



Breve brújula para el despliegue de Sistemas de Inteligencia Artificial por Administraciones Públicas con los Derechos Fundamentales como norte^(*)

A Brief Compass for the Deployment of Artificial Intelligence Systems by Public Administrations, with Fundamental Rights as the Guiding North

Matilde Carlón Ruiz^()**

Universidad Complutense de Madrid (Madrid, España)

Resumen: Las Administraciones Públicas no pueden quedar ajenas a las múltiples oportunidades que les ofrecen las herramientas de inteligencia artificial en su variedad y versatilidad. Estas oportunidades no están, sin embargo, exentas de riesgos, particularmente para los derechos fundamentales. El Reglamento de Inteligencia Artificial aprobado por la Unión Europea en junio de 2024 es consciente de ello, exigiendo a los organismos públicos que desplieguen sistemas de IA de alto riesgo que realicen una evaluación de impacto que se ofrece como un instrumento valioso que las Administraciones Públicas ponderen la decisión misma de implantar o no el sistema.

Palabras clave: Inteligencia Artificial - Sistemas de IA de Alto Riesgo - Evaluación de Impacto - Derechos Fundamentales – Derecho Administrativo - España

Abstract: Public administrations cannot ignore the numerous opportunities offered by artificial intelligence tools, given their variety and versatility. These opportunities are not, however, without risk, particularly with regard to the protection of fundamental rights. The Artificial Intelligence Regulation adopted by the European Union in June 2024 acknowledges this, requiring public authorities deploying high risk AI systems to carry out an impact assessment. This assessment offers a valuable tool for public administrations to weigh the very decision of whether or not to implement the system.

Key words: Artificial Intelligence – High Risk AI Systems – Impact Assessment – Fundamental Rights – Administrative Law – Spain

(*) Nota del Equipo Editorial: Este artículo fue recibido el 2 de noviembre de 2025 y su publicación fue aprobada el 6 de febrero de 2026.

(**) Abogada por la Universidad de Oviedo (Oviedo, España) (con premio extraordinario). Doctora por la misma Universidad. Catedrática de Derecho Administrativo en la Universidad Complutense de Madrid. Sus investigaciones se han dedicado a temas muy diversos, con especial atención a los servicios públicos, sectores regulados, contratación pública y control de reglamentos. Es técnico de la Comisión Nacional de los Mercados y la Competencia, ha sido Experto Nacional Destacado ante la Unión Europea y letrada del Tribunal Constitucional por adscripción temporal. ORCID: <https://orcid.org/0000-0002-9054-7047>. Correo electrónico: mcarlonr@der.ucm.es.



1. Introducción

La inteligencia artificial es un tema de moda. Tan es así que resulta por momentos un lugar común en toda conversación y un tema recurrente en seminarios y foros. Seguramente tenga que ver con ello la irrupción reciente de los modelos de lenguaje, con el consabido *ChatGpt* como paradigma, cuyas capacidades nos admiran y desconciertan a un mismo tiempo, haciéndonos creer que se ha roto definitivamente la barrera del famoso test de Turing.

Este fenómeno de omnipresencia y fascinación por la Inteligencia Artificial (en adelante, IA), que, de hecho, está más presente en nuestras vidas de lo que estamos dispuestos a creer, ha contagiado, como no podía ser de otro modo, la órbita de actuación de los poderes públicos, particularmente de las Administraciones públicas, que parecen impelidas a incorporar sistemas de IA en sus actividades a como haya lugar. En este escenario, parece conveniente ofrecer de forma sosegada alguna pauta que oriente a los actores públicos, en cuanto potenciales responsables del despliegue de sistemas de IA, expresiones sobre las que posteriormente repararemos, y a los propios ciudadanos, que pasarán a relacionarse con una Administración en mayor o menor medida transformada por estos sistemas. Y ello por cuanto, como se ha puesto de manifiesto reiteradamente, y resulta una obviedad, la incorporación de sistemas de IA al quehacer público introduce oportunidades y riesgos que es necesario ponderar.

Las páginas que siguen pretenden responder a este objetivo, de forma necesariamente sintética. Para ello partiremos del esquema que ofrece el ya famoso Reglamento Europeo de Inteligencia Artificial aprobado hace poco más de un año⁽¹⁾, cuyo espíritu parece haber contagiado la reciente normativa peruana de IA⁽²⁾.

2. Los tres rasgos de la IA: (i) heterogeneidad, (ii) inestabilidad e (iii) instrumentalidad y sus implicaciones en su uso por Administraciones públicas

Las Administraciones Públicas no pueden ni deben ser ajenas al fenómeno de la inteligencia artificial. El mejor

cumplimiento de los intereses generales que constituye su razón de ser puede encontrar un instrumento especialmente hábil en las múltiples facilidades y herramientas que se engloban bajo el concepto de inteligencia artificial. Su posición exorbitante en el ordenamiento reclama, sin embargo, que el uso o, mejor dicho, particularmente algunos de los usos, de sistemas de IA por Administraciones Públicas estén revestidos de especiales cautelas.

La heterogeneidad es, de hecho, uno de los rasgos que caracterizan a la inteligencia artificial. Conviene advertir desde este momento que bajo esta expresión se engloban múltiples manifestaciones de lo que no deja de ser un *software* avanzado, caracterizado por una mayor o menor autonomía de acción sobre la base de la aplicación a amplios conjuntos de datos de algoritmos, es decir, de instrucciones matemáticas precisas determinadas por un objetivo con mayor o menor grado de sofisticación, en función de capacidades de computación crecientes⁽³⁾. Son estos tres elementos (i) datos, (ii) algoritmos, (iii) computación, los que están en la base de cualquier sistema de inteligencia artificial desde su origen hace más de 80 años. De ahí precisamente que la ampliación cuantitativa y cualitativa de la disponibilidad de estos tres elementos en los últimos tiempos haya determinado los llamativos avances que han conducido a los actuales momentos de fascinación. Y de los avances que vendrán, porque a la heterogeneidad de las aplicaciones y herramientas reconocidas como IA se suma el rasgo de su inestabilidad y continua, incluso frenética, evolución, lo que suma complejidad al análisis del fenómeno y al proceso que tiende a juridificarlo.

Junto a la heterogeneidad y la inestabilidad, un tercer rasgo viene dado por el carácter

(1) Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial). Sobre el Reglamento se han publicado algunos comentarios generales, véase: Barrio Andrés, M. (Dir.). (2024a). *El Reglamento Europeo de Inteligencia Artificial*. Tirant lo Blanch, Barrio Andrés, M. (Dir.). (2024b). *Comentarios al Reglamento Europeo de Inteligencia Artificial*. La Ley, Cotino Hueso, L., & Simón Castellano, P. (Dirs.). (2025). *Tratado sobre el Reglamento Europeo de Inteligencia Artificial*. Aranzadi.

(2) Nos estamos refiriendo al Decreto Supremo 115-2025 que aprueba el Reglamento de la Ley 31814, Ley que promueve el uso de la inteligencia artificial en favor del desarrollo económico y social del país.

(3) Acerca del fenómeno IA Véase "Inteligencia Artificial" de Boden (2022).



instrumental de los sistemas de IA. Algo que, pareciendo obvio, conviene enfatizar para calibrar el alcance de esa juridificación cuando se trata de ordenar el uso de estos sistemas de IA por parte de Administraciones públicas.

La inteligencia artificial opera como una mera herramienta, que la Administración decide introducir o no y de cuyos resultados es en todo caso responsable. Esto es así tanto en los casos en los que las Administraciones utilicen sistemas de IA como mero apoyo en la adquisición de información o en la redacción de documentos, cuando los apliquen para el análisis de datos a efectos de orientación o consulta o, incluso, cuando se planteen la incorporación de los sistemas de IA para la propia adopción de decisiones⁽⁴⁾. La heterogeneidad de los sistemas se traduce en su versatilidad, siendo así que a las Administraciones públicas les corresponde decidir cuándo y para qué introducen estos sistemas, decisión en la que deberán ponderar beneficios y riesgos.

