RETECH Redes Territoriales de Especialización Tecnológica

RED ARGOS

Guía de clasificación de sistemas de inteligencia artificial en base al Reglamento de IA de la UE en los entornos de Smart Cities y Salud

Septiembre de 2025

Versión 1.0

Junta de Andalucía

Financiado por la Unión Europea NextGenerationEU

Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales

Plan de Recuperación, Transformación y Resiliencia

incibe_ Instituto Nacional de Ciberseguridad

Junta de Andalucía

Agencia Digital de Andalucía

Objeto del Expediente:

"ELABORACIÓN DE GUÍAS DE CIBERSEGURIDAD IA PARA ENTORNOS DE SMART CITIES Y SALUD" (CONTR 2024 829535).

Esta guía forma parte de la colección Guías de Ciberseguridad en IA para las Smart Cities y el sector Salud creada por la Agencia Digital de Andalucía como parte del Proyecto Red Argos, perteneciente al programa RETECH, de Redes Inteligentes de Especialización Tecnológica, en el marco del Plan de Recuperación, Transformación y Resiliencia, financiado por la Unión Europea-NextGenerationEU, a través de INCIBE.

Autor del documento: Agencia Digital de Andalucía

Tipo de documento: Guía

Fecha de elaboración: 05/09/2025

Fecha de última actualización: 03/10/2025

ÍNDICE

Índice de ilustraciones

Índice de tablas

1. INTRODUCCIÓN

La inteligencia artificial (IA) se ha convertido en un elemento central de la transformación digital europea, con aplicaciones que impactan directamente en la salud, la movilidad, la energía y la gestión urbana. En el ámbito sanitario, la IA permite mejorar la precisión diagnóstica, optimizar los recursos y personalizar los tratamientos; en los entornos urbanos, favorece una gestión más eficiente y sostenible de los servicios públicos. Sin embargo, junto a estos beneficios surgen riesgos que requieren una clasificación y una supervisión adecuadas. Un sistema de IA mal definido o insuficientemente controlado puede generar errores clínicos, sesgos en la toma de decisiones o fallos en infraestructuras críticas, afectando derechos fundamentales y la confianza social.

Para abordar estos desafíos, la Unión Europea adoptó el **Reglamento (UE) 2024/1689, relativo a la inteligencia artificial (RIA),** que establece normas armonizadas destinadas a garantizar que los sistemas de IA introducidos o utilizados en el mercado de la Unión sean seguros, transparentes y coherentes con los valores europeos. El RIA no establece una jerarquía cerrada de niveles de riesgo, sino un sistema de clasificación basado en la evaluación del riesgo potencial que un sistema puede generar en su contexto de uso, especialmente en relación con la salud, la seguridad y los derechos fundamentales.

A efectos de interpretación, esta clasificación distingue entre prácticas prohibidas, sistemas de alto riesgo, sistemas de riesgo limitado o de transparencia y sistemas de riesgo mínimo. No se trata de una clasificación alternativa o paralela, sino del propio enfoque estructural del RIA, que permite determinar el grado de supervisión, control y garantía necesario para cada tipo de sistema en función de su impacto potencial.

En el ámbito sanitario, el RIA se aplica de forma complementaria al **Reglamento (UE) 2017/745 sobre productos sanitarios (MDR)** y al **Reglamento (UE) 2017/746 sobre productos para diagnóstico in vitro (IVDR),** que mantienen la responsabilidad principal en materia de evaluación de conformidad y marcado CE de los productos.

Garantizar una clasificación precisa de los sistemas de IA constituye el punto de partida esencial para un uso seguro, responsable y conforme al Derecho de la Unión. En sectores como la salud y las ciudades inteligentes, donde los sistemas procesan datos sensibles y sustentan servicios esenciales, una clasificación incorrecta puede comprometer la seguridad jurídica, generar riesgos técnicos y limitar su acceso al mercado europeo.

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
Plan de Recuperación, Transformación y Resiliencia
incibe_ INSTITUTO NACIONAL DE CIBERSEGURIDAD
Junta de Andalucía
Agencia Digital de Andalucía

2. OBJETIVO Y ALCANCE

2.1. Objetivo

El objetivo de esta guía es proporcionar un marco metodológico claro y estructurado para la clasificación de los sistemas de inteligencia artificial (IA) que se desarrollan, comercializan o utilizan en los sectores de Salud y Ciudades Inteligentes, conforme a lo establecido en el **Reglamento (UE) 2024/1689 sobre inteligencia artificial (RIA)**. La guía tiene por finalidad facilitar la determinación del nivel de riesgo aplicable a cada sistema (prohibido, alto, limitado o de transparencia, o mínimo) en función de su finalidad prevista, su contexto de uso y su posible impacto sobre la salud, la seguridad o los derechos fundamentales de las personas.

En el sector sanitario, la aplicación del RIA debe entenderse en coherencia con el Reglamento (UE) 2017/745 sobre productos sanitarios (MDR) y el Reglamento (UE) 2017/746 sobre productos para diagnóstico in vitro (IVDR), garantizando que los procedimientos de evaluación de conformidad y el marcado CE se completen previamente a la clasificación RIA.

El documento está concebido para servir de apoyo a todos los actores implicados en el ciclo de vida de la IA como, proveedores tecnológicos, autoridades competentes, responsables de cumplimiento normativo, profesionales sanitarios, gestores municipales y equipos de ciberseguridad, con el fin de ofrecer criterios homogéneos y fundamentos prácticos y verificables para determinar con precisión el nivel de riesgo de cada sistema. Una clasificación adecuada constituye la condición necesaria para desplegar soluciones de IA de forma legal, segura y confiable en estos sectores críticos.

2.2. Alcance

El alcance de esta guía se limita a los sectores de Salud y Ciudades Inteligentes. Sus contenidos se aplican a los sistemas de inteligencia artificial introducidos en el mercado, puestos en servicio o utilizados en el territorio de la Unión Europea dentro de estos ámbitos.

Las orientaciones aquí recogidas tienen por objeto facilitar la identificación y la valoración del nivel de riesgo de los sistemas de IA que operan en entornos sanitarios y urbanos, atendiendo a sus particularidades técnicas, regulatorias y operativas. El documento abarca tanto los sistemas destinados a uso clínico, asistencial o de gestión sanitaria, como las soluciones implantadas en infraestructuras y servicios municipales relacionados con la movilidad, la energía, la seguridad, la sostenibilidad y la administración pública local.

