¿CUÁNDO "NO ES NO"? CRITERIOS PARA DEFINIR LOS SISTEMAS DE INTELIGENCIA ARTIFICIAL PROHIBIDOS EN LA UNIÓN EUROPEA

Por

LORENZO COTINO HUESO [1]
Catedrático de Derecho Constitucional
Universidad de Valencia

cotino@uv.es

Revista General de Derecho Administrativo 69 (2025)

RESUMEN: El Reglamento de Inteligencia Artificial de la Unión Europea se aplica desde febrero de 2025 respecto de las prohibiciones que establece su artículo 5. Este artículo prohíbe sistemas de IA que impliquen manipulación subliminal, explotación de vulnerabilidades, puntuación social, predicción delictiva basada en perfiles, rastreo masivo de imágenes faciales, inferencia de emociones en entornos laborales y educativos, categorización biométrica para inferir características sensibles y el uso de identificación biométrica remota en tiempo real en espacios públicos con fines policiales. En ocasiones las prohibiciones se formulan como absolutas y en otras ocasiones se prevén expresamente excepciones. Lo cierto es que más que un "no es no", en la mayoría de los supuestos se configura un "no, pero sí", por lo que las prohibiciones requieren una interpretación detallada en cuanto a sus elementos, alcance y exclusiones, así como respecto a las condiciones que permiten aplicar excepciones. El estudio analiza los Criterios establecidos por la Comisión Europea al respecto de cada prohibición. Entre ellas, tiene especial relevancia el uso policial de identificación biométrica remota en tiempo real y su detallado régimen de excepciones que requiere autorización judicial previa, limitaciones temporales, geográficas y personales, así como mecanismos de supervisión reforzada. Además, el estudio analiza la delgada línea que separa los sistemas prohibidos de los de alto riesgo

PALABRAS CLAVE: inteligencia artificial, sistemas prohibidos, Reglamento inteligencia artificial, Unión Europea, derechos fundamentales.

SUMARIO: I. La necesidad de concretar cuándo "no es no" en la inteligencia artificial o si se trata de un "no pero sí"; II. La prohibición de la manipulación subliminal y explotación de vulnerabilidades (artículo 5.1.a y b); III. Cuándo se considera un "social scoring" prohibido (artículo 5, 1 c); IV. Cuándo se da la prohibición sobre predicción delictiva sobre perfiles y personalidad del artículo 5, 1, d); V. Los requisitos de la prohibición de rastreo no selectivo de imágenes faciales, artículo 5, 1, e); VI. La prohibición de sistemas de inferencia de emociones en el trabajo y la educación(artículo 5.1.f); VII. La prohibición de sistemas de categorización biométrica para inferir características especialmente sensibles de las personas (artículo 5.1.g); VIII. El alcance de la prohibición del uso de sistemas de identificación biométrica a distancia en tiempo real en espacios

[1] Investigador de la Universidad Católica de Colombia. Valgrai. El presente estudio es resultado de investigación de los siguientes proyectos: proyecto "Derecho, Cambio Climático y Big Data", Grupo de Investigación en Derecho Público y TIC como investigador de la Universidad Católica de Colombia; MICINN Proyecto "Derechos y garantías públicas frente a las decisiones automatizadas y el sesgo y discriminación algorítmicas" 2023-2025 (PID2022-136439OB-I00) financiado por MCIN/AEI/10.13039/501100011033/; Convenio de Derechos Digitales-SEDIA Ámbito 5 (2023/C046/00228673) y Ámbito 6. (2023/C046/00229475).

públicos con fines policiales (art. 5.1. h); IX. Las excepciones a la prohibición del artículo 5.1. H); X. Requisitos, salvaguardias y condiciones para las excepciones del artículo 5.1. H); X. Para concluir.

WHEN DOES “NO MEANS NO”? CRITERIA FOR DEFINING ARTIFICIAL INTELLIGENCE SYSTEMS PROHIBITED IN THE EUROPEAN UNION

ABSTRACT: The European Union's Artificial Intelligence Regulation is applicable as of February 2025 with regard to the prohibitions established in Article 5. It prohibits AI systems that involve subliminal manipulation, exploitation of vulnerabilities, social scoring, crime prediction based solely on profiling, mass facial image tracking, emotion inference in work and educational environments, biometric categorization to infer sensitive characteristics, and the use of real-time remote biometric identification in public spaces for law enforcement purposes. Sometimes the prohibitions are formulated as absolute, while in others, explicit exceptions are provided. Rather than a strict ‘no means no’ approach, most cases fall into a nuanced ‘no, but with exceptions’ framework. Therefore, the prohibitions require interpretation concerning their specific elements, scope, and exclusions, as well as, where applicable, the potential applicability of exceptions. The study analyses the criteria established by the European Commission in relation to each prohibition. It is also worth emphasising the relevance of the police use of remote biometric identification in real time, which is subject to a strict exception regime requiring prior judicial authorization, temporal, geographical and personal limitations, as well as reinforced supervision mechanisms. In addition, the study takes into account the fine line between prohibited systems and high-risk systems.

KEYWORDS: artificial intelligence, prohibited systems, Artificial Intelligence Regulation, European Union, fundamental rights.

Fecha de recepción: 06/02/2025

Fecha de aceptación: 14/05/2025

I. LA NECESIDAD DE CONCRETAR CUÁNDO “NO ES NO” EN LA INTELIGENCIA ARTIFICIAL O SI SE TRATA DE UN “NO PERO SÍ”

El 4 de febrero de 2025 se dio a conocer el proyecto de Directrices de la Comisión Europea sobre prácticas prohibidas de inteligencia artificial (IA) en artículo 5 del

Reglamento (UE) 2024/1689 de IA (RIA)[2] que se adoptarán a través de Comunicación. El artículo 96 1 b), RIA habilita a la Comisión a adoptar directrices sobre la aplicación práctica de este artículo. Dos motivos impulsan a hacerlo. De un lado, la IA prohibida es la primera parte del RIA que se aplica, desde el 2 de febrero de 2025. Por otro lado, la extensión de este artículo 5, con sus 2000 palabras, su complejidad y los numerosos espacios que deja abiertos a la indefinición.

Es por ello que la Oficina de Inteligencia Artificial de la Comisión el 13 de noviembre puso en marcha un proceso de consulta sobre estas directrices respecto de la definición de sistemas de IA, consulta abierta a la participación hasta el 11 de diciembre de 2024.[3] Dicha consulta incluía un completo formulario de 42 páginas, con diez mil palabras en los que se preguntaban 31 cuestiones y se daban 43 espacios para realizar contribuciones sobre los criterios de interpretación y delimitación de las prohibiciones.[4] La consulta solicitaba también ejemplos prácticos adicionales para mayor claridad sobre los aspectos prácticos y los casos de uso.

Pues bien, el resultado de la consulta y los estudios relacionados han dado lugar a un proyecto de Comunicación con un extenso anexo titulado *Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)*[5]. Se trata de un estructurado y prolijo documento de 140 páginas y casi 65 mil palabras que se numeran en 434 párrafos que aquí se indican entre paréntesis. El documento está en inglés, si bien he dispuesto una versión traducida automatizadamente al españoI.[6] Cabe recordar que el RIA tiene 106 mil palabras y, como se ha dicho el artículo 5 que se interpreta en este anexo 2000 palabras de gran complejidad jurídica. Obviamente estas páginas no son un estudio sobre el artículo 5 RIA y las prohibiciones que establece, sino que, sobre la base de un conocimiento de dicho artículo se centra en extraer el contenido

---

[2] Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas sobre inteligencia artificiaI y se modifican determinados reglamentos.

[3] https://digital-strategy.ec.europa.eu/es/news/commission-launches-consultation-ai-act-prohibitions-and-ai-system-definition

[4] El documento ya no está disponible en la red, si bien he dispuesto un acceso al mismo en https://www.uv.es/cotino/varios/Prohibitions-Survey-2024_13_11_2024_EN.docx

[5] https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act

acceso al documento en https://ec.europa.eu/newsroom/dae/redirection/document/112367

[6] https://disco.uv.es/disco?fileman:es:principal:user:web/varios::4ni#:~:text=C_2025_884_1_EN_annexeespan%CC%83oI.pdf

más relevante de los criterios interpretativos que se derivan de estas amplias Directrices de la Comisión. Me permito remitir a los trabajos más completos sobre el particular.[7]

Baste recordar ahora que el artículo 5 RIA si se me permite, establece el "no es no" de la IA. Ahí se prohíben, esencialmente, sistemas IA que faciliten la manipulación, explotación, control social o vigilancia masiva, por considerarse inaceptables e incompatibles con los valores y derechos fundamentales de la Unión. En concreto ello se articula a través de ocho prohibiciones de sistemas de IA[8], cuatro de las prohibiciones en principio son verdaderos "no es no", esto es, son absolutas (letras a, b, c y e). Sin embargo, el resto se tratan más bien de un "no pero sí", esto es, se contemplan excepciones a las prohibiciones(letras d, f, g y, en especial h). Cabe destacar que el sistema de excepciones respecto de la letra h (prohibición de la identificación biométrica remota en tiempo real en espacios públicos para fines policiales) ocupa buena parte de todo el artículo 5.