No cabe duda de que, en este amplio abanico de manifestaciones, los riesgos inherentes al uso de sistemas de IA por Administraciones Públicas no son unívocos. Como no cabe dudar de que algunos de ellos pueden ser especialmente intensos, particularmente en relación con algunos ámbitos de actuación pública. Así se refleja en el ya citado Reglamento europeo de Inteligencia Artificial, que se muestra plenamente consciente del papel que pueden jugar las Administraciones Públicas en el desarrollo y utilización cualificada de sistemas de inteligencia artificial, imponiendo en ocasiones reglas específicas.

Antes de identificar algunas de estas especialidades, es necesario ofrecer algunas aclaraciones sobre el planteamiento regulatorio de este Reglamento.

3. Claves del planteamiento regulatorio del Reglamento Europeo de Inteligencia Artificial: la IA como producto

El Reglamento europeo de Inteligencia Artificial aborda la regulación de los sistemas de IA considerándolos como productos. Entender esto es clave. De ahí que su enfoque esté basado en establecer una armonización a nivel europeo de los

requisitos y condiciones de los sistemas de IA a efectos de garantizar su libre circulación sobre la base de la necesaria protección no solo de la salud y la seguridad, que son los valores típicamente concernidos cuando se trata de armonizar la regulación de mercancías, sino también de los derechos fundamentales, por la particular condición de este producto singular⁽⁵⁾.

Esta particularidad del “producto IA” no deja de generar tensiones en el modelo de regulación armonizador por cuanto se escapa del esquema típico y es difícil de materializar⁽⁶⁾. Instrumento capital al efecto es la evaluación de impacto en los derechos fundamentales que deben llevar a cabo los “organismos de Derecho público” con carácter previo al despliegue del grueso de los sistemas de IA que el propio Reglamento de Inteligencia Artificial (en adelante, RIA) califica de alto riesgo. Una evaluación de impacto que ofrece una clara oportunidad a las Administraciones públicas para racionalizar la decisión de si implantar o no un concreto sistema de IA y, de hacerlo, para pautar algunos elementos fundamentales de su despliegue, como es en particular el relativo a la tantas veces evocada necesaria supervisión humana de los sistemas.

Comprender el alcance de este instrumento requiere desentrañar lo que supone en el marco del RIA la condición de responsable del despliegue de un sistema en correlación con la consideración de determinados sistemas como de alto riesgo.

3.1 La peculiar figura del responsable del despliegue

En los términos del art. 3.4 del Reglamento, se reconoce como responsable de despliegue a la persona física o jurídica o “autoridad pública, órgano u organismo (sic)” que utilice

(4) Sobre los múltiples usos de los sistemas de IA por las Administraciones públicas, Véase: “La aplicación de la inteligencia artificial en la actividad formal e informal de la Administración. Ponencia presentada en el XVII Congreso de la Asociación Española de Profesores de Derecho Administrativo” de Bermejo Vera (2024) y “El uso de sistemas basados en inteligencia artificial por las Administraciones públicas: Estado actual de la cuestión y algunas propuestas *ad futurum* para un uso responsable” de Berning Prieto (2023). Desde una perspectiva con intención de exhaustividad Véase el libro selectivo “Inteligencia artificial y sector público: Retos, límites y medios” dirigido por Gamero Casado (2023) y “AI Watch: European landscape on the use of artificial intelligence by the public sector” de Tangi, L., Van Noordt, C., Comberty, M., Gattwinkel, D., & Pignatelli, F. (2022).

(5) Planteando, como dilema, la presencia de ambos objetivos en el planteamiento del RIA, véase: “El Reglamento europeo de inteligencia artificial: ¿garantía del mercado o defensa de los derechos?” de Presno Linera (2025).

(6) Véase “La regulación de la inteligencia artificial en Europa a través de la técnica armonizadora del nuevo enfoque, en RGDA” de Álvarez García y Tahirí Moreno (2023).



un sistema de IA de alto riesgo bajo su propia autoridad, siendo así que el concepto de “sistema de IA” se define en el apartado 1 del mismo art. 3 en términos que reflejan las características de la IA más arriba descritas, poniendo énfasis en la capacidad de inferencia de los sistemas y precisamente en la de “adaptación tras el despliegue”⁽⁷⁾.

Cualificando la posición del mero usuario, con la figura del responsable del despliegue, introducida por el Parlamento Europeo en la redacción final del Reglamento, se quiere perfilar la posición de quien decide introducir en sus procedimientos “bajo su autoridad”, sistemas de inteligencia artificial de alto riesgo, implicándose en el proceso de aplicación, y asumiendo correlativamente la responsabilidad consiguiente. Es esta, pues, una figura original y por sí misma muy reveladora de la peculiaridad del “producto IA”.

El responsable del despliegue asume, pues, una serie de obligaciones específicas en el marco del RIA frente a las que asume el proveedor del sistema, que es quien lo desarrolla por sí mismo o para el que es desarrollado, de modo que se ponga en servicio con su propio nombre o marca, en los términos del apartado 3 del mismo art. 3 del RIA. Y ello en el bien entendido de que responsable del despliegue y proveedor pueden llegar a coincidir cuando concurra alguno de los supuestos del art. 25 del RIA, bajo coordenadas en las que no será infrecuente que se encuentren Administraciones públicas: cuando pongan su nombre en un sistema previamente introducido en el mercado o puesto en servicio, cuando lo modifiquen sustancialmente o cuando modifiquen la finalidad prevista con el efecto de que un sistema que no fuera originalmente reconocible como de alto riesgo pase a tener tal consideración en el marco del art. 6 del RIA. Es este el de la “finalidad prevista” un requisito capital en el marco del RIA que determina la clasificación de los sistemas en función de un análisis de riesgos.

3.2. La clasificación de los sistemas de IA desde un punto de vista funcional

El RIA asume, en efecto, una perspectiva funcional al abordar el análisis de riesgos que está en su base, tanto para su diseño como para su aplicación. Los sistemas de IA se clasifican y disciplinan en función de su “finalidad prevista”, lo cual redundará en el carácter instrumental que antes destacábamos. Bajo este esquema determinados sistemas se reconocen de alto riesgo, y sobre ellos se proyectan una serie de exigencias impuestas de forma diferenciada a proveedores y responsables del despliegue. El solape de papeles que el art. 25 concibe no puede, sin embargo, conducir a confundirlos, ya que a los

primeros se les imponen obligaciones en el plano del diseño de los sistemas, mientras que las obligaciones que se imponen a los segundos se corresponden con la fase de utilización, aun cuando el mentado principio del ciclo de vida obliga a reconocer una comunicación permanente entre ambas fases.

Conforme al análisis de riesgos efectuado por el RIA, siempre a la luz de los valores de salud, seguridad y protección de los derechos fundamentales ya señalados, determinados sistemas quedan prohibidos (art. 5), otros son considerados de alto riesgo (art. 6), con la consecuencia de someter a sus proveedores y responsables de despliegue a una serie de obligaciones, según anticipamos y desarrollaremos, y los restantes se consideran en principio inocuos y, por ende, de libre comercialización y puesta en servicio, sin perjuicio de la imposición de determinadas obligaciones de transparencia en relación con el uso de ciertos sistemas (art. 50).

Los sistemas de alto riesgo pueden estar integrados en un producto o ser independientes. A estos últimos se refiere el art. 6.2 y, correlativamente, el Anexo III. No es de descartar que las Administraciones públicas se vean involucradas en el manejo de sistemas de IA integrados, de hecho, su uso es hoy por hoy extensivo en la vida común para las utilidades más variadas, pero no cabe duda de que es en relación con el despliegue de sistemas independientes en el que se hace más significativo el papel de las Administraciones públicas como responsables del despliegue.