Esta guía no sustituye las disposiciones generales del **Reglamento (UE) 2024/1689 relativo a la inteligencia artificial (RIA)**, ni las de la normativa sectorial aplicable. Actúa como instrumento de referencia para su correcta aplicación en los ámbitos indicados y aporta criterios consistentes y verificables que permitan una clasificación coherente y comparable de los sistemas de IA en estos sectores.

En el sector sanitario, las orientaciones deben interpretarse en coherencia con el **Reglamento (UE) 2017/745 relativo a los productos sanitarios (MDR)** y el **Reglamento (UE) 2017/746 relativo a los productos para diagnóstico in vitro (IVDR).** Los procedimientos de evaluación de conformidad y marcado CE se realizan con carácter previo a la clasificación conforme al RIA.

Quedan fuera del alcance de esta guía los sistemas de IA desarrollados o utilizados con fines militares, de defensa o de seguridad nacional, así como los utilizados exclusivamente en investigación científica o experimental que no estén destinados a su puesta en el mercado ni a su uso en entornos reales dentro de la Unión.

3. DEFINICIONES

Término	Definición
Reglamento de Inteligencia Artificial (RIA)	Reglamento (UE) 2024/1689, que establece el marco jurídico armonizado en materia de inteligencia artificial en la Unión Europea, basado en un enfoque de clasificación por niveles de riesgo. Define obligaciones diferenciadas según el impacto potencial de los sistemas de IA en la salud, la seguridad y los derechos fundamentales.
Reglamento de productos sanitarios (MDR)	Reglamento (UE) 2017/745, que regula los productos sanitarios en la UE, incluidos los dispositivos médicos con software como componente independiente. Establece criterios de clasificación, requisitos de seguridad y procedimientos de evaluación de conformidad y marcado CE.
Reglamento de productos sanitarios para diagnóstico in vitro (IVDR)	Reglamento (UE) 2017/746, que regula los productos sanitarios para diagnóstico in vitro en la UE. Abarca reactivos, equipos, software y otros productos destinados al examen de muestras humanas con fines médicos, estableciendo requisitos de evaluación de conformidad y trazabilidad (UDI).
Sistema de inteligencia artificial (IA)	Según el artículo 3.1 del RIA, se refiere a un sistema basado en máquina que, para objetivos explícitos o implícitos, infiere a partir de entradas cómo generar salidas como predicciones, recomendaciones, decisiones o contenidos, que pueden influir en entornos físicos o virtuales.
Producto sanitario	Todo instrumento, aparato, equipo, software, implante, reactivo, material u otro artículo destinado por el fabricante a fines médicos específicos, como diagnóstico, prevención, control, tratamiento o alivio de una enfermedad, conforme a la definición del MDR.
Producto sanitario para diagnóstico in vitro (IVD)	Producto sanitario constituido por reactivos, calibradores, materiales de control, equipos, software u otros artículos destinados por el fabricante al examen in vitro de muestras procedentes del cuerpo humano, con fines médicos, conforme a la definición del IVDR.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

Término	Definición
Software como producto sanitario (SaMD)	Software que cumple por sí mismo la definición de producto sanitario, al tener una finalidad médica prevista, tal como se contempla en la Regla 11 del Anexo VIII del MDR y en el artículo 2 del IVDR. Puede constituir un dispositivo médico autónomo o formar parte de un sistema de diagnóstico in vitro.
Proveedor	Persona física o jurídica, autoridad pública, agencia u organismo que desarrolla un sistema de IA o lo introduce en el mercado o lo pone en servicio bajo su propio nombre o marca, asumiendo la responsabilidad por su conformidad (RIA, art. 3.2).
Responsable de despliegue	Persona física o jurídica, autoridad pública, agencia u organismo que utiliza un sistema de IA bajo su autoridad, salvo cuando dicho uso sea de carácter personal y no profesional (RIA, art. 3.4).
Importador	Persona física o jurídica establecida en la Unión que introduce en el mercado un sistema de IA procedente de un tercer país, asegurando que cumple con los requisitos del RIA y, en su caso, de legislación sectorial como MDR o IVDR (RIA, art. 3.6).
Distribuidor	Persona física o jurídica en la cadena de suministro, distinta del proveedor o del importador, que pone a disposición en el mercado un sistema de IA sin intervenir en su desarrollo, garantizando que el producto lleva el marcado CE y va acompañado de la documentación pertinente (RIA, art. 3.7).
Evaluación de conformidad	Procedimiento mediante el cual se demuestra que un producto cumple con los requisitos aplicables de la normativa de la Unión (RIA, MDR, IVDR). Puede incluir auditorías, pruebas, revisión de la documentación técnica y, cuando proceda, la intervención de un organismo notificado.
Marcado CE	Marca que indica que un producto cumple con los requisitos esenciales de seguridad, salud y protección de la normativa de la UE. Es obligatoria para la comercialización de productos sanitarios y de diagnóstico in vitro, incluido software con finalidad médica, y se exige antes de la aplicación complementaria del RIA.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

Término	Definición
Clasificación de riesgo (RIA)	Proceso normativo por el cual un sistema de IA se categoriza en una de cuatro clases: riesgo inaceptable (prohibido), alto riesgo, riesgo limitado o riesgo mínimo, en función de su finalidad, contexto de uso e impacto potencial en la salud, la seguridad y los derechos fundamentales.
Prácticas prohibidas / Riesgo inaceptable	Categoría de sistemas o usos de IA expresamente prohibidos por el RIA por ser contrarios a los derechos fundamentales, la dignidad humana o la seguridad, tales como la manipulación subliminal, la puntuación social generalizado o el reconocimiento biométrico remoto en tiempo real en espacios públicos (RIA, art. 5).
Sistema de alto riesgo	Sistema de IA sujeto a obligaciones reforzadas por estar incluido en el Anexo I del RIA (como componente de seguridad de productos regulados por MDR o IVDR) o en el Anexo III (aplicaciones críticas en sanidad, infraestructuras, empleo, etc.).
Sistema de riesgo limitado o de transparencia	Sistema de IA sujeto a requisitos de transparencia conforme al artículo 50 del RIA, cuando interactúa directamente con personas físicas, genera o manipula contenidos sintéticos, o realiza reconocimiento de emociones o categorización biométrica, sin estar clasificado como de alto riesgo ni constituir práctica prohibida.
Sistema de riesgo mínimo	Sistema de IA cuyo impacto en la salud, la seguridad o los derechos fundamentales es irrelevante y que no está sujeto a obligaciones regulatorias específicas bajo el RIA. Ejemplos: filtros de spam, motores de búsqueda simples, videojuegos con IA.
Documentación técnica	Conjunto de documentos que describen las características, diseño, datos, pruebas, controles y prestaciones de un sistema de IA o de un producto regulado por MDR/IVDR. Es esencial para demostrar conformidad con los requisitos aplicables.
Gestión de riesgos	Proceso sistemático de identificación, análisis, mitigación y seguimiento de los riesgos asociados al sistema de IA, exigido de manera reforzada en los sistemas de alto riesgo y complementario en los regulados por MDR e IVDR.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

