas de IA que se considere que son de riesgo bajo, a partir de la evaluación realizada por la ORA y su Consejo, en colaboración con las unidades/equipos/responsables de cada sistema, no se propone el despliegue de ninguna medida por entenderse que son herramientas puramente instrumentales que en sí mismas ni comportan riesgos ni son las que tiene en mente, en ningún caso, el art. 16.1 I) LTBGCV.

Octava.- En la medida en que se considera que el concepto de sistema algorítmico que maneja la ley ha de quedar íntimamente vinculado a esta evaluación del riesgo y, en concreto, a una evaluación realizada siguiendo los estándares europeos en la materia, y especialmente los definidos previsiblemente por el futuro reglamento europeo que regulará los usos de la inteligencia artificial en Europa (RIA), tal y como se indica en el Informe, todos los sistemas expresamente incluidos en el art. 6.3ª RIA en conexión con su

Anexo III son considerados de alto riesgo y, en consecuencia, en el caso en que la Generalitat o el sector público instrumental valenciano los emplee han de quedar sujetos a las obligaciones de gestión y transparencia indicadas. En concreto, en todo caso lo serán los posibles sistemas de identificación biométrica, aquellos empleados para la gestión y priorización de emergencias, los que determinen la aplicación policial y judicial de determinados contenidos legales y especialmente los que se concreten en la elaboración de perfiles de los ciudadanos (lo que ha de trasladarse a los perfiles de ciudadanos para un uso estrictamente administrativo y de gestión más allá de cuestiones policiales), los que puedan emplearse para la admisión o no en programas o centros de enseñanza (así como para la evaluación y seguimiento del aprendizaje), los relativos a servicios de empleo y evaluación del rendimiento, así como todos los sistemas de evaluación, revisión, concesión o revocación de prestaciones y servicios esenciales de asistencia pública. Por último, también los usos en el ámbito de la justicia y penitenciario lo serían. Aunque no se considere de alto riesgo per se en la actual versión del RIA, es evidente que de esta categorización, debido a la existencia de perfilados implicados en ello, todos los sistemas que los incorporen y, especialmente, los empleados para la persecución del fraude, blanqueo, controles en materia fiscal o de seguridad social, inspección de trabajo o sanciones administrativas de cualquier tipo son también particularmente conflictivos y han de quedar integrados, cuando supongan un perfilado y ayuden a la adopción de decisiones públicas que pueden tener un elevado nivel de impacto en los derechos y situaciones jurídicas de los ciudadanos, en el grupo de los algoritmos públicos de alto impacto, por consideraciones analógicas evidentes.

Novena.- Más allá de la lista que pueda contener al versión final del RIA sobre qué tipo de algoritmos o sistemas de inteligencia artificial empleados por el sector público, a la luz de sus previsiones, han de quedar necesariamente incluidos como de alto riesgo, hay que entender que esa lista en ningún caso resultará exhaustiva, y que a partir de los criterios tanto del propio RIA como de la evolución de los mismos y de las mejores prácticas europeas en la materia, se ha de realizar una evaluación del riesgo de cada algoritmo que, en los términos desarrollados por este informe, ha de tener en cuenta:

- Métricas de impacto y error, así como estudios y análisis de posibles sesgos:
- Una consideración cuidada de los posibles impactos diferenciales, en cada caso, sobre grupos especialmente protegidos o vulnerables.
- Impactos jurídicos significativos y diferenciados derivados del hecho de que el algoritmo sea empleado por los poderes públicos.
- Impacto de la personalización completa y de los perfilados, especialmente sensibles en el sector público.
- Grado de individualización de las decisiones y hasta qué punto el sistema algorítmico o de IA llega o no, predetermina o no y hasta qué punto, esa individualización. Esto conduce a un menor impacto en ciertos casos (mera mejora de sistemas administrativos) y un mayor impacto en otros (supuestos de generalización y masividad, problemas derivados con la afección o reafirmación de sesgos sistémicos y estructurales), con riesgo de errores masivos, que en esos casos han de requerir de un estudio de impacto particularmente detallado

- Determinación concreta y exacta del grado de automatización e intervención o supervisión humana de cada automatismo y evaluación de lo que significa, en cada caso, a efectos del riesgo del sistema, incluyendo la evaluación del grado real y efectivo de supervisión humana en cada caso.
- Por último, y de gran importancia para el caso de los sistemas automatizados empleados por los poderes públicos, como es el caso, esta evaluación del riesgo concreto de cada sistema ha de evaluar y valorar, en cada caso, cómo se integra el empleo y uso del sistema dentro del procedimiento administrativo o procesos administrativos en cuestión.

Décima.- Respecto de la información que ha de ser incluida en el Registro y, en su caso, a efectos de transparencia, el informe ha realizado también un exhaustivo repaso a qué información se considera que se ha de proporcionar respecto de los sistemas algorítmicos y de inteligencia artificial en razón de la obligación legal contenida en el art. 16.1 l) LTBGCV. En este sentido, el informe concluye la necesidad de extremar la transparencia y la información proporcionada en determinados casos; en concreto, cuanto mayor sea la complejidad del sistema, así como cuanto más intervenga el sistema en la toma de decisiones que afecte a la ciudadanía. Asumiendo, también, la existencia de límites, tanto legales como en ocasiones derivados de la propia preservación de otros viene jurídicos y derechos como los datos de carácter personal y su protección.