De hecho, la obligación de realizar una evaluación de impacto del sistema sobre los derechos fundamentales previa al despliegue que el art. 27 del RIA impone a las Administraciones públicas solo se refiere al grueso de los sistemas independientes de alto riesgo. Este instrumento es capital para

(7) Un sistema de IA se define como “sistema basado en una máquina diseñado para funcionar con distintos niveles de autonomía, que puede mostrar capacidad de adaptación tras el despliegue y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar información de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales”. Véase Comisión Europea, Directrices relativas a la definición de sistema de inteligencia artificial establecida en el Reglamento de Inteligencia Artificial, 29 de julio de 2025, C(2025) 5053 final.



ponderar la oportunidad de introducir un concreto sistema de IA en la actividad administrativa en el marco de la lógica riesgo/beneficio. Y se engarza, en función del principio del ciclo de vida, con las exigencias previas que se imponen a los proveedores de los sistemas antes de introducirlos en el mercado o de ponerlos en servicio.

3.3. Las obligaciones de proveedores y responsables del despliegue de sistemas de alto riesgo

Con vistas a asegurar la fiabilidad de los sistemas de alto riesgo bajo la lógica de disciplinar sus posibles afecciones sobre la salud, la seguridad o, por lo que particularmente nos interesa, los derechos fundamentales, el RIA impone una serie de requisitos que pivotan, no casualmente, sobre la obligación nuclear de exigir que dispongan de un sistema de gestión de riesgos (art. 9). En este sistema de gestión de riesgos confluyen, en efecto, otros requisitos entre los que interesa destacar, en particular, las exigencias relativas a la disposición y gobernanza de los datos (art. 10) -la inteligencia artificial se explica por la tríada: (i) datos, (ii) algoritmos y (iii) capacidad de computación; no lo olvidemos-, así como la incorporación de medidas de vigilancia humana (art. 14).

El cumplimiento de estos requisitos se impone, evidentemente, a los proveedores, de modo que cuando las Administraciones Públicas sean proveedoras de los sistemas de alto riesgo, deberán asegurar su cumplimiento y el de las obligaciones que se desgranán en el art. 16, y, entre ellas, con la de someterse a una evaluación de conformidad previa a su puesta en servicio bajo los cánones establecidos en el art. 43, que apuesta decididamente por un modelo de autoevaluación. Esta evaluación de conformidad no ha de ser confundida en ningún caso con la evaluación de impacto en los derechos fundamentales que se impone al responsable del despliegue y se desenvuelve, pues, en la fase de utilización de los sistemas de alto riesgo. Algo que conviene retener al recordar que proveedor y responsable del despliegue pueden coincidir si se dan las circunstancias previstas en el art. 25.

Si no se cumplen estas circunstancias, las Administraciones que se limiten a ser responsables del despliegue solo quedan vinculadas por las obligaciones que se contienen en el art. 26 -cuyo alcance tiene también que ver con la puesta en práctica del cumplimiento de los requisitos que se imponen a los sistemas de alto riesgo- y por la exigencia de desarrollar con carácter previo al despliegue la evaluación que centra nuestra atención. Exigencia que, en rigor, el art. 27 impone a los “organismos de Derecho público” (sic)., según expresión que no debe interpretarse en sentido estricto, proyectada sobre la taxonomía de personificaciones jurídicas

contempladas por la Ley 40/2015 para excluir las formas de personificación jurídico-privadas encuadrables en el sector público. Bien al contrario, su vocación es dar también cobertura a poderes públicos encuadrables en el poder legislativo o judicial, tal y como cabe confirmar cuando se repasan algunos de los supuestos que el RIA identifica como sistemas de alto riesgo.

4. La evaluación de impacto en los derechos fundamentales como obligación peculiar de las Administraciones que desplieguen sistemas de alto riesgo: el análisis de riesgos como perspectiva

El art. 27 del RIA, introducido por el Parlamento europeo en el texto final del Reglamento, impone, en efecto, a los responsables del despliegue del grueso de los sistemas de IA de alto riesgo “independientes” que sean “organismos de Derecho Público” la obligación de que antes del despliegue lleven a cabo una evaluación del impacto que su utilización pueda tener en los derechos fundamentales.

El contenido de esta evaluación queda desmenuzado en el apartado 1 del artículo de referencia, siendo así que su apartado 2 ofrece precisiones sobre los supuestos en los que la evaluación puede quedar suplida o debe ser reeditada, siempre conforme a su lógica funcional, mientras que el apartado 5 busca garantizar su debido deslinde, no siempre fácil, con la evaluación de impacto de la protección de datos previsto en el art. 35 del Reglamento General de Protección de Datos (en adelante RGPD)⁽⁸⁾. Los apartados 2 y 5 ofrecen pautas procedimentales en punto a la necesaria notificación de sus resultados a la autoridad de vigilancia del mercado conforme a un modelo que la Oficina de IA está llamada a definir a nivel europeo con vistas a su simplificación.

(8) Aprobado como Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE.



Resulta bien significativo que la obligación de llevar a cabo una evaluación de impacto de derechos fundamentales se imponga específicamente a los responsables de despliegue de sistemas de alto riesgo que sean “organismos de Derecho público”, junto con “las entidades privadas que prestan servicios públicos”. Ello es prueba de que conforme al análisis de riesgo llevado a cabo por el RIA para establecer los tres niveles de ordenación más arriba descritos, se ha concluido que la utilización del grueso de sistemas de alto riesgo por parte de Administraciones públicas entraña un riesgo potencial para los derechos fundamentales que requiere una ponderación previa a su despliegue en la que se trata de replicar, en el ámbito de cada uso concreto, el mismo tipo de análisis. Es más, una parte muy significativa de los sistemas calificados como de alto riesgo vienen identificados por finalidades genuinamente públicas, siendo la práctica totalidad de los restantes susceptibles de un uso típico o potencial por Administraciones públicas.

4.1. El análisis de riesgos como trasfondo de la clasificación de los sistemas independientes de alto riesgo

La elaboración del RIA se apoya, en efecto, en un análisis de riesgos basado en un concepto asentado de riesgo, que combina la gravedad del posible perjuicio con la probabilidad de que se produzca.

La metodología y criterios de este análisis aparecen reflejados en el art. 7, en el que se establecen los parámetros conforme a los cuales la Comisión puede llegar a modificar en el futuro los concretos supuestos que contempla el Anexo III en atención al rápido ritmo del desarrollo tecnológico y a los posibles cambios en los usos de los sistemas de IA, siempre dentro de los ocho ámbitos definidos en el propio Anexo. La modificación de estos ámbitos requerirá una revisión del Reglamento mismo en los términos del art. 112.

La finalidad prevista del sistema, la probabilidad de su uso, la naturaleza y cantidad de los datos utilizados, la preexistencia de perjuicios y, en todo caso, su intensidad y capilaridad forman parte de los criterios a tener en cuenta, a los que se suman otros especialmente significativos a nuestros efectos. De una parte, el grado de autonomía con el que actúe el sistema y la posibilidad de que un ser humano anule una decisión o recomendaciones que puedan dar lugar a un perjuicio y la medida en que las personas potencialmente perjudicadas por el sistema dependan del resultado generado por el mismo, “en particular porque, por motivos prácticos o jurídicos, no sea sensato (sic) poder renunciar a dicho resultado”. Y, de otra, lo que resulta especialmente relevante en la órbita de las

actuaciones públicas, la medida en que exista un desequilibrio de poder o las personas potencialmente afectadas se encuentren en una situación de vulnerabilidad frente al responsable del despliegue, por razones, entre otras, de autoridad. En su redacción actual, el Anexo III pretende expresar, en definitiva, el delicado equilibrio en el que se desenvuelve el planteamiento del RIA entre protección y oportunidad; entre seguridad y mercado o, más propiamente a nuestros efectos, entre seguridad y eficacia en el cumplimiento de funciones públicas⁽⁹⁾.