4. ÁMBITO DE APLICACIÓN Y ROLES IDENTIFICADOS

La presente sección establece las condiciones en las que resulta aplicable el Reglamento (UE) 2024/1689 relativo a la inteligencia artificial (RIA), define los supuestos excluidos y describe los actores que intervienen en el ciclo de vida de los sistemas de IA. Su propósito es ofrecer un marco inequívoco que permita a las organizaciones del ámbito sanitario y de las Smart Cities determinar si un sistema concreto de inteligencia artificial está sujeto al RIA y cómo deben proceder en materia de cumplimiento normativo.

Se aclara que las disposiciones del RIA se aplican sin perjuicio de otras normas sectoriales de la UE que regulen aspectos de seguridad o protección de datos (GDPR, MDR, IVDR, NIS 2, DORA cuando aplique).

4.1. Ámbito de aplicación

De conformidad con el artículo 2 del Reglamento (UE) 2024/1689, el ámbito de aplicación comprende todos los sistemas de inteligencia artificial que se introduzcan en el mercado, se pongan en servicio o se utilicen dentro del territorio de la Unión Europea, con independencia del lugar en que esté establecido el proveedor o el responsable de despliegue.

Asimismo, se consideran comprendidos los sistemas cuyos resultados o salidas se utilicen en la Unión, aun cuando su desarrollo, entrenamiento o explotación se haya realizado fuera de ella. Este principio de aplicabilidad territorial y funcional garantiza que cualquier sistema de IA con efectos en la UE quede sujeto a las disposiciones del RIA.

En el sector sanitario, la aplicabilidad del RIA debe interpretarse en coherencia con el Reglamento (UE) 2017/745 (MDR) y el Reglamento (UE) 2017/746 (IVDR). Cuando un sistema de IA constituye un producto sanitario o un producto sanitario para diagnóstico in vitro, se aplican de forma complementaria los procedimientos de clasificación y evaluación de conformidad previstos en los artículos 52 a 54 del MDR y 48 a 50 del IVDR, incluyendo, cuando proceda, la intervención de organismo notificado, el marcado CE y la asignación del UDI. Una vez cumplidos estos requisitos previos, el sistema debe ser analizado adicionalmente conforme a la metodología de clasificación del RIA, sin que ello sustituya las obligaciones derivadas de la normativa sectorial de productos sanitarios.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

En Ciudades Inteligentes, el RIA resulta aplicable a sistemas de IA utilizados por ayuntamientos, empresas públicas, concesionarias y operadores de servicios urbanos cuando sus salidas se emplean en la gestión de movilidad (control adaptativo de semáforos, regulación de túneles), energía y redes (gestión de demanda, microrredes, alumbrado), agua y residuos (telecontrol de estaciones de bombeo, optimización de rutas), medio ambiente (calidad del aire, ruido), seguridad urbana (analítica de vídeo con fines no policiales) y administración pública local (gestión de citas, priorización de trámites). Igualmente quedan incluidos los sistemas prestados como servicio en la nube cuando sus resultados se utilizan en la Unión, con independencia del país del proveedor.

4.1.1. Definición de sistema de inteligencia artificial conforme al RIA

El artículo 3(1) del RIA define un sistema de inteligencia artificial (IA) como: *un sistema basado en una máquina que, para objetivos explícitos o implícitos, infiere a partir de las entradas que recibe cómo generar salidas como predicciones, contenidos, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales.*

Interpretación práctica:

- Basado en máquina: incluye tanto software independiente como IA integrada en hardware (p. ej., robots quirúrgicos o sensores de movilidad urbana).
- Inferencia a partir de entradas: va más allá del simple procesamiento de datos; implica el uso de algoritmos capaces de detectar patrones, aprender de ejemplos o realizar razonamientos.
- Generación de salidas: las salidas pueden adoptar múltiples formas: diagnósticos médicos, recomendaciones de tratamiento, predicciones de tráfico, alertas de consumo energético, etc.
- Influencia en entornos físicos o virtuales: la IA puede incidir directamente en decisiones clínicas, en la gestión hospitalaria o en la organización de servicios urbanos.

Esta definición coincide con la establecida en ISO/IEC 22989:2022, lo que refuerza la coherencia entre el marco normativo europeo y los estándares internacionales de terminología en IA.

4.1.2. Condiciones generales de aplicabilidad

El ámbito de aplicación del RIA está regulado principalmente en su artículo 2. De forma general, el Reglamento se aplica a todo sistema de inteligencia artificial que sea introducido en el mercado de la Unión, puesto en servicio o utilizado en su territorio, con independencia de que el proveedor o el responsable de despliegue esté o no establecido en un Estado miembro.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

Las condiciones generales pueden resumirse en tres ejes fundamentales:

I. **Territorialidad**: el RIA se aplica a los sistemas de IA comercializados, desplegados o utilizados dentro del territorio de la Unión Europea. Esto incluye tanto sistemas desarrollados en la UE como en terceros países, siempre que los resultados de salida se empleen en la Unión. Este principio incorpora el criterio de **extraterritorialidad**, de modo que también se incluyen los sistemas cuyos resultados se utilicen en la Unión, aunque hayan sido desarrollados o desplegados fuera de ella.

   - *Ejemplo en salud*: una aplicación de diagnóstico por imagen desarrollada en EE. UU. que genera informes utilizados por hospitales europeos debe cumplir con el RIA.