Pues bien, en todos los casos y en especial por la técnica legislativa empleada, es necesaria la interpretación para delimitar el alcance de la prohibición. Asimismo, es preciso determinar las exclusiones de ámbito y alcance de cada precepto y, en el caso de las excepciones, se hace preciso interpretar y concretar el sistema de excepciones a las prohibiciones y los requisitos que se establecen en muchos supuestos con habilitaciones a la regulación legal de los Estados o de la propia UE.

Cabe adelantar que los sistemas que no están efectivamente prohibidos por el artículo 5 RIA, pasan por lo general a ser sistemas IA -permitidos- de "alto riesgo" del Anexo III,[9] que he analizado en otros estudios. Como ha examinado excepcionalmente Palma Ortigosa hay una "delgada línea entre los sistemas de inteligencia artificial

[7] Me remito en particular a los estudios sobre el artículo 5 RIA que serán citados en Cotino Hueso, L. y Simó Castellanos, P. (coords.), *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea*, Aranzadi La Ley, Cízur Menor, 2024. Se dedican un comentario al "artículo 5" en Barrio Andrés, Moisés, Comentarios al Reglamento Europeo de Inteligencia Artificial, La Ley, Madrid, 2024, pp. 180-205.

[8] Manipulación subliminal o engañosa (letra a); explotación de vulnerabilidades (letra b); sistemas de puntuación social (letra c); evaluaciones de riesgo delictivo (letra d); creación de bases de datos de reconocimiento facial (letra e); inferencia de emociones en lugares de trabajo y centros educativos (letra f); categorización biométrica para inferir características sensibles (letra g); uso de identificación biométrica remota en tiempo real en espacios públicos para fines policiales (letra h).

[9] Como se recuerda en nº 37, los sistemas de reconocimiento de emociones no prohibidos en el artículo 5.1º f) RIA, serán considerados sistemas de alto riesgo conforme al artículo 6.2º y al anexo III.1º c). De manera similar, ciertos sistemas de puntuación basados en IA, como los utilizados en la evaluación crediticia o en la valoración del riesgo en seguros de salud y vida, serán clasificados como de alto riesgo siempre que no entren en el ámbito de prohibición previsto en el artículo 5.1º c). Asimismo, los sistemas de IA empleados para determinar la elegibilidad de las personas a prestaciones y servicios esenciales de asistencia pública, incluyendo la sanidad y la seguridad social, serán calificados como de alto riesgo. No obstante, si tales sistemas generan una puntuación social inadmisible y cumplen los criterios establecidos en el artículo 5.1º c), su comercialización, puesta en servicio y uso estarán expresamente prohibidos.

prohibidos y de alto riesgo en el Reglamento de IA"[10]. Su estudio y ahora estas Directrices de la Comisión Europea que aquí se destilan suponen un *manual* de cuándo no es no en IA y cuándo se trata de un no, pero sí. Espero que resulte de utilidad.

Resulta necesario hacer una advertencia en todo caso. Hay insistir por su importancia en algo que es clave, que un sistema IA no esté prohibido y esté regulado como sistema de alto riesgo por el RIA no quiere decir, ni mucho menos, que esté permitido. Puede haber muchos otros motivos para que un sistema no sea legal: que sea contrario a derechos fundamentales o al derecho a no ser discriminado, que no tenga base legal, por ejemplo, desde el punto de vista de la protección de datos o, por ejemplo, que no cumpla los requisitos del Derecho de la UE. De igual modo, hay que insistir, como hace expresamente el Considerando 63 RIA[11] y el nº 46 de estos Criterios, que el RIA no es base legal de legitimación para el tratamiento de datos personales, esto es, que si un sistema es regulado como de alto riesgo por el RIA, no implica que esta regulación por un Reglamento de la UE sirva de base para hacer ese tratamiento de datos que requiere de una base legal. Esta regla general es clara y considero que debe aplicarse a todos los sistemas de alto riesgo, siendo especialmente relevante respecto de los del Anexo III.[12]

Estas Directrices son útiles más allá del ámbito de las prohibiciones. No en vano, son las primeras Directrices interpretativas que perfilan otros conceptos básicos del RIA como la "puesta en el mercado", "puesta en servicio" o "uso" de un sistema de IA. También abordan el ámbito personal del reglamento como los proveedores e implementadores o usuarios. Igualmente, las exclusiones de aplicación del RIA (seguridad nacional, la defensa y los fines militares, la cooperación judicial y policial con terceros países, la investigación y el desarrollo, la actividad personal no profesional y los sistemas de IA con licencias libres y de código abierto).

Cabe por último señalar que los Criterios afirman expresamente que "Se esfuerzan por interpretar las prohibiciones de una manera proporcionada que logre los objetivos del RIA de proteger los derechos fundamentales y la seguridad, al tiempo que promueven la innovación y proporcionan seguridad jurídica." (nº 5).

---

[10] El mismo será publicado próximamente en *IDP. Revista de Internet, Derecho y Política*, nº 42 2025.

[11] "El hecho de que un sistema de IA sea clasificado como un sistema de IA de alto riesgo en virtud del presente Reglamento no debe interpretarse como indicador de que su uso sea legal con arreglo a otros actos del Derecho de la Unión o del Derecho nacional compatible con el Derecho de la Unión [...] No debe entenderse que el presente Reglamento constituye un fundamento jurídico [...] salvo que el presente Reglamento disponga específicamente otra cosa."

[12] Mi estudio "Alcance y delimitación de los sistemas de alto riesgo en el Reglamento de inteligencia artificial", en Cotino Hueso, L. y Simó Castellanos, P. (coords.), *Tratado sobre el Reglamento … cit.* Acceso

II. LA PROHIBICIÓN DE LA MANIPULACIÓN SUBLIMINAL Y EXPLOTACIÓN DE VULNERABILIDADES (ARTÍCULO 5.1.A Y B)

El artículo 5.1º RIA[13] prohíbe la comercialización, puesta en servicio o uso de sistemas de IA que empleen técnicas subliminales, manipuladoras o engañosas que alteren significativamente el comportamiento de las personas, limitando su capacidad de decisión informada y causando o pudiendo causar perjuicios graves (letra a). Asimismo, se prohíbe la IA que explote vulnerabilidades derivadas de la edad, discapacidad o situación socioeconómica para alterar el comportamiento de la persona y que tenga el potencial de generar daño (letra b). Se trata de unas prohibiciones que requieren no pocas concreciones. Así, uno de los aspectos más problemáticos del artículo 5. 1 a y b RIA es la determinación de cuándo una técnica de IA "altera de manera sustancial el comportamiento de una persona o un colectivo de personas" de una manera que pueda calificarse como manipulativa o explotadora. Y para ello se introducen criterios específicos (nº 58-145).

Primero, se apunta el concepto de "probabilidad razonable" de daño (nº 94 y ss.), lo que implica que no es necesario demostrar que una persona específica ha sido manipulada o explotada para que la prohibición sea aplicable. Basta con que el sistema de IA tenga una capacidad verificable de producir este efecto de manera consistente con el comportamiento de la persona o grupo afectado. Se propone reforzar la exigencia de que la determinación del nexo causal entre la técnica empleada y el daño significativo se base en criterios técnicos y científicos universalmente aceptados (nº 94).