Esta misma tensión se manifiesta en el peculiar supuesto que contempla el art. 6.3 del RIA, que habilita al propio proveedor para que autoexcluya su sistema de la consideración de alto riesgo, siempre y cuando el mismo no plantee verosíblemente un riesgo importante para la salud, la seguridad o los derechos fundamentales de las personas físicas: en particular, lo que es muy significativo, “al no influir sustancialmente en el resultado de la toma de decisiones”, lo cual sucederá cuando se den una o varias de las condiciones que el apartado 3 desmenuza, por referencia a los casos en los que el sistema tenga por objeto llevar a cabo una tarea limitada, mejorar el resultado de una actividad humana ya realizada, detectar patrones en la toma de decisiones sin que su finalidad sea “sustituir la evaluación humana previamente realizada sin una revisión humana adecuada, ni influir en ella” o que tenga por objeto llevar a cabo una labor meramente preparatoria. Todo ello con la excepción absoluta de los sistemas que lleven a cabo la elaboración de perfiles de personas físicas en los términos de la normativa de protección de datos, que siempre se considerarán de alto riesgo.

La mera afección a la toma de decisiones, sin necesariamente implicar la toma misma de la decisión por parte del sistema de IA, se considera, pues, la frontera material que determina el ámbito en el que puede o no considerarse de alto riesgo un sistema⁽¹⁰⁾.

(9) Para mayor información en términos no exentos de posible crítica, véase “Truly risk-based regulation of artificial intelligence: How to implement the EU’s AI Act.” De Ebers (2024).

(10) Tal y como aclara el propio RIA, cuando se cumplen las características perfiladas en el art. 6.3, “el sistema de IA solo añade un nivel adicional a la actividad humana, entrañando por consiguiente un menor riesgo”, como sería el caso de los sistemas dirigidos a mejorar el lenguaje de los documentos ya redactados, siendo así que solo es utilizado “tras una evaluación previamente efectuada por un



4.2. Los sistemas de alto riesgo desplegados por Administraciones públicas que quedan sometidos a la evaluación de impacto

Tal y como ha quedado apuntado, la práctica totalidad de los supuestos de sistemas de alto riesgo contemplados en el Anexo III involucran a Administraciones públicas, de forma exclusiva o potencial. Y todos ellos, a excepción del relativo a las infraestructuras críticas⁽¹¹⁾, están sometidos a la obligación del art. 27 del RIA precisamente cuando son desplegados por organismos de Derecho público. Siendo este el contexto que ahora interesa, es oportuno repasar los distintos supuestos a la luz de los derechos fundamentales potencialmente afectados en cada caso, algunos de los cuales han sido puestos de manifiesto por el propio RIA en sus considerandos. Y ello en el bien entendido de que el RIA, como norma europea que es, reclama volver la vista a la Carta de Derechos Fundamentales de la Unión Europea como fuente primaria de reconocimiento de los derechos cuya posible vulneración debe ser evaluada⁽¹²⁾.

Conviene advertir, en todo caso, que las previsiones que nos ocupan no entrarán en vigor hasta el 2 de agosto de 2026, según dispone el art. 113 del RIA.

a) El RIA reconoce de alto riesgo tres supuestos de sistemas de biometría que potencialmente pueden ser utilizados por Administraciones Públicas por considerar que implican usos críticos de datos biométricos, partiendo de la consideración de estos como una categoría especial de datos personales sensibles, sin llegar a una intensidad determinante de la prohibición que contempla el art. 5. De que estos son los sistemas de IA de más alto riesgo en la consideración del RIA da cuenta que son los únicos para los que el art. 43 del RIA contempla un régimen de heterocontrol del cumplimiento de las obligaciones impuestas por él frente al régimen general de autoevaluación.

El primero de los supuestos clasificados se refiere a los sistemas de identificación biométrica remota. Siempre que no se utilicen en tiempo real en espacios abiertos, hay que

entender, por ser este un supuesto prohibido por el art. 5.1.h), y dejando a salvo los sistemas utilizados con fines de verificación biométrica cuyo objetivo sea confirmar que una persona física es quien dice ser o darle acceso a un espacio, servicio o dispositivo. Entre estos sistemas están los utilizados para la búsqueda selectiva de un sospechoso en el marco de una investigación a los que se refiere el art. 26.10 para exigir que el responsable del despliegue obtenga una autorización judicial o administrativa previa o en plazo máximo de cuarenta y ocho horas, descartando en todo caso su uso indiscriminado desvinculado de una concreta investigación penal o una amenaza real y actual o previsible de infracción penal, o de una búsqueda de una persona desaparecida concreta.

En la consideración del RIA, las imprecisiones técnicas de estos sistemas pueden conducir a resultados sesgados con efecto discriminatorio en ámbitos tan sensibles como la edad, la etnia, el sexo o la discapacidad⁽¹³⁾, todo lo cual conduce a que estén sometidos a obligaciones reforzadas de trazabilidad y vigilancia humana. El derecho a la propia imagen se ve frontalmente afectado por estos sistemas, en todo caso.

Se consideran igualmente de alto riesgo los sistemas destinados a la “categorización” biométrica en función de atributos o características sensibles protegidos en virtud del art. 9.1 RGPD que se basen en la inferencia a partir de los mismos. Es decir, a partir de datos que:

humano y no pretende sustituirla o influir en ella sin una revisión adecuada por parte de un ser humano”. No ya la sustituibilidad del ser humano, sino el grado de influencia en sus decisiones, es el que determina el ámbito de consideración como de alto riesgo de los sistemas.

- (11) El RIA clasifica como de alto riesgo los sistemas destinados a ser utilizados como componentes de seguridad en la gestión y funcionamiento de las infraestructuras digitales críticas —de las que se enumeran en el anexo, punto 8, de la ya Directiva (UE) 2022/2557, del Parlamento Europeo y del Consejo de 14 de diciembre de 2022 relativa a la resiliencia de las entidades críticas y por la que se deroga la Directiva 2008/114/CE del Consejo—, del tráfico rodado o del suministro de agua, gas, calefacción y electricidad, en tanto en cuanto un fallo en los mismos puede poner en peligro la vida y la salud de las personas a gran escala y alterar de forma apreciable el desarrollo habitual de las actividades económicas y sociales. Estos supuestos quedan, sin embargo, extramuros de la exigencia de evaluación de impacto en los derechos fundamentales, lo que no deja de ser coherente con el hecho de que la posible afección de derechos fundamentales por estos sistemas no dejaría de ser indirecta
- (12) Para más detalles del impacto de la IA en los derechos fundamentales véase “Derechos fundamentales e Inteligencia Artificial” de Presno Linera (2022).
- (13) Así lo expresa el cdo. 54, que se cuida de precisar, en todo caso, que los sistemas biométricos destinados a ser utilizados exclusivamente a efectos de posibilitar la ciberseguridad y las medidas de protección de datos personales no deben considerarse sistemas de alto riesgo.



Revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, y el tratamiento de datos genéticos, datos biométricos dirigidos a identificar de manera unívoca a una persona física, datos relativos a la salud o datos relativos a la vida sexual o la orientación sexual de una persona física (2016).

El simple repaso del perfil de los datos protegidos da cuenta de los derechos fundamentales potencialmente afectados: igualdad, libertad de pensamiento y de expresión, libertad religiosa, libertad sindical, intimidad y propia imagen. Finalmente, también se consideran de alto riesgo los sistemas destinados al reconocimiento de emociones que no entren en el ámbito de la prohibición contemplada en el art. 5.1.f), que se refiere a aquellos que puedan “inferir” las emociones de una persona física en los lugares de trabajo y en los centros educativos, salvo cuando se introduzcan por razones médicas o de seguridad. Sistemas que pueden afectar a la libertad de pensamiento y a la intimidad, en una primera aproximación.

b) El RIA identifica también como de alto riesgo varios sistemas empleados en el ámbito educativo, lo que interpela sin duda a las Administraciones públicas, tanto en relación con centros públicos como privados.