   - *Ejemplo en Smart Cities*: un software de optimización de tráfico diseñado en un país asiático, pero implementado en Europa entra plenamente en el ámbito de aplicación del RIA.

II. **Principio de extraterritorialidad**: la protección de los derechos fundamentales europeos se extiende más allá de las fronteras de la UE. Así, cualquier sistema que produzca efectos en la Unión queda sometido al Reglamento, incluso si el proveedor no tiene presencia física en Europa.

III. **Actividad regulada**: el Reglamento no se limita al momento de comercialización, sino que cubre también la puesta en servicio y el uso de los sistemas. Esto significa que incluso entidades que no sean las desarrolladoras iniciales (hospitales, ayuntamientos, proveedores de servicios urbanos) están sujetas a las disposiciones.

4.1.3. Condiciones específicas para considerar

Además de las condiciones generales, existen casuísticas particulares que deben analizarse con atención:

- **Sistemas de propósito general (GPAI)**: modelos reutilizables en diferentes ámbitos (por ejemplo, modelos fundacionales de lenguaje o visión). Están dentro del ámbito del RIA cuando sus implementaciones concretas puedan generar riesgos significativos en salud o en Smart Cities.

- **Usos duales**: sistemas desarrollados para investigación o fines generales que, al reutilizarse en entornos sanitarios o urbanos, impactan directamente en ciudadanos y, por tanto, entran en el RIA.

- **Integración en productos regulados**: cuando la IA constituye parte de un producto sometido a otra legislación (como un dispositivo médico), la obligación de cumplimiento del RIA se extiende al conjunto.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

- **Servicios de IA en la nube (AIaaS)**: comprendidos siempre que los resultados se utilicen dentro de la UE, independientemente de la localización del proveedor.
- **Cadenas de subcontratación**: todos los actores implicados en el diseño, integración o despliegue de un sistema de IA mantienen responsabilidades diferenciadas conforme a su rol y la clasificación del riesgo potencial del sistema.

4.1.4. Exclusiones - sistemas a los que no les aplica el RIA

Conforme al **artículo 2, apartado 5, del Reglamento (UE) 2024/1689**, quedan excluidos del ámbito de aplicación los siguientes sistemas:

- **Defensa, seguridad y fines militares:** los sistemas de IA utilizados exclusivamente con fines militares, de defensa o seguridad nacional, cuya regulación corresponde a los Estados miembros.
- **Investigación y desarrollo científico:** los sistemas de IA empleados únicamente en el marco de la investigación y el desarrollo, mientras no se introduzcan en el mercado ni se pongan en servicio dentro de la Unión.
- **Cooperación policial o judicial internacional:** determinados sistemas empleados en el marco de acuerdos internacionales con garantías jurídicas equivalentes.

4.2. Roles identificados

El RIA identifica a los siguientes operadores en el ciclo de vida de los sistemas de inteligencia artificial:

- **Proveedor** (artículo 3.2 del RIA): Una persona física o jurídica, autoridad pública, agencia u otro organismo que desarrolla un sistema de IA o lo introduce en el mercado o lo pone en servicio bajo su propio nombre o marca. Se trata del actor que asume la autoría o responsabilidad principal del sistema, ya sea por haberlo creado o por comercializarlo con su identidad.
- **Responsables de despliegue** (artículo 3.4 del RIA): persona física o jurídica, autoridad pública, agencia u otro organismo que utiliza un sistema de IA bajo su autoridad, salvo cuando el uso sea personal y no profesional. Este rol corresponde a la entidad que pone en funcionamiento el sistema dentro de sus operaciones, aunque no sea el desarrollador original.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

- **Importador** (artículo 3.6 del RIA): Una persona física o jurídica establecida en la Unión que introduce en el mercado un sistema de IA procedente de un tercer país. El importador actúa como responsable de garantizar que los sistemas desarrollados fuera de la UE cumplen con las disposiciones aplicables antes de ser comercializados en Europa.
- **Distribuidor** (artículo 3.7 del RIA): persona física o jurídica de la cadena de suministro, distinta del proveedor o importador, que pone a disposición en el mercado un sistema de IA. Su función es la comercialización o reventa de sistemas ya fabricados, sin haber intervenido en su desarrollo, aunque con la obligación de verificar su conformidad antes de su distribución.
- **Representante autorizado** (artículo 3.8 del RIA): Una persona física o jurídica establecida en la Unión que ha recibido un mandato escrito de un proveedor situado fuera de la UE para actuar en su nombre en relación con el cumplimiento del Reglamento. Este papel es esencial para asegurar la comunicación con las autoridades europeas cuando el proveedor se encuentra en un país tercero.

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA SECRETARÍA DE ESTADO DE TELECOMUNICACIONES E INFRAESTRUCTURAS DIGITALES
Plan de Recuperación, Transformación y Resiliencia
incibe_ INSTITUTO NACIONAL DE CIBERSEGURIDAD
Junta de Andalucía
Agencia Digital de Andalucía

5. EVALUACIÓN DE CASO DE USO

La evaluación de caso de uso constituye la fase preliminar del proceso de clasificación de sistemas de inteligencia artificial conforme al Reglamento (UE) 2024/1689 (en adelante, RIA). Esta fase tiene por objeto recopilar y analizar la información técnica, funcional y contextual necesaria para determinar si un sistema se encuentra sujeto al ámbito de aplicación del Reglamento y, en caso afirmativo, para preparar su clasificación formal.

En los sectores de salud y ciudades inteligentes, esta evaluación adquiere relevancia estratégica. Estos ámbitos concentran sistemas que procesan datos sensibles, sostienen infraestructuras críticas y afectan directamente a derechos fundamentales. Una identificación inadecuada puede derivar en errores de clasificación con consecuencias graves: sanciones administrativas, impacto reputacional o daños para pacientes y ciudadanos.

5.1. Identificar y categorizar los sistemas de IA

La correcta clasificación de un sistema de inteligencia artificial requiere disponer de información completa sobre sus características técnicas, su finalidad prevista y su contexto de despliegue. Esta información constituye la base documental que sustentará el análisis de aplicabilidad del RIA y la posterior clasificación de riesgo conforme a la metodología establecida en el apartado 7.