En segundo lugar, es necesario *diferenciar con precisión entre influencia permitida y manipulación prohibida*, esto es, entre persuasión o influencia que sí que es lícita y manipulación prohibida, dejando claro que la mera influencia sobre la toma de decisiones humanas no basta para considerar ilícito un sistema de IA. La clave radica en si el sistema menoscaba de manera apreciable la capacidad de las personas para tomar decisiones informadas (nº 84). Así, el artículo 5 distingue entre los sistemas de IA que simplemente ejercen una influencia sobre el usuario y aquellos que distorsionan materialmente su comportamiento de manera que menoscaban su autonomía individual. A efectos de interpretación, se deben valorar los siguientes elementos: el grado de control del usuario sobre la información que recibe. Un sistema de IA puede ser legal en un entorno y prohibido en otro, dependiendo de la capacidad de la persona afectada para detectar y resistir la influencia que ejerce el sistema. También, la variabilidad en la

[13] Sobre el tema puede seguirse "El contenido de las llamadas "técnicas subliminales" y las vulnerabilidades de grupo específico de personas en el Reglamento de inteligencia artificial", por Luis Miguel González de la Garza, en Cotino Hueso, L. y Simó Castellanos, P. (coords.), *Tratado sobre el Reglamento.. cit*.

prohibición de técnicas subliminales en función del tipo de estímulo. La inmersión en realidad virtual o interfaces cerebro-máquina amplifica la influencia de la IA, reduciendo la capacidad de discernimiento del usuario y aumentando el riesgo de manipulación encubierta. En estos casos, el criterio de evaluación debe ser más restrictivo, esto es, favorable a considerar la prohibición, dado el alto riesgo de alteración de la percepción y el comportamiento sin plena consciencia del usuario.

En tercer lugar, hay que *determinar el daño significativo: tipos, umbral de importancia y nexo causal*. Para aplicar la prohibición prevista en el artículo 5.1.a, el daño causado debe ser significativo y se establecen tres criterios fundamentales:

a) Tipos de daños considerados. Los perjuicios derivados de la manipulación o explotación de vulnerabilidades mediante IA incluyen: daños físicos, esto es, lesiones, deterioro de la salud y afectación de la integridad corporal (nº 87). También, daños psicológicos: efectos adversos sobre la salud mental, que pueden ser acumulativos y no inmediatamente perceptibles, pero con impacto prolongado (nº 88) y perjuicios o pérdidas económicas, exclusión financiera o inestabilidad (nº 89).

b) La determinación del umbral de importancia del daño (nº 91-93). Un daño se considera significativo cuando afecta de manera sustancial la autonomía o el bienestar de las personas y para ello hay que tener en cuenta la gravedad del daño y perjuicio que ha resultado o que es razonablemente probable que resulte del uso del sistema de IA. Asimismo, el contexto y los efectos acumulativos son relevantes para valorar daño, puesto que no siempre es inmediato y puede derivarse de interacciones repetidas con el sistema. En cuanto a la escala e intensidad del daño, se han de tener en cuenta el número de personas afectadas y la intensidad del perjuicio. Además, la vulnerabilidad de las personas afectadas incrementa la percepción del daño, particularmente en el caso de colectivos como niños, ancianos, personas con discapacidad o en situación de desventaja socioeconómica. Finalmente se valora el daño según si los efectos son irreversibles o con consecuencias a largo plazo tienen una mayor probabilidad de ser considerados significativos.

c) Respecto de la relación causal y previsibilidad del daño (nº 94-96), para que un sistema sea prohibido no se exige que el daño ya se haya producido; basta con que sea razonablemente probable en función del diseño del sistema y sus impactos previsibles. Para establecer la relación causal, se considera la capacidad del sistema de IA para generar una distorsión del comportamiento que derive en daño significativo. Asimismo, se evalúa si el proveedor o implementador del

sistema podría haber previsto y mitigado el daño mediante medidas y salvaguardas preventivas suficientes.

III. CUÁNDO SE CONSIDERA UN "SOCIAL SCORING" PROHIBIDO (ARTÍCULO 5, 1 C)

El artículo 5. 1º c) prohíbe la comercialización, puesta en servicio o uso de sistemas de IA que clasifiquen o evalúen a individuos o colectivos en función de su comportamiento social o características personales, inferidas o predichas, cuando dicha puntuación genere un trato perjudicial, injustificado o desproporcionado en contextos distintos a aquellos en los que se originaron los datos. Esta prohibición del *social scoring* o crédito social es una de las más discutidas[14] y la Comisión aporta criterios interpretativos adicionales (nº 146-183) para considerar o no si el sistema IA está o no prohibido. Así, para que un sistema de IA quede comprendido dentro de la prohibición del artículo 5.1.c, deben concurrir simultáneamente tres condiciones acumulativas: la comercialización, puesta en servicio o uso del sistema de IA, que el sistema realice una evaluación o clasificación basada en el comportamiento social o en características personales o de personalidad y que se produzca un trato perjudicial o desfavorable en contextos sociales no relacionados y/o trato desproporcionado respecto del comportamiento social.

Por cuanto la *comercialización, puesta en servicio o uso* del sistema de IA, se requiere que el sistema sea ofrecido en el mercado, desplegado o utilizado activamente para la evaluación o clasificación de individuos. Así, la prohibición afecta tanto a proveedores como a usuarios que implementen o utilicen tales sistemas en sus operaciones.

Respecto de la *"evaluación o elaboración de perfiles"*, el sistema debe realizar una "puntuación social", esto es, procesar datos sobre el comportamiento social o las características personales de las personas físicas o grupos, ya sea mediante información directa, inferida o predicha (nº 151 y ss.). Cabe matizar que una evaluación implica "apreciación o juicio sobre una persona o grupo de personas" y que va más allá que una "simple clasificación" de personas o grupos (nº 153). Evaluación incluye también el concepto de "elaboración de perfiles" propio de la protección de datos (nº 153).

El tercer elemento de la prohibición es que esté *basada en el comportamiento social o en características personales o de personalidad*. En este sentido, el concepto de

[14] Sobre el tema puede seguirse ; "La prohibición de sistemas de inteligencia artificial que evalúan y clasifican a las personas a partir de datos que no guardan relación con el contexto donde se generaron y que provocan discriminaciones", por Miguel Ángel Presno Linera, en Cotino Hueso, L. y Simó Castellanos, P. (coords.), *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea*, Aranzadi La Ley, Cízur Menor, 2024.

"comportamiento social" se interpreta de manera amplia e incluye (nº 157): acciones y hábitos en la vida cotidiana (por ejemplo, cumplimiento de normas de tráfico, participación en eventos comunitarios, actividades en redes sociales). Interacciones con entidades públicas y privadas (por ejemplo, historial de pagos, relaciones con instituciones estatales, comportamientos financieros) y, también, la conducta en el ámbito digital (por ejemplo, publicaciones en redes sociales, patrones de navegación en internet, análisis de actividad en línea).

En paralelo, las características personales (nº 158) se detalla que abarcan aspectos como la edad, género, orientación sexual, origen étnico, estado de salud, situación económica y antecedentes profesionales, entre otros. Se considera especialmente problemático cuando estos datos son combinados en sistemas de clasificación intersectoriales, lo que puede reforzar sesgos y generar discriminación estructural. Además, estas características personales, bien pueden ser datos introducidos o inferidos por el sistema IA e incluso "características predichas".

Asímismo, para que se configure una práctica de social scoring prohibido se requiere que se de una alternativa. Bien, que *la puntuación social genere un trato perjudicial o desfavorable en contextos sociales no relacionados* (nº 160 y ss.). Así, se señalan varios ejemplos (nº 166). [15] Es de interés indicar que, siguiendo el nº 175, los sistemas de puntuación en el ámbito financiero (como los historiales de crédito) no están prohibidos si su uso se restringe a la determinación de riesgos de impago en un préstamo, esto es, que su uso no se extrapole a conclusiones más amplias sobre su comportamiento social en otros contextos o áreas sociales, como el acceso a empleo o servicios básicos, por ejemplo, que una puntuación crediticia determine el acceso a educación o vivienda.

Si no se cumple el requisito anterior, también se considerará social scoring prohibido si la puntuación social genera un trato desproporcionado en relación con el comportamiento social. La prohibición también puede concurrir (nº 167) poque el sistema genere alternativamente un trato desproporcionado respecto de la gravedad del comportamiento social, lo que significa que la evaluación conduce a una sanción excesiva o discriminatoria en comparación con la conducta evaluada (por ejemplo, la negación de un servicio esencial por una infracción menor como no devolver un libro a tiempo a la biblioteca). Se señalan variados ejemplos.