Se consideran, en efecto, de alto riesgo varios sistemas relacionados con la evaluación de las personas. Sea para determinar el acceso o la admisión a centros de educación y formación profesional a todos los niveles o para distribuir a las personas entre dichos centros. Sea para evaluar los resultados del aprendizaje, particularmente cuando los resultados así obtenidos se utilicen para orientar el proceso de aprendizaje. Sea para evaluar el nivel de educación adecuado que recibirá una persona o al que podrá acceder. Se considerarán también de alto riesgo los sistemas destinados a ser utilizados para el seguimiento y la detección de comportamientos prohibidos por parte de los estudiantes durante los exámenes en el contexto de los centros, todo ello en función de los efectos discriminatorios que puedan suponer.

Según argumenta el considerando 56 del RIA, el potencial efecto discriminatorio de este tipo de sistemas “cuando no se diseñan y utilizan correctamente” puede llegar a afectar a la capacidad de subsistencia de la persona, por condicionar su trayectoria formativa y profesional, y, en todo caso, puede violar el derecho a la educación y a la formación y el derecho a no sufrir discriminación, “además de perpetuar patrones históricos de discriminación”. Todo ello como excepción a un planteamiento en el que se pondera la importancia del despliegue de sistemas de IA en el ámbito educativo con el fin de fomentar una educación y formación digitales de alta calidad que asegure la tan deseada alfabetización mediática, que, como veremos, en el contexto del RIA se cualifica en relación con todos los involucrados en el manejo de los sistemas.

c) Determinados sistemas de IA dirigidos a la selección y gestión de empleados públicos quedan condicionados por

el RIA en su consideración de alto riesgo, por cuanto el reglamento europeo considera como tales varios sistemas relativos a la gestión de trabajadores, sin distinguos.

Se clasifican de alto riesgo los sistemas que se destinen a la contratación o selección de personas físicas, especialmente para publicar anuncios de empleo específicos, analizar y filtrar solicitudes de empleo y evaluar a candidatos, así como los que se utilicen para tomar decisiones que afecten a las condiciones de las relaciones “de índole laboral” -expresión que no debería impedir que se aplique igualmente a relaciones funcionariales- o para la promoción o rescisión de dichas relaciones, para la asignación de tareas a partir de comportamientos individuales o rasgos o características personales o para supervisar y evaluar el rendimiento y la conducta de las personas en el marco de dichas relaciones. El argumento último que ofrece el RIA para justificar esta clasificación es que pueden afectar a los medios de subsistencia de las personas y a los derechos de los trabajadores y perpetuar patrones históricos de discriminación, llegando a socavar derechos fundamentales, como la protección de datos personales y la intimidad, tal y como razona el considerando 57.

d) Las autoridades públicas aparecen expresamente reconocidas en el RIA como responsables del despliegue de algunos sistemas declarados de alto riesgo en relación con la prestación, estable o bajo condiciones de emergencia, de servicios que se califican de esenciales.

El Anexo III del RIA se refiere, en efecto, en particular, para reconocerlos como de alto riesgo, a los sistemas destinados, literalmente a:

Ser utilizados por las autoridades públicas o en su nombre para evaluar la admisibilidad de las personas físicas para acceder a prestaciones y servicios esenciales de asistencia pública, incluidos los de asistencia sanitaria, así como para conceder, reducir o retirar dichos servicios y prestaciones o reclamar su devolución (2024).

En este contexto el Reglamento se refiere expresamente a muy variadas prestaciones que se consideran necesarias para que las



personas puedan participar plenamente en la sociedad o mejorar su nivel de vida, como puedan ser los servicios de asistencia sanitaria, las prestaciones de seguridad social y los servicios sociales que garantizan protección en supuestos como la maternidad, la enfermedad, los accidentes laborales, la dependencia o la vejez, la pérdida de empleo, la asistencia social o las ayudas a la vivienda. Que se trate de prestaciones que suelen responder a situaciones de dependencia y vulnerabilidad “respecto de las autoridades responsables” y que pueden afectar a derechos fundamentales como el derecho a la protección social, a la no discriminación, a la dignidad humana o a la tutela judicial efectiva, según especifica el considerando 58 del RIA, está en la base de la clasificación como de alto riesgo de este amplio catálogo de sistemas de IA.

Esta clasificación se aplica igualmente a los sistemas destinados a ser utilizados para la evaluación y la clasificación de las llamadas de emergencia realizadas por personas físicas o el establecimiento de prioridades en el envío de servicios de primera intervención en situaciones de emergencia, como es el caso de la policía, bomberos y servicios de asistencia médica, incluyendo de forma expresa los “sistemas de triaje de pacientes en el contexto de la asistencia sanitaria de urgencia”. El propio RIA argumenta que se trata de sistemas que adoptan decisiones en situaciones sumamente críticas para la vida y la salud de las personas y de sus bienes, con lo que debe concluirse que los mismos sistemas no se consideran de alto si se aplican en situaciones ordinarias.

Consciente de la directa incidencia de ambas clasificaciones en un amplio espectro de actuaciones públicas, el propio Reglamento se cuida de advertir que su aplicación no debe conducir a obstaculizar:

El desarrollo y el uso de enfoques innovadores en la Administración pública, que se beneficiarían de una mayor utilización de sistemas de IA conformes y seguros, siempre y cuando dichos sistemas no conlleven un alto riesgo para las personas jurídicas y físicas (2024).

e) El apartado 6 del anexo III identifica como de alto riesgo una serie de sistemas englobables en el ámbito de la “aplicación de la ley”, expresión que el RIA refiere únicamente a supuestos de aplicación de la ley penal. Ello no ha de impedir, sin embargo, que se tomen como referencia en relación con la aplicación de la ley administrativa sancionadora, por cuanto tienen como fin último garantizar plenamente la protección de los derechos fundamentales en presencia.

El RIA identifica a estos efectos sistemas dirigidos a evaluar los riesgos de que una persona física sea víctima de infracciones penales –supuesto en el que entraría el modelo Viogen diseñado para la protección de mujeres víctimas de maltratos- y otros que se relacionan con la obtención o valoración de pruebas, como es el caso de los que consisten en polígrafos o similares y los destinados a evaluar la fiabilidad de las pruebas durante la investigación o el enjuiciamiento de infracciones penales. Así como los que tengan como objetivo evaluar la probabilidad de que una persona física cometa una infracción o reincida en la comisión de una infracción atendiendo no solo a la elaboración de perfiles de personas físicas referida en el art. 3.4 de la Directiva (UE) 2016/680⁽¹⁴⁾, o evaluar rasgos y características de la personalidad o comportamientos delictivos pasados de personas físicas o grupos y, finalmente, los sistemas destinados a elaborar esos mismos perfiles de personas físicas durante la detección, la investigación o el enjuiciamiento de infracciones penales.