5.1.1. Entendimiento y análisis del propósito y contexto

Este paso aborda toda la información que se necesita recopilar para evaluar si el sistema entra en la definición de IA y para sustentar su clasificación:

I. **Propósito y alcance funcional: Propósito y alcance funcional del sistema**

   Es necesario identificar claramente la finalidad prevista del sistema, las tareas principales que realiza (predicción, clasificación, generación, control o recomendación), los resultados de salida que genera y las áreas de decisión o procesos que apoya o automatiza. Esta información permite determinar si el sistema se ajusta a la definición de inteligencia artificial establecida en el artículo 3.1 del RIA.

II. **Naturaleza del sistema y nivel de autonomía**

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
Plan de Recuperación, Transformación y Resiliencia
incibe_ INSTITUTO NACIONAL DE CIBERSEGURIDAD
Junta de Andalucía
Agencia Digital de Andalucía

Debe determinarse si el sistema es específico para un caso de uso concreto o si constituye un modelo de inteligencia artificial de propósito general (GPAI) conforme a los artículos 51 y 52 del RIA. Asimismo, resulta relevante identificar el nivel de autonomía del sistema en la toma de decisiones: asistencial (la decisión final corresponde a un operador humano), semiautónomo (el sistema toma decisiones bajo supervisión humana) o autónomo (el sistema opera sin intervención humana directa).

III. Usuarios y colectivos afectados

Es necesario identificar los perfiles de usuarios profesionales que operan el sistema, los colectivos o personas físicas cuyos derechos o intereses pueden verse impactados por los resultados del sistema, y si existe interacción directa con ciudadanos, pacientes u otros colectivos vulnerables. Esta información resulta esencial para evaluar el impacto potencial en derechos fundamentales.

IV. 5.1.4 Entorno de uso y nivel de criticidad

Debe documentarse el sector de aplicación (salud o ciudades inteligentes), el contexto operativo específico (hospital, centro de salud, red de transporte, infraestructura urbana), el nivel de criticidad del servicio que soporta el sistema (bajo, medio o alto) y los riesgos preliminares asociados al entorno de despliegue (seguridad física, privacidad, discriminación, salud pública). Esta información permite encuadrar el sistema en los ámbitos contemplados en el Anexo III del RIA.

V. Tecnología y tratamiento de datos

Es necesario identificar los tipos de datos de entrada procesados por el sistema (estructurados, no estructurados, biométricos, clínicos, procedentes de sensores urbanos), su naturaleza jurídica (personales, sensibles conforme al artículo 9 del Reglamento (UE) 2016/679, anonimizados o sintéticos), su origen (internos, externos o procedentes de terceros) y las técnicas algorítmicas empleadas. Debe indicarse asimismo si existen dependencias técnicas con proveedores externos o servicios en la nube, y si el sistema presenta requisitos específicos de explicabilidad o interpretabilidad.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

En el caso de sistemas desplegados en el ámbito sanitario, debe determinarse específicamente si el sistema constituye un producto sanitario conforme al Reglamento (UE) 2017/745 o un producto sanitario para diagnóstico in vitro conforme al Reglamento (UE) 2017/746. En tal caso, debe identificarse la clasificación del producto según las reglas del Anexo VIII del MDR o del IVDR, el estado del procedimiento de evaluación de conformidad, la intervención de organismos notificados (si procede) y el código de identificador único de producto (UDI), si ya hubiera sido asignado.

VI. Interacciones con personas y capacidades de generación de contenido

Es relevante identificar los canales de interacción del sistema con usuarios o personas afectadas, sus capacidades de generación de contenido sintético (texto, audio, imagen o vídeo), la incorporación de funcionalidades de reconocimiento biométrico o reconocimiento de emociones, y la existencia de medidas de transparencia previstas para informar sobre el uso del sistema. Esta información resulta determinante para evaluar la aplicabilidad de las obligaciones de transparencia establecidas en el artículo 50 del RIA.

VII. Usos duales y riesgos preliminares

Debe indicarse si el sistema presenta usos duales (empleo simultáneo o alternativo en investigación y en aplicación real), los riesgos preliminares detectados por el equipo responsable de la evaluación, y cualquier otra información relevante para la caracterización del sistema.

5.1.2. Verificación de excepciones de aplicación del RIA

No todos los sistemas de inteligencia artificial se encuentran sujetos al ámbito de aplicación del RIA. El artículo 2 del Reglamento establece un conjunto de excepciones que excluyen determinados sistemas del cumplimiento de sus disposiciones.

Entre las excepciones principales artículo 2 del RIA, los siguientes supuestos quedan fuera del ámbito de aplicación del Reglamento:

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

- **Sistemas utilizados exclusivamente con fines militares, de defensa o seguridad nacional:** Los sistemas de inteligencia artificial empleados exclusivamente con fines de defensa, seguridad nacional o actividades militares no se encuentran sujetos al RIA. La regulación de estos sistemas corresponde a los Estados miembros en virtud de sus competencias exclusivas en materia de seguridad nacional conforme al artículo 4.2 del Tratado de la Unión Europea.
  Ejemplo: un sistema de IA utilizado para analizar imágenes de satélite con fines de defensa.

- **Sistemas empleados únicamente en investigación y desarrollo:** Los sistemas de inteligencia artificial que se utilicen únicamente en el marco de actividades de investigación y desarrollo científico, sin ser puestos en servicio ni comercializados en la Unión Europea, están excluidos del ámbito de aplicación del RIA. Esta exclusión se fundamenta en la necesidad de preservar la libertad de investigación científica, garantizando al mismo tiempo que los sistemas que lleguen al mercado o sean desplegados en entornos reales queden sujetos a las obligaciones del Reglamento.
  En el caso específico de la investigación sanitaria, un sistema de inteligencia artificial utilizado exclusivamente en ensayos clínicos o estudios de investigación biomédica, sin comercialización ni uso en práctica clínica asistencial, queda excluido del RIA mientras mantenga este carácter exclusivamente experimental. Si posteriormente el sistema se introduce en el mercado como producto sanitario, deberá someterse tanto a las obligaciones del MDR o IVDR como a las del RIA de manera complementaria.
  Ejemplo: un prototipo experimental de IA para análisis genómico que aún no se usa en la práctica clínica.