Undécima.- Respecto de la manera de presentar la información, su grado de profundidad y comprensibilidad, el informe aboga, en la línea de los mejores sistemas europeos, por proporcionar la información en capas, estructurando la información que se hace pública de modo que un nivel básico se presente debidamente explicado y enmarcado para su comprensión por cualquier persona, con independencia del nivel de conocimiento que tenga en la materia, pero con capas de mayor información y detalle respecto de los datos que maneja el sistema y su lógica, que pueden hacerse disponibles o no al público en general, según la evaluación que se haga, y que en todo caso han de estar disponibles para la revisión experta y auditoría, tanto interna como externa.

Duodécima.- Por lo que se refiere a la información que en todo caso ha de proporcionarse de cada algoritmo, se parte de la base de lo ya establecido en el Real Decreto 203/2021, por el que se aprueba el Reglamento de actuación y funcionamiento del sector público por medios electrónicos, por lo que los contenidos mínimos han de comprender: descripción del diseño y funcionamiento del sistema; mecanismos de rendición de cuentas y transparencia; así como los datos utilizados para su configuración y aprendizaje. El informe detalla, además, la concreción de estas obligaciones en distintos ámbitos y según su normativa sectorial, con la conclusión general de que este es el mínimo explicativo que habrá de contener el Registro de Algoritmos Públicos y que, caso por caso, y dependiendo de la evaluación de riesgo que se haga y de lo determinado por la legislación sectorial en la materia en su caso y la comisión de evaluación, habrá que hacer con más profundidad y exhaustividad, respetando siempre la protección de datos y, en su caso, otros

derechos. A estos efectos, el diseño propuesto de información por capas, así como el hecho de que respecto de los algoritmos de riesgo medio la información vaya al registro pero no se publique necesariamente, permite un mejor balance del sistema de protección y de garantías y armonizar los posibles intereses en conflicto. En este sentido, el informe detalla también la información que se considera que, por defecto, ha de incluirse sobre el diseño y funcionamiento de cada sistema, con numerosos ejemplos comparados y recomendaciones basadas en extremar la transparencia siempre que sea posible, siguiendo estos ejemplos, pero valorando posibles conflictos y necesidades de eficacia y eficiencia administrativa o colisión con otros derechos o bienes jurídicos. En todo caso, las decisiones que limiten la transparencia habrán de ser motivadas, razón por la cual la estructura organizativa propuesta está diseñada para que estos procesos puedan vehicularse. Así, podría valorarse en este sentido, de modo coherente con este diseño, en su caso, que el Registro contemplara por defecto la obligación de publicación completa del código fuente en las capas más profundas y especializadas, con obligación de justificación y aportación de las razones, jurídicas y técnicas que en su caso puedan explicar que en ciertos casos no se haga así, incluyendo los debidos informes en este sentido del referido Consejo del Registro de Algoritmos Públicos de la Generalitat Valenciana. De manera análoga a este tipo de decisiones, cualquier restricción en el acceso o la transparencia ha de pasar también por el Consejo y ser debidamente explicada, razonada y, en definitiva, motivada.

Decimotercera.- Se propone el establecimiento de un punto de contacto y de la identificación en todo caso, siempre, tanto en el registro como en la información publicada en el portal de transparencia, de un responsable administrativo como mecanismo de garantía y aseguramiento del cumplimiento tanto de la legislación básica estatal en materia de procedimiento administrativo como de las exigencias derivadas de las normas europeas y nacionales en materia de protección de datos, pero también de la propia norma valenciana en materia de transparencia algorítmica.

Decimocuarta.- La contratación administrativa referida a soluciones de IA o a sistemas algorítmicos ha de incorporar, desde los pliegos de contratación, todas estas exigencias para que los proveedores de estas soluciones, además de cumplir con las obligaciones legales, estén también obligados por el contrato a respetar en todo caso el modelo y diseño de transparencia algorítmica de la Generalitat valenciana, así como las mejoras en materia de gobernanza algorítmica y de riesgos que se derivan del mismo. En este sentido, se propone también el avance en la redacción de clausulados tipo para toda la administración que sirvan como base para estos pliegos de contratación y faciliten la labor de los distintos órganos.

Decimoquinta.- Respecto de los límites concretos a la información aportada, y más allá de lo expuesto en la conclusión duodécima, la necesidad de motivación y la labor del Consejo, el informe propone una serie de medidas que se fundamentan en las prácticas de algunos países europeos ya consolidadas (especialmente, las del ICO inglés), de manera que con carácter general se considera que puede justificarse la limitación de la información publicada en relación con los sistemas algorítmicos siempre que existan riesgos

de ciberseguridad, propiedad intelectual o de que la transparencia absoluta facilite que se “engañe” al sistema (*game the system*). Estas tres grandes categorías serían las que, con carácter general, permitirían excepcionar el régimen normal y ordinario de transparencia propuesto en el informe, pero con necesidad de concreción y explicación caso por caso, y de traslación de las reflexiones sobre sus riesgos y por qué tendría sentido la limitación de manera particularizada respecto de cada sistema. En este punto, se insiste en la necesidad de justificar, de manera exhaustiva, todas aquellas razones que sirvan de base a la limitación de la transparencia de los sistemas algorítmicos, motivación sin la cual no podría ser aceptada la restricción, así como la idoneidad del sistema institucional y organizativo propuesto para llevar a cabo esta valoración y motivación con todas las garantías, en su caso, en cada su situación (con base en la labor del Consejo propuesto en el diseño institucional sugerido en este informe, actuando en todo caso en conexión con los responsables/órganos/equipos de cada sistema algorítmico).