[15] Así, se menciona la selección de declaraciones fiscales para inspección basada en variables ajenas a criterios fiscales, como hábitos sociales o conexiones a internet. La detección de fraude en prestaciones sociales mediante factores no vinculados a la determinación de derechos legales - como la nacionalidad del cónyuge o el uso de plataformas sociales. También, la puntuación de desempleados para ayudas estatales basada en datos sensibles o irrelevantes como información sanitaria, estado civil o aspectos personales no relacionados con la empleabilidad en la determinación de acceso a beneficios estatales.

Se apuntan además otros criterios de interpretación y criterios de aplicación respecto de la prohibición del crédito social. Así:

- la prohibición recae tanto en la procedencia pública o privada del sistema IA de crédito social (nº 170).

- La prohibición no aplica o excluye (nº 173-178) a personas jurídicas salvo que su evaluación se base en características personales o comportamiento social.

Tampoco se considera aplicable respecto de prácticas legítimas y específicas, como la calificación crediticia, la prevención del fraude financiero o la evaluación proporcional del comportamiento en plataformas digitales, siempre que cumplan la normativa de la UE y no generen trato desproporcionado o injustificado. Además, la prohibición no afecta a sistemas de IA empleados en seguridad pública, siempre que su uso esté justificado y se ajuste al marco legal aplicable.

Finalmente (nº 179-183) se recuerda que la prohibición del crédito social interactúa con otras normas y, por ejemplo, puede infringir la Directiva 2005/29/CE sobre prácticas comerciales desleales si distorsionan el comportamiento económico de los consumidores, así como vulnerar el RGPD cuando el tratamiento de datos carezca de base legal o implique decisiones automatizadas sin salvaguardias. Asimismo, la Directiva (UE) 2023/2225 sobre crédito al consumo prohíbe evaluar la solvencia con datos sensibles o extraídos de redes sociales. Finalmente, los sistemas de IA empleados en la lucha contra el blanqueo de capitales y financiación del terrorismo deben cumplir la legislación sectorial aplicable.

IV. CUÁNDO SE DA LA PROHIBICIÓN SOBRE PREDICCIÓN DELICTIVA SOBRE PERFILES Y PERSONALIDAD DEL ARTÍCULO 5, 1, D)

El artículo 5 1 d), RIA prohíbe que los sistemas de IA evalúen o predigan el riesgo de que una persona física cometa un delito basándose únicamente en la elaboración de perfiles o en la evaluación de rasgos y características de la personalidad.[16] Pese a establecerse como una prohibición absoluta, porque expresamente no se establecen excepciones, es muy importante su delimitación. No en vano, si no se dan los

[16] Sobre el tema puede seguirse "La regulación de los sistemas policiales predictivos en el Reglamento Inteligencia Artificial", por Fernando Miró Llinares y Mario Santisteban Galarza, en Cotino Hueso, L. y Simó Castellanos, P. (coords.), *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea*, Aranzadi La Ley, Cízur Menor, 2024.

presupuestos del no es no, muy posiblemente el sistema afín no prohibido caerá dentro de diversos supuestos de alto riesgo.[17]

Los criterios fijados por la Comisión (nº 184 y ss.) recorren los motivos de la prohibición y la necesidad de que las personas sean juzgadas por su comportamiento real y no por predicciones derivadas de perfiles basados en patrones históricos de criminalidad o en atributos personales. Pero la prohibición no implica una negativa absoluta a las técnicas predictivas penales, sólo cuando se basan exclusivamente en la elaboración de perfiles.

La propuesta de Comunicación detalla los elementos constitutivos de la prohibición. Para que se active la prohibición del artículo 5.1.d, deben concurrir simultáneamente dos elementos: la existencia de una evaluación del riesgo o predicción de criminalidad (nº 189) y que esta se base únicamente en el perfil de una persona física. Como se ha dicho se sigue el concepto de elaboración de perfiles del RGPD y cabe tener en cuenta que también puede aplicarse a los "perfiles de grupo" (p. 196) por ejemplo, terroristas, gángsters, etc. a partir de datos históricos. Asimismo, puede basarse en la evaluación de sus rasgos y características de personalidad. Es determinante que sólo están prohibidas cuando se basan "únicamente" en perfiles o características de la personalidad. Como aclara el considerando 42 RIA, puede ya haber una sospecha razonable respecto de la persona física de que se trate, esto es, que haya habido una evaluación humana, que normalmente se basará en hechos pertinentes objetivos y verificables. Así pues, el RIA deja abierta la posibilidad de que se tengan en cuenta otros elementos en la evaluación del riesgo que hay que analizar caso por caso (nº 199-202) y se mencionan diversos ejemplos que sí que estarían prohibidos.[18]

---

[17] Así, si no está prohibida, posiblemente pueda considerarse que es un sistema de IA para evaluar el riesgo de reincidencia delictiva (anexo III.6.d), siempre que no se limiten únicamente a perfiles o rasgos de personalidad, sino que integren antecedentes delictivos u otros factores objetivos relacionados con la reincidencia. En su caso podría ser un sistema de IA para evaluar la fiabilidad de las pruebas en investigaciones penales (anexo III.6.c). También podría ser considerado un sistema IA de elaboración de perfiles en investigaciones penales (anexo III.6.e), no prohibido cuando el perfilado es parte de una investigación en curso basada en hechos y datos objetivos, no utilizado únicamente para predecir delitos futuros. También el sistema podría encuadrarse en alto riesgo por evaluar el riesgo de que una persona cometa delitos en el contexto de migración y control fronterizo (anexo III.7.b), siempre que no se base en el perfilado personal. Igualmente podría caer bajo los polígrafos y herramientas similares para detección de engaños (anexo III.6.b y 7.a), siempre que no derive en una clasificación de personas basada solo en perfilado. Quizá, también como sistema IA en la administración de justicia (anexo III.8.a), siempre que no se utilicen para predecir criminalidad futura, sin fundamento en hechos objetivos.

[18] Así, si una autoridad policial utiliza un sistema de IA para predecir el comportamiento delictivo en delitos como el terrorismo basándose únicamente características personales como la edad, la nacionalidad, la dirección, el tipo de coche y el estado civil de las personas. También se considera prohibido si se utiliza una herramienta predictiva de IA para revisar las declaraciones de la renta de todos los contribuyentes con el fin de predecir posibles delitos fiscales e identificar los casos que requieren una investigación más profunda y esto se hace únicamente sobre rasgos de personalidad, como la doble nacionalidad, el lugar de nacimiento, el número de hijos, y variables

Son de especial interés los criterios de los *sistemas IA que quedarían fuera de esta prohibición,* pese a que se trata de una prohibición formulada con carácter absoluto. Así, los sistemas IA no quedan prohibidos cuando se utilizan para apoyar la evaluación humana de la implicación de una persona en una actividad delictiva (nº 203-206), siempre que dicha evaluación se base en hechos objetivos y verificables directamente relacionados con un delito. Así se acepta que la IA sea herramienta de análisis auxiliar en la toma de decisiones y son de alto riesgo. La función de la IA en estos contextos debe ser de apoyo a la evaluación humana, sin sustituir el criterio jurídico ni la valoración probatoria basada en hechos verificables. En estos casos, los sistemas de IA se consideran de alto riesgo (anexo III, punto 6 d) RIA).

Aunque las autoridades policiales son los principales usuarios de estos sistemas, la prohibición también puede aplicarse a *agentes privados*, ello sucede cuando la ley les otorga funciones públicas de seguridad (nº 207 y ss). Sin embargo, quedarían fuera de la prohibición las actividades empresariales que incluyen la detección de fraudes financieros o la evaluación de riesgos de seguridad en transacciones comerciales, siempre que su finalidad no sea predecir la criminalidad individual, sino proteger intereses legítimos de manera proporcional y conforme a la normativa aplicable.

También quedan fuera de la prohibición los sistemas de IA destinados a la predicción geoespacial de la delincuencia, siempre que no impliquen la evaluación individualizada del riesgo de una persona específica. Es decir, no está prohibido por este precepto el uso la IA para identificar patrones delictivos en determinadas zonas y guiar estrategias de prevención policial. Pero sí que estaría prohibido si el sistema combina esta predicción con la elaboración de perfiles personales. Tampoco estaría prohibida la evaluación de empresas, asociaciones u ONG en relación con posibles delitos, salvo que elaboren perfiles de individuos concretos para predecir su riesgo criminal. La prohibición no impide el uso de IA para la evaluación del riesgo de reincidencia en la ejecución de penas, siempre que las predicciones se basen en elementos objetivos y verificables, como el historial delictivo comprobado y el comportamiento penitenciario.