- **Sistemas utilizados en cooperación policial o judicial internacional:** Determinadas actividades de cooperación internacional en materia policial y judicial pueden quedar fuera del alcance del Reglamento, siempre que se rijan por acuerdos específicos con garantías adecuadas.
  Ejemplo: Un sistema compartido en un proyecto piloto entre la Europol y un país tercero para investigación criminal.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

5.1.3. Documentación general

Para cada sistema evaluado, debe conservarse evidencia documental de la evaluación de aplicabilidad realizada y su resultado (sujeto al RIA o excluido), los criterios técnicos y jurídicos empleados para determinar la inclusión o exclusión del ámbito del RIA, las referencias normativas aplicadas (con mención específica al artículo 2 del RIA y apartados pertinentes), y la fecha de realización de la evaluación junto con la identificación de los responsables designados.

5.1.3.1. Documentación específica para el ámbito sanitario

En el ámbito sanitario, debe documentarse adicionalmente la determinación fundamentada de si el sistema constituye producto sanitario conforme al MDR o producto de diagnóstico in vitro conforme al IVDR, la clasificación preliminar del producto según las reglas del Anexo VIII del MDR o del IVDR, el estado del procedimiento de evaluación de conformidad y marcado CE (si ya se hubiera iniciado), la identificación del organismo notificado designado (en caso de ser preceptiva su intervención) y el código de identificador único de producto (UDI), si ya hubiera sido asignado.

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA SECRETARÍA DE ESTADO DE TELECOMUNICACIONES E INFRAESTRUCTURAS DIGITALES
Plan de Recuperación, Transformación y Resiliencia
incibe_ INSTITUTO NACIONAL DE CIBERSEGURIDAD
Junta de Andalucía
Agencia Digital de Andalucía

6. CLASIFICACIÓN DE SISTEMAS DE IA SEGÚN EL RIA

El Reglamento (UE) 2024/1689 sobre inteligencia artificial (en adelante, RIA) establece un marco regulatorio basado en el riesgo que constituye el eje central de todo el sistema de clasificación. Este marco permite determinar la categoría de riesgo aplicable a cada sistema en función de su impacto potencial sobre la salud, la seguridad y los derechos fundamentales.

El RIA no establece explícitamente una jerarquía cerrada de niveles de riesgo denominados. Sin embargo, su estructura normativa articula diferentes categorías de sistemas de inteligencia artificial en función del riesgo que representan. A efectos de facilitar la comprensión y aplicación práctica del Reglamento, esta guía identifica y describe cuatro categorías principales que se derivan de la lectura sistemática de sus disposiciones:

I. **Riesgo inaceptable**: sistemas incompatibles con los valores y derechos fundamentales de la Unión, cuyo uso está prohibido.

II. **Alto riesgo**: sistemas con potencial de causar perjuicios graves a la seguridad, la salud o los derechos de las personas, sujetos a un régimen regulatorio reforzado.

III. **Riesgo limitado:** sistemas cuyo riesgo es moderado y que quedan sujetos a obligaciones de transparencia conforme al artículo 50 del RIA.

IV. **Riesgo mínimo:** sistemas con impacto reducido que pueden desarrollarse y utilizarse sin restricciones regulatorias específicas adicionales.

Esta estructuración responde a la lógica interna del RIA y permite determinar la categoría de riesgo aplicable a cada sistema.

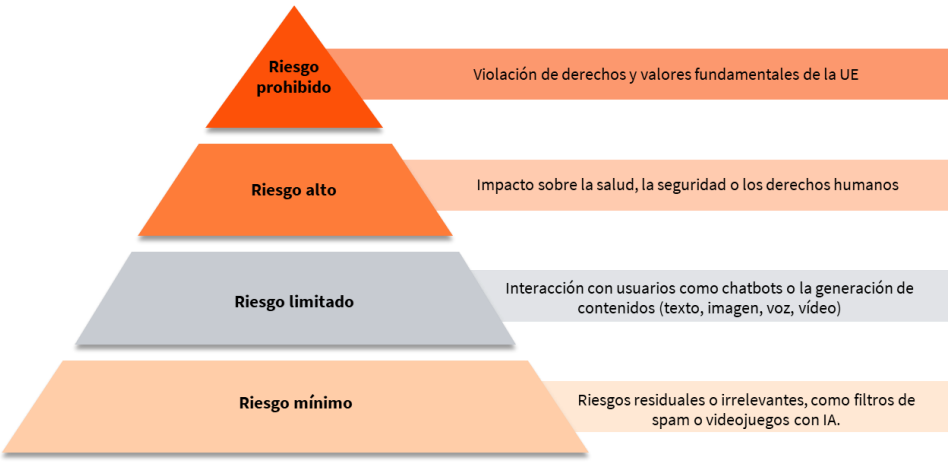

Ilustración 1. Niveles de riesgo del RIA

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA SECRETARÍA DE ESTADO DE TELECOMUNICACIONES E INFRAESTRUCTURAS DIGITALES
Plan de Recuperación, Transformación y Resiliencia
incibe_ INSTITUTO NACIONAL DE CIBERSEGURIDAD
Junta de Andalucía
Agencia Digital de Andalucía

Cada una de estas categorías conlleva consecuencias distintas en cuanto al marco normativo aplicable, la necesidad de control y las medidas de gestión.

6.1. Enfoque basado en niveles de riesgo

Se describe el enfoque metodológico que prioriza la evaluación del nivel de riesgo que un sistema de IA puede generar en función de su propósito, contexto de uso y potencial impacto en derechos fundamentales, seguridad y salud pública.

En el sector sanitario, debe subrayarse que la clasificación de un sistema de IA como producto sanitario o producto de diagnóstico in vitro se rige previamente por los Reglamentos (UE) 2017/745 (MDR) y 2017/746 (IVDR). La determinación del nivel de riesgo conforme al RIA se aplica de manera complementaria, una vez cumplidos los procedimientos de evaluación de conformidad y marcado CE previstos en dichos Reglamentos.

La clasificación es especialmente crítica en sectores como **Salud** o **Smart Cities**, donde la inteligencia artificial se utiliza para:

- Analizar y procesar datos sensibles (p. ej., historiales médicos, datos biométricos).
- Tomar decisiones que afectan al acceso de ciudadanos a servicios esenciales (diagnóstico, movilidad, energía).
- Gestionar infraestructuras críticas (hospitales, redes de transporte, servicios urbanos).