El considerando 59 apela al desequilibrio de poder y los efectos sobre la propia libertad de las personas –estrictamente, físicas, conviene advertirlo- que es consustancial al uso de este tipo de sistemas de IA por parte de las autoridades encargadas de la aplicación de la ley, enfatizando, correlativamente, los riesgos que implicaría no someterlos a las reglas que el propio RIA establece para los sistemas de alto riesgo, por cuanto podrían conducir a señalar a personas “de manera discriminatoria, incorrecta e injusta” e impedir el ejercicio de importantes derechos procesales, como el derecho a la tutela judicial efectiva y a un juez imparcial, así como al derecho de defensa y a la presunción de inocencia si los sistemas no son lo suficientemente transparentes y

(14) Directiva (UE) 2016/680 del Parlamento europeo y del Consejo de 27 de abril de 2016 relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, y a la libre circulación de dichos datos y por la que se deroga la Decisión Marco 2008/977/JAI del Consejo, cuyo art. 3.4 define la «elaboración de perfiles» como “toda forma de tratamiento automatizado de datos personales consistente en utilizar datos personales para evaluar determinados aspectos personales de una persona física, en particular para analizar o predecir aspectos relativos al rendimiento profesional, situación económica, salud, preferencias personales, intereses, fiabilidad, comportamiento, ubicación o movimientos de dicha persona física”.



explicables ni están suficientemente bien documentados⁽¹⁵⁾. Se trata, en definitiva, de disciplinar los sistemas cuya precisión, fiabilidad y transparencia sean especialmente importantes para evitar consecuencias adversas, conservar la confianza de la población (sic) y garantizar la rendición de cuentas y una compensación efectiva.

Sin perder de vista el principio de proporcionalidad que late en la clasificación de los sistemas como de alto riesgo, no se puede negar que muchos de estos argumentos son perfectamente trasladables al ámbito de las infracciones administrativas. El hecho de que el mismo considerando 59 concluya negando que deban considerarse de alto riesgo los sistemas de IA destinados específicamente a que las autoridades fiscales y aduaneras los utilicen en procesos administrativos, significadamente en “tareas administrativas de análisis de información” en materia de lucha contra el blanqueo de capitales, no impide que desde la normativa administrativa o, en todo caso, por parte de la autoridad administrativa competente, en su caso por vía contractual, se incorporen algunas garantías relevantes extraídas del RIA para conjurar riesgos equivalentes⁽¹⁶⁾.

f) En el ámbito de la migración, el asilo y la gestión del control fronterizo, que por definición involucra a autoridades públicas, el RIA califica de alto riesgo diversos sistemas de IA.

Este es el caso de los polígrafos y similares; los sistemas que se utilicen por “las autoridades públicas competentes” (sic) para evaluar un riesgo que plantee una persona física que tenga la intención de entrar en el territorio de un Estado miembro o que ya haya entrado en él, como puedan ser riesgos de salud, seguridad o de inmigración irregular; los sistemas destinados a ayudar a estas mismas autoridades a examinar las solicitudes de asilo, visado o permiso de residencia o las correspondientes reclamaciones con el fin de comprobar el cumplimiento de los requisitos exigibles por parte de los solicitantes, con inclusión de la evaluación conexa de la fiabilidad de pruebas y, finalmente, aquellos cuyo objetivo sea detectar, reconocer o identificar a personas físicas, con excepción de la verificación de documentos de viaje.

El RIA vincula esta clasificación con el objetivo de proteger a quienes se encuentren en situación de especial vulnerabilidad y dependencia, lo que justifica que este tipo de sistemas deba

estar revestido de especiales garantías, particularmente asegurando su precisión y transparencia y la ausencia de sesgos, en consideración a los múltiples derechos fundamentales que pueden verse afectados por su uso. Particularmente los de libre circulación, no discriminación, a la intimidad personal y a la protección de datos personales y el de buena administración, según refiere el cdo. 60 del RIA, que no deja de invocar distintas normas europeas e internacionales que deben ser respetadas en todo caso, lo que sirve por sí mismo como recordatorio de que las propias instituciones europeas pueden ser proveedoras o responsables de despliegue de estos tipos de sistemas⁽¹⁷⁾.

g) El último epígrafe del Anexo III del RIA, el octavo, identifica como de alto riesgo una serie de sistemas que engloba bajo el enunciado “Administración de justicia y sistemas democráticos”, por referencia estricta a sistemas que se aplican a la actividad jurisdiccional o de resolución alternativa de conflictos, de una parte, y al desenvolvimiento de los procesos electorales, de otra. No obstante, ambos y en particular el primero, invitan a trasladar criterios de referencia para el manejo de herramientas de IA por parte de Administraciones públicas, tanto en las actuaciones de ejercicio de potestades públicas, como en el de revisión de sus actos.

El RIA identifica, en efecto, como de alto riesgo, de una parte, los sistemas utilizados por una autoridad judicial, o en su nombre, para “ayudarle”, literalmente, “en la investigación e interpretación de hechos y de la ley, así como en la aplicación de la ley a un conjunto concreto de hechos o a ser utilizados de forma similar en una resolución alternativa de litigios”.

(15) El mismo considerando 59 concluye significativamente afirmando que:

No debe ignorarse el impacto del uso de herramientas de IA en los derechos de defensa de los sospechosos, en particular la dificultad para obtener información significativa sobre el funcionamiento de dichos sistemas y la consiguiente dificultad para impugnar sus resultados ante los tribunales, en particular por parte de las personas físicas investigadas.

(16) *Al respecto, véase* “Propuesta de cláusulas contractuales tipo para la contratación de inteligencia artificial por parte de organismos públicos” de Public Buyers. (2023).

(17) El propio RIA, en su Anexo X, menciona el Reglamento por el que se establece un Sistema Europeo de Información y Autorización de Viajes (SEIAV) -Reglamento (UE) 2018/1240 del Parlamento y del Consejo, de 12 de septiembre de 2018, en que se contempla un sistema de reconocimiento biométrico de personas en las fronteras de la Unión.



El RIA identifica así un supuesto de uso de herramientas de IA en el ámbito jurídico verdaderamente revelador, ya que asume la posibilidad de utilizar sistemas de IA en la interpretación y la aplicación de la ley, si bien limitado, significativamente, a un mero carácter auxiliar y no sustitutivo. Carácter auxiliar que debe integrar, en todo caso, un grado de implicación superior a los supuestos que el considerando 61 del Reglamento deja expresamente extramuros de la consideración como sistema de alto riesgo como la anonimización o pseudoanonimización de resoluciones judiciales, documentos o datos o lo que se califican de “tareas administrativas”, asumiendo implícitamente la relevante distinción entre informática instrumental e informática decisonal. Para evitar equívocos, el considerando se anima a afirmar rotundamente que “la utilización de herramientas de IA puede apoyar el poder de decisión de los jueces o la independencia judicial, pero no debe sustituirlas: la toma de decisiones finales debe seguir siendo una actividad humana”.

Bajo esta lógica habrá que interpretar, sin duda, las previsiones del Real Decreto-ley 6/2023, de 19 de diciembre, por el que se aprueban medidas urgentes para la ejecución del Plan de Recuperación, Transformación y Resiliencia en materia de servicio público de justicia, función pública, régimen local y mecenazgo, que contempla el manejo de sistemas de IA en las actividades de la Administración de justicia en distintas escalas. Su art. 57 distingue -frente a las actuaciones automatizadas y proactivas que se identifican en el art. 56⁽¹⁸⁾ -, las que califica como “actuaciones asistidas”, que se definen en el apartado 1 como aquellas para las que “el sistema de información de la Administración de Justicia genera un borrador total o parcial de documento complejo basado en datos, que puede ser producido por algoritmos, y puede constituir fundamento o apoyo de una resolución judicial o procesal”.

El carácter auxiliar de estas “actuaciones asistidas” es enfatizado en los sucesivos apartados 2 y 3, en los que se reafirma que “en ningún caso el borrador documental así generado constituirá por sí una resolución judicial o procesal, sin validación de la autoridad competente”, de modo que “los sistemas de la Administración de Justicia asegurarán que el borrador documental sólo se genere a voluntad del usuario y pueda ser libre y enteramente modificado por éste” (apartado 2), de modo que, según refiere el apartado 3:

La constitución de resolución judicial o procesal requerirá siempre la validación del texto definitivo, por el juez o jueza, magistrado

o magistrada, fiscal o letrado o letrada de la Administración de Justicia, en el ámbito de sus respectivas competencias y bajo su responsabilidad, así como la identificación, autenticación o firma electrónica que en cada caso prevea la ley, además de los requisitos que las leyes procesales establezcan (2024).