Finalmente, cabe tener en cuenta otras normas como las de protección de datos (RGPD y Directiva 2016/343 sobre la presunción de inocencia establece que cualquier tratamiento de datos personales en el contexto penal debe evitar la estigmatización de los sospechosos y garantizar que las decisiones no se fundamenten en inferencias probabilísticas sin sustento factual.

---

opacas, especialmente información inferida que es predictiva y, por tanto, no objetiva y difícil de verificar. Otro ejemplo de sistema prohibido sería una herramienta policial de evaluación de riesgos de que niños y adolescentes se vean implicados en futuros delitos violentos y contra la propiedad y se evalúa a los niños en función de, por ejemplo, por estar vinculados a otra persona con una evaluación de alto riesgo, como un hermano o un amigo o el nivel de riesgo de los padres.

V. LOS REQUISITOS DE LA PROHIBICIÓN DE RASTREO NO SELECTIVO DE IMÁGENES FACIALES, ARTÍCULO 5, 1, E)

El artículo 5.1.e) RIA (y considerando 43) impone una prohibición teóricamente absoluta o categórica a la comercialización, puesta en servicio y uso de sistemas de inteligencia artificial destinados a la creación o ampliación de bases de datos de reconocimiento facial mediante la extracción no selectiva de imágenes faciales de Internet o grabaciones de vídeovigilancia.[19] Los criterios (nº 222 y ss.) repasan los *requisitos de la prohibición*. Así, se deja claro que para que la prohibición sea aplicable, deben cumplirse cuatro condiciones acumulativas: (i) la actividad debe implicar la comercialización, puesta en servicio o uso de un sistema de IA; (ii) el propósito debe ser la creación o expansión de bases de datos de reconocimiento facial; (iii) el método de recopilación debe consistir en técnicas de scraping no selectivo; y (iv) las fuentes de las imágenes deben ser Internet o grabaciones de vídeovigilancia.

En cuanto a la técnica empleada (iii), se prohíbe el *scraping* no selectivo, que implica la extracción masiva de datos de sitios web sin un enfoque individualizado. En consecuencia, la prohibición no se aplica a la extracción selectiva y dirigida a personas concretas (nº 225), por ejemplo, en investigaciones policiales donde se buscan imágenes de sospechosos o víctimas específicas. Tampoco estaría prohibido, por ser selectivo, la recopilación de imágenes de forma dirigida para identificar a víctimas de trata de personas. Pero si un sistema va recopilando imágenes por segmentos para construir progresivamente una base de datos global, se considerará una maniobra para eludir la prohibición.

Otro de los elementos de la prohibición es que el propósito debe ser la creación o expansión de *"bases de datos de reconocimiento facial"* (ii). Cabe delimitar que se trata de bases que permiten cotejar rostros humanos con imágenes almacenadas con el fin de identificar a individuos (nº 226). No es necesario para la Comisión que la única función de la base de datos sea el reconocimiento facial, basta con que tenga esa capacidad. De manera similar, el *scraping* de imágenes faciales se entiende como la extracción automatizada de datos mediante rastreadores web, *bots* u otras herramientas informáticas.

Cabe delimitar cuáles son las *fuentes de las imágenes prohibidas y no prohibidas* (iv). En principio están prohibidas fuentes de Internet o grabaciones de vídeovigilancia. Se concreta al respecto (nº 232 y ss.) que está prohibida la inclusión de fotografías en redes

[19] Puede seguirse al respecto, "El resto de sistemas de inteligencia artificial prohibidos o inaceptables en el Reglamento", por Pere Simó Castellanos en Cotino Hueso, L. y Simó Castellanos, P. (coords.), *Tratado sobre el Reglamento de Inteligencia Artificial…cit.*

sociales, pues no se cuenta con el consentimiento para bases de datos de reconocimiento facial. Tampoco las imágenes de vídeovigilancia captadas en espacios públicos mediante *scraping* no selectivo. No obstante, la búsqueda inversa de imágenes, en la que se introduce una fotografía para encontrar coincidencias en la web, no se considera "*scraping* no selectivo" en sentido estricto y, por lo tanto, su tratamiento jurídico requiere un análisis más detallado. Se excluye de la prohibición (nº 234 y ss.) la extracción masiva de otros datos biométricos distintos del rostro, como muestras de voz. Tampoco se prohíbe por el artículo 5 el almacenamiento de bases de datos de imágenes faciales cuando no se emplean para el reconocimiento de personas, como en el entrenamiento de modelos de IA para la generación de rostros sintéticos. Y en todo caso, además de otra normativa concurrente, habría de cumplir los requisitos de transparencia establecidos en el artículo 50 RIA. Tampoco entran en la prohibición *per se* las bases de datos preexistentes creadas antes de la entrada en vigor de la prohibición, sin perjuicio de su cumplimiento de otras normas como el RGPD.

VI. LA PROHIBICIÓN DE SISTEMAS DE INFERENCIA DE EMOCIONES EN EL TRABAJO Y LA EDUCACIÓN(ARTÍCULO 5.1.F)

El artículo 5.1.f) y el considerando 44 RIA prohíben o establecen requisitos y límites al uso de sistemas de inteligencia artificial destinados a inferir emociones en los ámbitos del trabajo y la educación, con la única excepción de aquellos implementados por razones médicas o de seguridad. [20] Sobre estas técnicas se recuerdan las serias dudas científicas y sus potenciales peligros (nº 241). De nuevo, son de todo interés los Criterios (nº 239 y ss.) que fija la Comisión al respecto.

Se delimita *el concepto de inferencia de emociones* (nº 244 y ss.). Así, engloba tanto la identificación previa de una emoción como su deducción a partir de datos biométricos. Aunque el texto del RIA menciona solo la inferencia, el considerando 44 sugiere una interpretación amplia que incluya la identificación como parte del proceso. Así, quedan prohibidos los sistemas que no solo detectan expresiones faciales o cambios en la voz, sino también aquellos que interpretan estados emocionales o intenciones a partir de estos datos. Se advierte que no puede eludirse la prohibición acudiendo a términos como "actitudes". Esto es, también estarán prohibidos sistemas IA que determinen, por ejemplo, si una persona muestra enfado en base a sus datos biométricos.

La prohibición es sólo para el lugar de trabajo y las instituciones educativas. Respecto del trabajo (nº 254), no hay distinción entre sectores, tipos de contrato o modalidad

[20] Puede seguirse al respecto, "El resto de sistemas de inteligencia artificial prohibidos o inaceptables en el Reglamento", por Pere Simó Castellanos en Cotino Hueso, L. y Simó Castellanos, P. (coords.), *Tratado sobre el Reglamento de Inteligencia Artificial…cit.*

laboral. Se incluye tanto el trabajo presencial como el remoto y se extiende a procesos de selección y contratación, en coherencia con otras disposiciones del RIA sobre gestión de empleados. En cuanto a las instituciones educativas (nº 255), se incluyen todos los niveles y modalidades de enseñanza, tanto públicas como privadas, incluyendo la formación profesional y continua.

Los criterios detallan las *excepciones por razones médicas y de seguridad.* Es importante el criterio de que la referencia a fines médicos debe restringirse al uso de productos sanitarios con marcado CE. En consecuencia, no valdrían como excepción y estarían prohibidas en el trabajo o la educación las herramientas de monitoreo del bienestar general, como sistemas que midan el estrés o el agotamiento de los empleados. En cambio, sí que entrarían en la excepción y no estarían prohibidos los sistemas IA de reconocimiento de emociones con sistemas clínicamente validados con marcado CE para detectar trastornos específicos de salud mental (nº 257). En el caso de la excepción por seguridad, se concreta que la excepción se limita exclusivamente a la protección de la vida y la salud, excluyendo cualquier otro propósito. Por ejemplo, sí que estarían prohibidos en trabajo y educación estos sistemas de reconocimiento con la finalidad de la prevención de fraudes o la protección de bienes (nº 258). En todo caso, se advierte, las excepciones requieren la concurrencia de los principios de necesidad y proporcionalidad, con evaluaciones previas que justifiquen la falta de alternativas menos impactantes en derechos e intereses. Sin perjuicio de las posibles excepciones posibles a la prohibición, se recuerda que el art. 2.11 RIA permite a los estados regulación más favorable a los trabajadores, lo que podría incidir en esta materia.