El RIA adopta un enfoque regulatorio **proporcional y gradual**, en el que las exigencias aumentan en función del nivel de riesgo identificado. No se evalúa únicamente la tecnología subyacente, sino el **propósito previsto, el contexto de uso y el impacto potencial** sobre las personas.

Los ejes interpretativos principales son:

- **Finalidad del sistema**: cuál es el objetivo para el que se ha diseñado (incluida la naturaleza de la función: predicción, clasificación, generación, control o recomendación).
- **Contexto de aplicación**: entorno operativo y criticidad del servicio (ej. hospital, espacio público, servicios administrativos).
- **Usuarios previstos**: profesionales, autoridades públicas, usuarios, colectivos vulnerables.
- **Impacto en derechos y seguridad**: sobre salud, seguridad y derechos fundamentales.
- **Grado de autonomía** y control humano efectivo.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

6.1.1. Sistemas de riesgo inaceptable / prácticas prohibidas

De conformidad con el artículo 5 del RIA, se consideran de riesgo inaceptable, y por tanto prohibidos, aquellos sistemas o prácticas de inteligencia artificial que contravengan los valores y derechos fundamentales de la Unión Europea. Estas prácticas no pueden introducirse en el mercado, ponerse en servicio ni utilizarse en la Unión Europea, salvo las **excepciones** previstas para casos muy concretos.

Las practicas prohibidos incluyen las siguientes:

Práctica prohibida	Descripción	Qué abarca en la práctica	Excepción expresa (si aplica)	Referencia
Manipulación subliminal o deliberadamente manipuladora/engañosa que altere sustancialmente el comportamiento y cause, o pueda causar razonablemente, perjuicios considerables	Prohibida la comercialización, puesta en servicio o uso de IA que utilice técnicas subliminales o manipuladoras para alterar de forma sustancial el comportamiento, mermando la capacidad de decisión informada, con perjuicio considerable real o razonablemente probable	Interfaces o técnicas que induzcan decisiones no informadas con perjuicio considerable	No se prevé excepción sectorial; se evalúa el umbral de *alteración sustancial* y *perjuicio considerable*	Art. 5. 1.a
Explotación de vulnerabilidades de personas o colectivos por edad, discapacidad o situación social/económica para alterar sustancialmente su comportamiento con perjuicio considerable	Prohibida la IA que explote dichas vulnerabilidades con ese efecto nocivo	Cualquier diseño o funcionalidad que se aproveche de la vulnerabilidad para inducir conductas perjudiciales	No se prevé excepción sectorial	Art. 5.1. b
Puntuación social con trato perjudicial en contextos no relacionados, o injustificado o desproporcionado	Prohibidos los sistemas que evalúan o clasifican personas por comportamiento social o rasgos, cuando la puntuación derive en los efectos descritos	Sistemas de reputación o puntuación cívica que condicionen acceso a servicios, beneficios o	No se prevé excepción sectorial	Art. 5.1.c

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

Práctica prohibida	Descripción	Qué abarca en la práctica	Excepción expresa (si aplica)	Referencia
		trato institucional		
Evaluación o predicción del riesgo de delinquir basándose únicamente en perfilado o rasgos/rasgos de personalidad	Prohibido el uso de IA para predecir la criminalidad sobre esa base	Puntación social sin hechos objetivos vinculados a una actividad delictiva concreta	Permitido cuando solo apoya una valoración humana basada en hechos objetivos y verificables directamente relacionados con una actividad delictiva	Art. 5.1. d
Creación o ampliación de bases de datos de reconocimiento facial mediante “scraping” no selectivo de internet o CCTV	Prohibida la extracción indiscriminada de imágenes faciales para construir o ampliar bases de datos	Recolección masiva y no selectiva de rostros desde fuentes abiertas o CCTV	No se prevé excepción general	Art. 5.1. e
Inferencia de emociones en trabajo o educación	Prohibido inferir emociones en lugares de trabajo y centros educativos	Análisis de rostro, voz o fisiología para deducir estado emocional en estos entornos	Permitido excepcionalmente si el sistema está destinado a fines médicos o de seguridad	Art. 5.1. f
Categorización biométrica para inferir atributos sensibles (raza, opiniones políticas, religión/creencias, afiliación sindical, vida u orientación sexual)	Prohibidos los sistemas que clasifiquen individualmente para deducir esos atributos a partir de datos biométricos	Modelos que etiquetan personas por atributos sensibles a partir de biometría	No se prohíbe el etiquetado/filtrado lícito de conjuntos de datos biométricos ni la categorización de biometría en garantía del cumplimiento del Derecho	Art. 5.1. g
Identificación biométrica remota “en tiempo real” en espacios públicos con fines de garantía del cumplimiento del Derecho	Prohibido el uso salvo que sea **estrictamente necesario** y para fines tasados: búsqueda selectiva de víctimas o desaparecidos; prevención de	Uso policial de identificación biométrica en vivo y en espacios de	Excepción cerrada y sometida a estricta necesidad y proporcionalidad. Delitos graves	Art. 5.1.h y Anexo II

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

Práctica prohibida	Descripción	Qué abarca en la práctica	Excepción expresa (si aplica)	Referencia
	amenaza específica e inminente a la vida o seguridad física, o amenaza real y actual o previsible de atentado terrorista; localización e identificación de un sospechoso de delitos graves	acceso público	referenciados en **Anexo II**	

Tabla 1: Prácticas prohibidas

6.1.1.1. Interpretaciones claves:

Para una correcta interpretación del artículo 5 del RIA, conviene precisar los siguientes aspectos:

- **Umbral de "alteración sustancial" y "perjuicio considerable":** En los supuestos contemplados, la prohibición no alcanza a cualquier técnica de persuasión o diseño persuasivo de interfaces (UX). El RIA establece que debe existir una distorsión significativa de la capacidad de decisión de la persona que, además, cause o pueda causar razonablemente un perjuicio considerable a individuos o colectivos. Ambos elementos deben concurrir de manera acumulativa.
- **Sistemas de puntuación social**: La prohibición de puntación social no se limita al ámbito de las autoridades públicas. Lo determinante es que la puntuación genere un trato perjudicial en contextos distintos de aquellos para los que se recogieron los datos, o que produzca consecuencias injustificadas o desproporcionadas respecto del comportamiento valorado, especialmente cuando se prolonga durante un periodo de tiempo determinado.
- **Predicción de criminalidad:** La prohibición es aplicable cuando la predicción se basa de forma exclusiva en el perfilado o en rasgos de personalidad. No obstante, el RIA permite que los sistemas de IA apoyen una valoración humana siempre que esta se fundamente en hechos objetivos, verificables y directamente relacionados con una actividad delictiva concreta. La clave diferenciadora es la existencia de una base fáctica sólida y comprobable.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