Las razones que avalan confesadamente en el considerando 61 del RIA la clasificación de este tipo de herramientas de IA como de alto riesgo son perfectamente extrapolables al ámbito del ejercicio de potestades administrativas de alcance decisor. A pesar de que en el ámbito de la actuación administrativa no se presentan en plenitud los riesgos de afección a:

La democracia, el Estado de Derecho, las libertades individuales y el derecho a la tutela judicial efectiva y a un juez imparcial” que vienen a justificar la clasificación como de alto riesgo, no cabe duda de que en el ejercicio de potestades administrativas es crítico evitar los posibles “sesgos, errores y opacidades que pudieran producirse (2024).

El segundo supuesto contemplado en el octavo y último epígrafe del Anexo III presenta un cariz distinto. Identifica como de alto riesgo los sistemas de IA destinados a ser utilizados para influir en el resultado de una elección o referéndum o en el comportamiento electoral de personas físicas que ejerzan el derecho al voto, en supuestos en los que pueden verse involucrados actores públicos, como puedan ser los medios de comunicación de titularidad pública o las instituciones públicas de análisis sociológico. La lógica de esta clasificación no es otra que proteger a los que ejercen el derecho de sufragio activo e, indirectamente, los valores de la democracia y el Estado de Derecho, según expresa el considerando 62 del RIA, que, coherentemente, excluye de forma expresa de esta consideración los sistemas a cuya información de salida no estén expuestas las personas físicas

(18) En los términos del art. 56, se entiende por actuación automatizada la actuación procesal producida por un sistema de información adecuadamente programado sin necesidad de intervención humana en cada caso singular, precisando que los sistemas informáticos utilizados en la Administración de Justicia posibilitarán la automatización de las actuaciones de trámite o resolutorias simples, que no requieren interpretación jurídica, como el numerado o paginado de los expedientes o la comprobación de representaciones o la declaración de firmeza, de acuerdo con la ley procesal. Y entre ellas, se consideran proactivas las auto-iniciadas por los sistemas de información sin intervención humana “que aprovechan la información incorporada en un expediente o procedimiento de una Administración Pública con un fin determinado, para generar avisos o efectos directos a otros fines distintos, en el mismo o en otros expedientes, de la misma o de otra Administración Pública, en todo caso conformes con la ley”.



como puedan ser las herramientas utilizadas para organizar, optimizar o estructurar campañas políticas desde un punto de vista administrativo o logístico.

4.3. El contenido de la evaluación de impacto

Una vez identificados, aun cuando sea de forma somera, los distintos sistemas que, considerados de alto riesgo por el RIA, pueden ser típicamente desplegados por organismos de Derecho público, en los términos del art. 27, cumple aproximarse a las características de la evaluación de impacto en los derechos fundamentales que este precepto les exige completar con carácter previo al despliegue. Empezando por su contenido, que busca lograr una delimitación “precisa” de los riesgos generados por la utilización del sistema y las medidas de reacción previstas en el concreto contexto de su utilización.

El apartado 1 del artículo de referencia exige una serie de contenidos dirigidos, de una parte, a delimitar los términos de utilización del sistema, y, de otra, a identificar las medidas de control y reacción. En el primer bloque se integraría una descripción de los procesos del responsable del despliegue en los que se utilizará el sistema en consonancia con su finalidad prevista; una descripción del período de tiempo durante el que está previsto su uso y de la frecuencia de utilización prevista; la identificación de las categorías de personas físicas o grupos que puedan verse afectados por su utilización en un contexto específico y la de los riesgos de perjuicios específicos que puedan afectarles tomando en consideración la información facilitada por el proveedor en los términos del art. 13, que se refiere en particular a las instrucciones de uso del “producto IA”.

Entre los contenidos concentrados en identificar las medidas de control y reacción se incluyen la descripción de la aplicación de medidas de supervisión humana, de acuerdo con las instrucciones de uso, y la identificación de las medidas que deben adoptarse en caso de que se materialicen los riesgos que hayan sido identificados, incluyendo los acuerdos de gobernanza interna y los mecanismos de reclamación.

Aunque no lo especifique el precepto, los riesgos de que se trata son los riesgos de afección a los derechos fundamentales en presencia. Derechos fundamentales que, como hemos apuntado, pueden ser muy variados, pero entre los que puede concurrir comúnmente el derecho a la protección de datos personales, para el que, como es sabido, se ha concebido una evaluación específica, con sus propias coordenadas, en el art. 35 del RGPD y en el art. 27 de la Directiva (UE) 2016/680. Esta evaluación busca, de hecho, valorar si en el tratamiento de los datos en presencia concurre un riesgo que lo desaconseje, mientras que la evaluación del art. 27 del RIA parte de que el sistema a utilizar es de alto riesgo, pero no cabe duda de que hay espacios de solape -por lo que hace a la valoración del riesgo en la protección de datos- que justifican que el apartado 4 de aquel artículo prevea que cuando se haya llevado a cabo la

evaluación relativa a la protección de datos la evaluación del RIA la “complementará”.

4.4. Cuestiones procedimentales

Esta última referencia al sentido con el que se concibe la evaluación relativa a la protección de datos debe extrapolarse, aunque no se extraiga de la literalidad del art. 27 del RIA, al sentido debido de la evaluación que el mismo concibe. Y es que esta evaluación no puede ser un fin en sí mismo, sino que debe plantearse como un momento de debida reflexión que conduzca, en su caso, a la decisión de rechazar el despliegue del sistema de IA proyectado. Ese es necesariamente su sentido último, porque no cabe duda de que lo que se trata en último término es de ponderar si los beneficios a los que se aspira con la introducción de un sistema de IA del alto riesgo compensan frente a los riesgos sobre los derechos fundamentales que genera a la vista de las medidas de reacción disponibles.

Esta debida reflexión es lo que justifica la introducción de esta carga, que en todo caso se pretende aligerar a partir de varias reglas que contempla el mismo art. 27. Según precisa su apartado 2, la evaluación de impacto solo se requiere respecto del primer uso del sistema de IA de alto riesgo, siendo posible para el responsable del despliegue, ante casos similares, basarse en evaluaciones de impacto realizadas previamente o a evaluaciones de impacto existentes realizadas por los proveedores. En la misma lógica facilitadora abunda la previsión relativa a la posible exigencia paralela de la evaluación de protección de datos, según quedó descrita en los términos del apartado 4. Así como la regla que en el mismo apartado 2 exige, en el caso de cambios en el contenido de alguno de los elementos contemplados en el apartado 1, una mera actualización de la evaluación original, y no una nueva evaluación completa, como es lógico.

Con claro afán facilitador del proceso, se ha previsto que la Oficina de IA de la Comisión Europea elabore un modelo de cuestionario, también mediante una herramienta automatizada. Así lo compromete el apartado 5 del mismo art. 27, sin que ello deba conducir a entender la evaluación como un mero formalismo que la vacíe de sentido.



Una vez realizada la evaluación el responsable del despliegue notificará sus resultados a la autoridad de vigilancia del mercado, presentando el modelo de referencia debidamente cumplimentado. Solamente en el caso de que concluya que el despliegue del sistema es oportuno, hay que colegir en coherencia con nuestras consideraciones previas.

El considerando 96 –no así el propio art. 27- hace una llamada de atención, particularmente cuando los sistemas evaluados se utilicen en el sector público, a la posible participación de las “partes interesadas pertinentes”, entre las que menciona a los representantes de grupos de personas que probablemente vayan a verse afectados, expertos independientes u organizaciones de la sociedad civil. A pesar de que esta participación se plantee en términos meramente potestativos, resulta más que conveniente y debería, como la evaluación misma, ampliarse a todos los supuestos de valoración de cualesquiera sistemas de IA por Administraciones públicas, sean o no de alto riesgo, viniendo, de recabarse, a desmontar el argumento práctico más robusto de los que se han manejado para defender la naturaleza reglamentaria de los algoritmos⁽¹⁹⁾.