Además de las excepciones, cabe tener en cuenta *la posibles exclusiones a la prohibición por quedar fuera del ámbito de aplicación* (nº 264 y ss.). Así, no estarían prohibidos los sistemas de reconocimiento emocional que no utilizan datos biométricos, como aquellos basados en encuestas o análisis de texto. También se excluyen de la prohibición del artículo 5.1.f) los sistemas de IA diseñados para inferir estados físicos como el dolor o la fatiga, porque el RIA no los considera sistemas de reconocimiento emocional. Asimismo, tampoco están prohibidos por este precepto los sistemas de control de multitudes que analicen emociones colectivas sin individualización, como los que evalúan el nivel de estrés en eventos masivos, no entran en la prohibición, salvo que infieran emociones de personas identificables. Tampoco están dentro de la prohibición los robots asistenciales o los médicos que utilizan sistemas de reconocimiento de emociones durante un examen en su lugar de trabajo y los monitores de voz que analizan las llamadas de emergencia (nº 269).

VII. LA PROHIBICIÓN DE SISTEMAS DE CATEGORIZACIÓN BIOMÉTRICA PARA INFERIR CARACTERÍSTICAS ESPECIALMENTE SENSIBLES DE LAS PERSONAS (ARTÍCULO 5.1.G)

El artículo 5.1º g) RIA [21] prohíbe la comercialización, puesta en servicio o uso de sistemas de categorización biométrica que clasifiquen individualmente a las personas en función de datos biométricos para inferir su raza, opiniones políticas, afiliación sindical, convicciones religiosas o filosóficas, vida sexual u orientación sexual. No obstante, se excluyen de esta prohibición el etiquetado o filtrado de conjuntos de datos biométricos obtenidos lícitamente y su uso en el ámbito de la garantía del cumplimiento del Derecho. También la Comisión aporta importantes criterios y aspectos interpretativos y delimitadores sobre el alcance de esta prohibición (nº 271 y ss.), que se formula de manera categórica y absoluta, esto es, en teoría sin excepciones. En razón de este precepto, la prohibición gira alrededor de la definición de "sistema de categorización biométrica", que "clasifiquen *individualmente* a las personas físicas" y que se utilice para inferir datos que se pueden calificar de *sensibles*.

Así, en primer término hay que detenerse en *la definición de "sistema de categorización biométrica".* Se detalla que la categorización no consiste en identificar o verificar la identidad de la persona, sino en asignarla a una categoría o grupo basándose en sus rasgos biométricos. No entrarían pues en la definición, por lo que no estarían prohibidas por este artículo ciertas funcionalidades "auxiliares" y "estrictamente necesarias" por razones técnicas para otro servicio comercial, siempre que no se trate de un uso principal de categorización biométrica (nº 270). Se afirma como ejemplo permitido, los filtros que categorizan los rasgos faciales o corporales utilizados en los mercados en línea para permitir a un consumidor previsualizar un producto en sí mismo podrían constituir una característica accesoria de este tipo, ya que sólo pueden utilizarse en relación con el servicio principal que consiste en vender un producto (nº 280). También estarían permitidos los filtros integrados en los servicios de redes sociales en línea que categorizan los rasgos faciales o corporales para permitir a los usuarios añadir o modificar imágenes o vídeos también podrían considerarse una característica accesoria, ya que dicho filtro no puede utilizarse sin el servicio principal de los servicios de redes sociales consistente en compartir contenidos en línea.

En segundo término, *se delimita la "*categorización biométrica *individual*". Para que la prohibición se aplique, el sistema debe catalogar a cada persona física de manera *individual* en función de sus datos biométricos (no a una colectividad sin distinción). Por ejemplo, como categorización individual sería un sistema de IA que realizan

[21] Puede seguirse al respecto, "El resto de sistemas de inteligencia artificial… cit.

'Estimaciones de Atributos" (nº 282) por ejemplo 'edad, sexo, etnia", basándose en rasgos corporales, como la cara, la altura o el color de la piel, ojos y pelo (o una combinación de los mismos). También lo serían sistemas que personas y las señalen en función de un rasgo específico (por ejemplo, una cicatriz bajo el ojo derecho) o porque tienen un tatuaje en la mano derecha.

Un tercer criterio básico de esta prohibición es delimitar que *la prohibición sólo alcanza sistemas que tengan el objetivo de inferir características que podemos calificar de sensibles*. En concreto, el artículo menciona la raza, opiniones políticas, afiliación sindical, creencias religiosas o filosóficas, vida sexual u orientación sexual. La prohibición sólo recae sobre este tipo y finalidades de inferencias y no otras que en su caso puedan ser sensibles, pero no en este artículo. Así, expresamente se recuerda que no estaría prohibido por este artículo el etiquetado o filtrado de datos biométricos legalmente obtenidos, incluso con fines policiales, si su finalidad es, por ejemplo, corregir sesgos en los conjuntos de entrenamiento o asegurar representatividad de los datos. Tampoco la prohibición comprende el "etiquetado o filtrado de conjuntos de datos biométricos" cuando tiene el objetivo de es mejorar la calidad de los datos o evitar discriminaciones. De hecho, puede incluso ser obligatorio para cumplir requisitos de igualdad y evitar sesgos algoritmos.

Finalmente, se recuerda la relación de esta prohibición con la normativa de no discriminación y la de protección de datos (RGPD y Directiva). En concreto se recuerda la previsión expresa del artículo 11.3 de la Directiva (UE) 2016/680 de aplicación de la ley que prohíbe la elaboración de perfiles que dé lugar a discriminación.

VIII. EL ALCANCE DE LA PROHIBICIÓN DEL USO DE SISTEMAS DE IDENTIFICACIÓN BIOMÉTRICA A DISTANCIA EN TIEMPO REAL EN ESPACIOS PÚBLICOS CON FINES POLICIALES (ART. 5.1. H)

El artículo 5.1º h) RIA [22] prohíbe -en principio- "el uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público con fines de garantía del cumplimiento del Derecho". Esta prohibición ha sido desde el inicio del proceso de regulación del RIA uno de los temas más sensibles políticamente y sometidos a negociación. Ello se acaba reflejando con claridad en la muy amplia y compleja regulación tanto de lo que sería la prohibición como de sus excepciones y de las necesidades y requisitos para la regulación de las mismas. De ahí que los criterios

[22] Me permito destacar el excelente estudio El reconocimiento biométrico en el Reglamento de inteligencia artificial: exenciones, prohibiciones y especialidades de alto riesgo, por Leire Escajedo en Cotino Hueso, L. y Simó Castellanos, P. (coords.), *Tratado sobre el Reglamento de Inteligencia Artificial…cit.*

interpretativos de la Comisión al respecto sean tan importantes (nº 289-325). Sobre esta regulación y su superposición compleja con la normativa de protección de datos me remito a anteriores estudios.[23]

En primer término procede delimitar cuál es el alcance inicial de la prohibición. Y ésta se articula en diversos elementos acumulativos. Así, *se prohíbe sólo el uso, no la comercialización*, que se regula como "sistema de alto riesgo" en el anexo III de la ley. Se prohíbe la "identificación biométrica *a distancia"*. La definición comprende cualquier tecnología capaz de identificar a una persona a partir de datos biométricos como la cara, la voz, la postura o el ritmo cardíaco, siempre que se utilice con fines de identificación individual y comparación con bases de datos. Si no hay una base de datos contra la que contrastar, no puede considerarse identificación biométrica.

La identificación biométrica que está prohibida es *"a distancia".* La noción de identificación biométrica remota se basa en la captura y procesamiento sin participación activa de los individuos. Ello a diferencia de la verificación biométrica, que requiere consentimiento y un proceso activo del sujeto. La distancia implica que la persona no interactúa de manera consciente o activa con el sistema. Se subraya por la Comisión que no basta la mera notificación de la presencia de cámaras; se necesita la participación activa ante el dispositivo para quedar fuera de la definición de "a distancia" (nº 304 y ss.). Esto excluye sistemas como los de desbloqueo de dispositivos personales o control de acceso en edificios.

Asímismo, la identificación prohibida ha de ser *"en tiempo real".* Ello implica la captura, comparación e identificación sin una demora significativa que permite una intervención instantánea de las autoridades (nº 310). Puede admitirse un breve lapso, pero no uno tal que la persona abandone ya el lugar de grabación. Por el contrario serían de alto riesgo el uso retrospectivo (análisis a posteriori).