5. Alfabetización y supervisión humana como mecanismos claves de garantía

No podemos cerrar esta brújula sobre el despliegue de los sistemas de IA por parte de Administraciones Públicas sin hacer un par de apreciaciones sobre dos cuestiones de importancia que han surgido a lo largo de estas reflexiones: la de la alfabetización en IA y la de la tan traída y llevada supervisión humana.

El despliegue de sistemas de IA por Administraciones Públicas es un reto para las mismas como organizaciones y, en último término, para sus autoridades y el personal a su servicio. Por ello es crucial desplegar esfuerzos para que todos los que estén involucrados en la utilización de los sistemas de IA tengan un “nivel suficiente” de alfabetización en materia de IA que tome en consideración sus conocimientos técnicos, su experiencia, su educación y su formación, así como el contexto previsto de uso y las personas o grupos respecto de los que se utilizarán los sistemas.

Todo ello “en la mayor medida posible”, según precisa el art. 4 del RIA, y en el bien entendido de que tal alfabetización se identifica por el art. 3.56 con las capacidades, los conocimientos y la comprensión que permiten -por lo que ahora interesa particularmente- a los responsables del despliegue “llevar a cabo un despliegue informado de los sistemas de IA

y tomar conciencia de las oportunidades y los riesgos que plantea la IA, así como de los perjuicios que puede causar”. Tal y como expresa elocuentemente el considerando 20 del RIA, de lo que se trata es de dotar a los responsables del despliegue de “los conceptos necesarios para tomar decisiones con conocimiento de causa en relación con los sistemas de IA”.

Entre los que han de tomar decisiones, por momentos, cruciales, en relación con los sistemas de IA están aquellos a los que se les encomiende la supervisión humana de los resultados que arroje. Supervisión humana que se ofrece como un mantra salvífico que ha de proteger de la materialización de los riesgos que los sistemas presenten⁽²⁰⁾, pero que ha de ser necesariamente protocolizada para evitar que se desenvuelva con automatismos -la famosa “palanca de hombre muerto”- que inhabiliten su sentido propio. Separarse del sentido de lo propuesto por la máquina es un riesgo para el humano que ha de tomar la decisión. Un riesgo del que deriva una responsabilidad. Protocolizar ese momento dándole razonables garantías en caso de que incurra en un error es imprescindible para que esa función se desenvuelva con la virtualidad para la que ha sido concebida.

6. A modo de breve conclusión

Llegados a este punto, se puede confirmar la reflexión apuntada en la introducción. Las herramientas de inteligencia artificial, en su heterogeneidad y versatilidad, ofrecen múltiples aplicaciones a las Administraciones Públicas que pueden facilitar, e incluso optimizar, el logro de los intereses generales. Pero estas aplicaciones no están exentas de riesgos, particularmente para los derechos fundamentales.

Esta conclusión queda perfectamente acreditada a la vista del repaso de los sistemas que el RIA considera de alto riesgo, sometiéndolos a un instrumento

(19) Para conocer una posición dialéctica sobre esta polémica, véase, “Algoritmos y reglamentos” de Arroyo Jiménez (2020).

(20) Así lo plantea, por ejemplo, Ponce (2020); véase “Reserva de humanidad y supervisión humana de la Inteligencia Artificial”.



de control previo cual es la evaluación de impacto en los derechos fundamentales. Este instrumento, cuyas claves procedimentales y de contenido han quedado explicadas, tiene como primera virtud su carácter obligatorio, que impone a las Administraciones Públicas una reflexión sustanciada previa a la decisión de introducir un sistema de IA de alto riesgo en sus procedimientos. De este modo la Administración podrá ponderar si los beneficios previstos compensan los riesgos potenciales y protocolizar cuestiones claves como la famosa supervisión humana en conexión con la debida alfabetización de quienes manejen los sistemas.

El RIA diseña así un instrumento promotor cuyas virtudes prácticas están por comprobar.

Referencias Bibliográficas

Álvarez García, A., & Tahirí Moreno, J. (2023). La regulación de la inteligencia artificial en Europa a través de la técnica armonizadora del nuevo enfoque. *Revista General de Derecho Administrativo (RGDA)*, (63).

Arroyo Jiménez, L. (2020, febrero 25). *Algoritmos y reglamentos*. Almacén de Derecho. <https://almacendederecho.org/algoritmos-y-reglamentos>

Barrio Andrés, M. (Dir.). (2024a). *El Reglamento Europeo de Inteligencia Artificial*. Tirant lo Blanch.

Barrio Andrés, M. (Dir.). (2024b). *Comentarios al Reglamento Europeo de Inteligencia Artificial*. La Ley.

Bermejo Vera, J. (2024, enero 25–27). *La aplicación de la inteligencia artificial en la actividad formal e informal de la Administración* [Ponencia]. XVII Congreso de la Asociación Española de Profesores de Derecho Administrativo, Universidad de Vigo. <https://congresoaepepdavigo2024.es/ponencias/>

Berning Prieto, A. D. (2023). El uso de sistemas basados en inteligencia artificial por las Administraciones públicas: estado actual de la cuestión y algunas propuestas ad futurum para un uso responsable. *Revista de Estudios de la Administración Local y Autonómica (REALA)*, (20). <https://doi.org/10.24965/reala.11247>

Boden, M. A. (2022). *Inteligencia artificial*. Turner.

Carlón Ruiz, M. (2025). *Las Administraciones públicas ante la inteligencia artificial*. Tirant lo Blanch.

Cotino Hueso, L., & Simón Castellano, P. (Dir.). (2025). *Tratado sobre el Reglamento Europeo de Inteligencia Artificial*. Aranzadi.

Comisión Europea. (2025). *Directrices relativas a la definición de sistema de inteligencia artificial establecida en el Reglamento de Inteligencia Artificial* (C(2025) 5053 final).

Ebers, M. (2024). *Truly risk-based regulation of artificial intelligence: How to implement the EU's AI Act*. European Union Law Working Papers, (101), Stanford–Vienna Transatlantic Technology Law Forum. Vienna Transatlantic Technology Law Forum.

Gamero Casado, E. (Dir.). (2023). *Inteligencia artificial y sector público: Retos, límites y medios*. Tirant lo Blanch.

Ponce, J. (2022). *Reserva de humanidad y supervisión humana de la inteligencia artificial*. El Cronista del Estado Social y Democrático de Derecho, (100).

Presno Linera, M. (2022). *Derechos fundamentales e inteligencia artificial*. Tecnos. <https://doi.org/10.20318/dyl.2023.7729>

Presno Linera, M. (2025). El Reglamento europeo de inteligencia artificial: ¿garantía del mercado o defensa de los derechos? *Revista Española de Derecho Constitucional*, (135). Presno Linera, M. (2025). El Reglamento europeo de inteligencia artificial: ¿garantía del mercado o defensa de los derechos? *Revista Española de Derecho Constitucional*, (135). <https://doi.org/10.18042/cepc/redc.135.01>

Public Buyers Community. (2023, septiembre 29). *Propuesta de cláusulas contractuales tipo para la contratación de inteligencia artificial por parte de organismos públicos*. <https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/eu-model-contractual-ai-clauses-pilot-procurements-ai>

Tangi, L., Van Noordt, C., Combetti, M., Gattwinkel, D., & Pignatelli, F. (2022). *AI Watch: European landscape on the use of artificial intelligence by the public sector* (EUR 31088 EN, JRC129301). Publications Office of the European Union. <https://doi.org/10.2760/39336> 