Y para que esté prohibida, también debe ser *en "espacios de acceso público"* (nº 313 y ss.). Ello es bien relevante y alcanza cualquier espacio (público o privado) accesible a un número indeterminado de personas, con independencia de restricciones como un billete o un requisito de edad. Todo ello independientemente de su titularidad pública o privada. No entran en la prohibición los espacios en línea (no físicos), las zonas restringidas de fábricas u oficinas administrativas, los controles fronterizos y otros lugares donde el acceso está predeterminado para un grupo limitado de personas.

[23] "Reconocimiento facial automatizado y sistemas de identificación biométrica bajo la regulación superpuesta de inteligencia artificial y protección de datos", en Balaguer, Francisco y Cotino, Lorenzo (2023): *Derecho público de la inteligencia artificial*, F. Jiménez Abad-Marcial Pons, acceso , más breve, "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 2022, pp. 68-79. acceso

Por último, la prohibición inicialmente concebida requiere que la identificación biométrica sea con *fines policiales y en el ámbito de aplicación de la ley*. Se trata de finalidades de prevención, investigación, detección o enjuiciamiento de delitos, además de la ejecución de sanciones. También estarán en la prohibición incluso cuando tales actividades las realicen entidades privadas que actúen en nombre de la autoridad.

IX. LAS EXCEPCIONES A LA PROHIBICIÓN DEL ARTÍCULO 5.1. H)

Una vez conocido el alcance inicial de la prohibición, es especialmente relevante todo el sistema de excepciones y requisitos a la misma. Más que un no es no, claramente esta prohibición se configura como un "no pero sí". De ahí que sean muy relevantes los criterios de la Comisión (nº 326 y ss.) que detallan elementos interpretativos sobre las excepciones. Hay que insistir que estas excepciones del artículo 5 a la prohibición no suponen una habilitación y una permisión, sino que como luego se expone es necesaria una regulación legal específica estatal o de la UE para que sí pudieran darse sistemas de identificación biométrica a distancia en tiempo real en espacios públicos con fines policiales. Y estas leyes específicas que puedan darse, además, habrán de cumplir los muchos requerimientos de artículos 5.2 a 7 RIA que luego se comentan.

Pues bien, respecto de toda posible excepción a la prohibición, el punto de partida de cualquier regulación habilitadora es la "estricta necesidad" de lograr "un interés público sustancial" que "compense los riesgos" que entrañan para los derechos fundamentales (considerando 33, nº 328). Otra premisa a las excepciones es que sólo se excepciona la identificación biométrica remota en tiempo real para casos muy concretos: búsqueda de víctimas o personas desaparecidas; prevención de amenazas graves e inminentes (incluyendo atentados terroristas); y localización o identificación de sospechosos de determinados delitos graves (listados en anexo II y con pena máxima de al menos cuatro años). Así pues, cualquier otro fin policial no contemplado no puede ser la base de una excepción a la prohibición y, por tanto, la prohibición sería plena. Y cuando el objetivo coincida con los enumerados, el despliegue del sistema debe justificarse por ser la única vía para conseguir dicho fin, valorando la gravedad y urgencia del supuesto (búsqueda de víctimas de delitos graves, amenaza real e inminente de atentado, etc.).

Procede delimitar los *delitos que habilitan la excepción*. Respecto de *la búsqueda selectiva de personas desaparecidas* (nº 329 y ss.) se señala que la policía solo interviene cuando la desaparición encaja en un supuesto legal concreto (por ejemplo, si existen indicios de que la persona está en peligro). En otros casos (p. ej., desaparición voluntaria de un adulto), podría no activarse una búsqueda policial que permita el uso de estos sistemas. Tampoco se incluirían en la excepción las búsquedas administrativas, por ejemplo, en asistencia social. Asímismo, se señala que sólo se permitirían

búsquedas selectivas, esto es, el sistema debe orientarse únicamente a la víctima específica, sin generar un rastreo indiscriminado de personas.

Por cuanto a la interpretación de la excepción por *amenazas inminentes o previsibles contra la vida* (nº 337 y ss.). Se deja a la evaluación de cada Estado miembro el análisis de la "amenaza específica, sustancial e inminente" para la vida o la seguridad física, así como de la "amenaza real y actual o de atentado terrorista". En todo caso, la amenaza debe ser concreta (no hipotética) y debe requerir reacción inmediata. Respecto de las amenazas por *terrorismo* (nº 345 y ss.), la amenaza debe referirse concretamente a un atentado terrorista, no a la amenaza terrorista en abstracto. Se recuerda además que el RIA exige que la amenaza sea "real y previsible", un criterio tomado del TJUE que en su jurisprudencia sobre medidas de seguridad nacional exige circunstancias concretas y verificables, evitando justificaciones basadas en simples sospechas o modelos predictivos no fundamentados (nº 347).

Por cuanto a la excepción basada en la *localización e identificación de sospechosos o condenados* por delitos graves (nº 349 y ss.), hay que tener en cuenta el listado de delitos del anexo II y que la pena máxima aplicable en el Estado miembro sea de al menos cuatro años de privación de libertad. Se insiste en que un sospechoso es alguien con indicios fundados de haber cometido un delito. Esta excepción puede relacionarse con otras, por ejemplo, si un delincuente es también una persona desaparecida (por ejemplo, un menor víctima de trata).

X. REQUISITOS, SALVAGUARDIAS Y CONDICIONES PARA LAS EXCEPCIONES DEL ARTÍCULO 5.1. H)

Como se ha visto, pueden darse excepciones a la prohibición para utilizar sistemas de identificación biométrica remota en tiempo real en espacios públicos con fines policiales. El artículo 5. 2º y siguientes detalla toda una serie de requisitos y condiciones para estos supuestos y se aportan criterios respecto de los mismos (nº 357 y ss.). Dado que cada Estado puede activar o no las excepciones, se recuerda que podrá en su caso imponer restricciones adicionales más estrictas de lo exigido por el RIA (quien puede lo más, puede lo menos). Asimismo cada Estado podrá definir las normas de procedimiento para la autorización, supervisión y supervisión del uso de estos sistemas. Sobre la materia me permito recordar estudios que he realizado sobre las exigencias de "Una regulación legal y de calidad para los análisis automatizados de datos o con inteligencia artificial. Los altos estándares del Tribunal Constitucional alemán y otros tribunales, que no se cumplen ni de lejos en España"[24]

[24] En *Revista General de Derecho Administrativo*, *RGDA Iustel*, RGDA Iustel, nº 63, 2023. acceso

El sistema IA biométrico solo puede emplearse con la *finalidad limitada y "confirmación de identidad"* de una persona previamente identificada y no para "identificar", esto es, no cabe la identificación masiva o vigilancia generalizada. Se pretenden evitar usos arbitrarios, sino que el uso ha de ser exclusivo a partir de una base razonable para relacionar a la persona en uno de los supuestos excepcionales permitidos (p. ej., víctima desaparecida, sospechoso de un delito muy grave, etc.).

La regulación establece requisitos y condiciones y la Comisión fija criterios sobre los mismos. Así, sobre finalidad limitada de los sistemas y la "confirmación de identidad", sobre la exigencia de evaluación de proporcionalidad y daño potencial ex ante del riesgo a través de un FRIAS, sobre los temporales, geográficos y personales. También se concretan diversas medidas de transparencia y supervisión, el sistema debe registrarse en la base de datos de la UE. De especial interés es la garantía de la autorización previa por parte de una autoridad independiente. Asimismo, se establece la prohibición de tomar decisiones automatizadas con efectos jurídicos adversos sin intervención humana, evitando que el sistema determine unilateralmente consecuencias legales para las personas. Finalmente, se exige la notificación de la activación de las excepciones a las autoridades de vigilancia y protección de datos.

Así, *se requiere una evaluación de proporcionalidad y daño potencial ex ante del riesgo (nº 360*), sobre la base de la gravedad del perjuicio en caso de no emplear el sistema, del número y perfil de personas afectadas y la probabilidad de la amenaza y su impacto en derechos fundamentales. Esta evaluación debe insertarse dentro de la Evaluación de Impacto sobre los Derechos Fundamentales (FRIA) del artículo 27 RIA que es obligatoria A (nº 370 y ss.).[25] Debe hacerse antes de desplegar el sistema por la autoridad policial y debe abarcar riesgos para la libertad de reunión, la no discriminación y otros derechos, además de la privacidad y de datos personales y no sustituye a la también obligatoria evaluación de impacto de protección de datos.

Por cuanto a los *límites temporales, geográficos y personales* es preciso delimitar la la duración de modo justificado), el ámbito geográfico específico ("esta delimitación no debería abarcar una ciudad o un país enteros, sino que debería ser más específica", 366) y las personas objeto del sistema (p. ej., integrantes de un grupo terrorista identificado). También el El despliegue del sistema IA debe inscribirse en la *base de datos europea* (art. 49 RIA) antes de su uso. De manera excepcional y por urgencia

[25] Sobre el tema, me permito destacar los estudios "La evaluación de impacto de derechos fundamentales por quienes despliegan sistemas de inteligencia artificial en el Reglamento", por Eduard Chaveli Donet así como "Los sistemas de gestión de riesgos como obligación específica para los sistemas de inteligencia artificial de alto riesgo en el artículo 9 del Reglamento", entre otros estudios de Pere Simón Castellano, ambos en Cotino Hueso, L. y Simó Castellanos, P. (coords.), *Tratado sobre el Reglamento de Inteligencia Artificial…cit.*

debidamente justificada, puede iniciarse su utilización sin el registro, pero este debe completarse “sin demora indebida” (nº 378).

Respecto de la *autorización previa de una autoridad independiente,* es imprescindible respecto de cada uso individual y ha de hacerse por un órgano judicial o por una autoridad administrativa independiente (obviamente distinta de la propia policía), que evalúe la estricta necesidad y proporcionalidad de la medida. Es de interés precisar que no se requiere la autorización para la instalación del sistema, sino para “cada uso” del mismo de modo concreto (nº 388). La petición ha de estar motivada y dirigirse a la “autoridad del lugar donde tendrá lugar uso” (nº 395). En caso de urgencia real, se puede empezar a usar el sistema y luego solicitar la autorización en un plazo máximo de 24 horas. La solicita la “autoridad competente” de los fines de aplicación de la ley”. Si se deniega, todo uso finaliza inmediatamente y se destruyen todos los datos obtenidos. Cabe señalar que la autorización ha de tener en cuenta todos los requisitos de evaluación de proporcionalidad y delimitación espacial y temporal. Así las cosas, se da una “doble evaluación de la necesidad y proporcionalidad” (nº 381) en el FRIA y por la autoridad.

Supervisión obligatoria. Ninguna autoridad puede tomar *decisiones con efectos jurídicos adversos para una persona únicamente en base a los resultados del sistema de identificación biométrica* en tiempo real. Se requiere la verificación y supervisión por al menos dos humanos cualificados cumpliendo los requisitos de supervisión del artículo 14, salvo que la normativa nacional determine excepciones proporcionales.

Cada vez que se use un sistema RBI en tiempo real en espacios públicos con fines policiales, se *notificará a la autoridad de vigilancia del mercado y a la autoridad de protección de datos pertinentes*, sin incluir “datos operativos sensibles”.

Se recuerda finalmente que todos los demás usos de los sistemas de reconocimiento biométrico que no estén prohibidos por el artículo 5. 1. h) entran en la categoría de sistemas de IA de alto riesgo del Anexo III. Así sucede con los sistemas de verificación/autenticación biométrica y el uso retrospectivo de datos (en espacios de acceso público con fines policiales. Por ejemplo, un reconocimiento facial retrospectivo para comparar imágenes de sospechosos de delitos con imágenes faciales registradas en una base de datos criminal. Tampoco está prohibido y es de alto riesgo el uso de sistemas en tiempo real con fines policiales en un espacio privado (como en casa de alguien) o en línea (como el uso de una sala de chat o un juego en línea para identificar a un sospechoso de difundir material de abuso sexual infantil). Igualmente es de alto riesgo y no prohibido el uso de sistemas por privados, tanto en tiempo real como a posteriori, por ejemplo de un supermercado para identificar a conocidos ladrones de tiendas o un recinto deportivo para identificar a personas a las que se prohíbe la entrada

al recinto o en escuelas con fines de seguridad y asistencia escolar). De nuevo hay que recordar que aunque no estén prohibidos, en modo alguno el RIA implica su habilitación legal, sino que será necesaria la oportuna habilitación legal en razón de protección de datos o, en su caso, de cualquier otro derecho fundamental.

Asímismo, hay un sistema de supervisión y control imponiendo obligaciones de notificación y reporte a autoridades nacionales y a la Comisión Europea. Se deben designar una autoridad de vigilancia del mercado, las autoridades deben enviar informes anuales con datos sobre autorizaciones concedidas, decisiones judiciales y cumplimiento del marco normativo. La Comisión, con base en esta información, publicará informes anuales con datos agregados, sin incluir detalles operativos sensibles que puedan comprometer investigaciones. Estos informes anuales permitirán a la Comisión evaluar tendencias y asegurar la correcta aplicación del RIA.

XI. PARA CONCLUIR

Las prohibiciones establecidas en el artículo 5 del RIA marcan un punto de inflexión en la regulación de la IA, tanto para la UE como para el contexto global. Desde el 2 de febrero de 2025, estas prohibiciones son ya plenamente exigibles, convirtiéndose en el primer gran bloque normativo del RIA en entrar en vigor, con efectos inmediatos para todos los operadores que diseñan, implementan o comercializan sistemas de IA en el mercado europeo. Estas restricciones no son meras declaraciones de principios, sino que se sustentan en un sólido marco sancionador y de supervisión, que permitirá a las autoridades europeas y nacionales perseguir activamente su incumplimiento.

El amplio y complejo artículo 5 aparentemente sería un tajante “no es no” cuando prohíbe de sistemas IA que facilitan la manipulación, la explotación de vulnerabilidades, la vigilancia masiva o el control social inaceptable. Algunas prohibiciones son teóricamente absolutas y no admiten excepciones, mientras que otras permiten excepciones, como sucede particularmente en el ámbito de la identificación biométrica remota en tiempo real con fines policiales. Sin embargo, respecto de todas las prohibiciones es imprescindible precisar los conceptos, delimitar su alcance y posibles exclusiones de aplicación y analizar rigurosamente las excepciones permitidas y bajo qué requisitos. Es así que los Criterios adoptados por la Comisión Europea en febrero de 2025 son un enorme referente. No se trata de un mero *softlaw*, puesto que está previsto por el propio artículo 96 1 b), RIA. Hasta que lleguen las primeras sentencias o cambios normativos, algo que puede hacerse esperar mucho, estos criterios Constituyen el referente más sólido para la interpretación del RIA. De ahí que el estudio haya destilado los elementos más relevantes de estos Criterios.

Debe tenerse en cuenta la relación entre las prohibiciones del artículo 5 y la categoría de sistemas de alto riesgo definidos en el Anexo III del RIA. En muchos casos, si un sistema del artículo 5 no está prohibido, será de alto riesgo y quedará sujeto a las exigencias y salvaguardas del RIA que impone estrictos requisitos en cuanto a evaluación de impacto, transparencia, supervisión, registro gobernanza de datos y otras obligaciones. La distinción entre lo que es IA prohibida y la IA de alto riesgo es fundamental, sin embargo, hay una muy *delgada línea roja* entre ambos ámbitos, lo que exige un escrutinio constante y tener claros los criterios de uno y otro. Cabe señalar que se percibe en los criterios de la Comisión Europea, se intenta evitar una interpretación excesivamente severa de las prácticas prohibidas que pueda dificultar o frenar en seco la innovación en diversas áreas.

Será necesario esperar varios años para evaluar cómo operan en la práctica las prohibiciones del artículo 5 y determinar si requieren actualización, reducción o ampliación. Muy posiblemente lo que está por venir nos obligará a repensar prohibiciones en materia de IA generativa avanzada, de aprendizaje profundo no supervisado o respecto de los modelos de IA multimodal, que sin duda plantearán nuevos riesgos que no están claramente cubiertos por las actuales prohibiciones. Además, la rapidez con la que evoluciona este sector probablemente nos obligará a superar la ya difusa distinción entre sistemas prohibidos y de alto riesgo

La aplicación del artículo 5 recae ahora en las autoridades nacionales y europeas, así como bajo control jurisdiccional. El legislador europeo y nacional tiene también muchos deberes pendientes y capacidad regulatoria respecto de las prohibiciones, especialmente en lo relativo a identificación biométrica y sus múltiples posibilidades de excepción, pues más que un no es no, se trata de un claro, no pero sí. Su actuación bajo el principio de legalidad y calidad es imperiosa y debe estar realmente acompañada de un debate y deliberación democrática que no siempre parece que se vaya a producir.