- **Extracción facial no selectiva ("scraping"):** La construcción o ampliación de bases de datos de reconocimiento facial a partir de la recolección indiscriminada de imágenes en internet o de sistemas de CCTV queda prohibida de forma absoluta. El Reglamento no contempla excepciones de uso legítimo en este ámbito.
- **Inferencia de emociones:** El veto se aplica específicamente en entornos de trabajo y de educación. Se admite, sin embargo, la excepción para sistemas cuyo diseño y finalidad estén orientados a usos médicos o de seguridad. Es insuficiente una mera declaración comercial de finalidad médica; el sistema debe estar objetivamente concebido y validado para ese propósito.
- **Categorización biométrica de atributos sensibles:** La prohibición alcanza la deducción de atributos sensibles (como raza, orientación sexual u opiniones políticas) a partir de datos biométricos en la identificación de personas concretas. El Reglamento distingue esta práctica de otras actividades permitidas, como el etiquetado o filtrado de conjuntos de datos biométricos adquiridos de forma lícita con fines de desarrollo, o de las aplicaciones enmarcadas en actuaciones de garantía del cumplimiento del Derecho.
- **Identificación biométrica en tiempo real con fines policiales**: La regla general es la prohibición del uso de sistemas de identificación biométrica en vivo en espacios de acceso público. Únicamente se permite bajo un régimen de estricta necesidad y para finalidades tasadas: búsqueda de víctimas de delitos o desaparecidos, prevención de amenazas específicas e inminentes a la vida o a la seguridad física, o localización e identificación de sospechosos de delitos graves expresamente recogidos en el Anexo II.

Cualquier excepción deberá estar justificada, documentada y sometida a control de la autoridad competente.

6.1.2. Sistemas de alto riesgo

El artículo 6 del Reglamento (UE) 2024/1689 (RIA) establece que determinados sistemas de inteligencia artificial se consideran de alto riesgo debido a su potencial de afectar de manera significativa a la salud, la seguridad o los derechos fundamentales de las personas.

La base jurídica principal se encuentra en el **artículo 6**, que prevé dos vías de clasificación:

- **Anexo I:** cuando el sistema de IA es **un componente de seguridad** de un producto regulado por legislación de la Unión (actos de armonización) que requiere evaluación de conformidad por terceros antes de ser introducido en el mercado (incluyendo los **productos sanitarios** y **diagnóstico in vitro** en base al MDR/IVDR).

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA SECRETARÍA DE ESTADO DE TELECOMUNICACIONES E INFRAESTRUCTURAS DIGITALES
Plan de Recuperación, Transformación y Resiliencia
incibe_ INSTITUTO NACIONAL DE CIBERSEGURIDAD
Junta de Andalucía
Agencia Digital de Andalucía

- **Anexo III:** cuando el sistema de IA se utiliza en **ámbitos identificados como críticos** por su impacto en personas y colectivos, listados de forma exhaustiva en el Anexo III.

Adicionalmente, el **artículo 6.3** del RIA permite **modular la clasificación** de riesgo en casos justificados. Un sistema incluido formalmente en alguno de los ámbitos de los anexos I o III puede no considerarse de alto riesgo si su función es exclusivamente instrumental o preparatoria, si no produce efectos sustantivos sobre personas físicas y si esta circunstancia se documenta adecuadamente en el expediente técnico.

6.1.2.1. Sistemas recogidos en el Anexo I:

El **Anexo I** del RIA contiene una lista de productos ya regulados por actos de armonización de la Unión Europea. Cuando un sistema de IA se integra en estos productos **como componente de seguridad**, pasa a considerarse de alto riesgo **siempre que el producto requiera evaluación de conformidad por terceros.**

Cuando el sistema de IA constituye un producto sanitario (MDR) o un producto sanitario para diagnóstico in vitro (IVDR), debe seguirse la siguiente secuencia:

I. **Clasificación según MDR/IVDR:** Determinar la clase del producto conforme a las reglas del Anexo VIII del MDR o del IVDR.
II. **Evaluación de conformidad y marcado CE:** Completar el procedimiento de evaluación de conformidad correspondiente, incluyendo, cuando proceda, la intervención de organismo notificado, y obtener el marcado CE y el identificador único de producto (UDI).
III. **Clasificación complementaria según RIA:** Una vez cumplidos los pasos anteriores, evaluar si el sistema constituye además un sistema de IA de alto riesgo conforme al artículo 6 del RIA y los Anexos I y/o III.

La aplicación del RIA es complementaria y no sustitutiva de las obligaciones establecidas en el MDR y en el IVDR. Ambos marcos regulatorios deben cumplirse de forma coordinada.

Lista completa del Anexo I:

- Juguetes.
- Maquinaria.
- Ascensores y componentes de seguridad para ascensores.
- Equipos de protección personal.

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA SECRETARÍA DE ESTADO DE TELECOMUNICACIONES E INFRAESTRUCTURAS DIGITALES
Plan de Recuperación, Transformación y Resiliencia
incibe_ INSTITUTO NACIONAL DE CIBERSEGURIDAD
Junta de Andalucía
Agencia Digital de Andalucía

- Dispositivos médicos.
- Productos sanitarios para diagnóstico in vitro.
- Vehículos de motor y sus sistemas de seguridad.
- Aeronaves civiles.
- Sistemas ferroviarios.
- Buques y equipos marinos.
- Equipos de presión.

Un sistema de IA integrado en un producto del Anexo I **solo será de alto riesgo** cuando:

Criterio	Explicación	Ejemplo en Salud	Ejemplo en Smart Cities
Producto regulado por actos de armonización de la UE	El sistema de IA forma parte de un producto sujeto a legislación sectorial de la UE que exige certificación previa.	Software de IA integrado en un dispositivo médico de diagnóstico por imagen regulado por el MDR.	Algoritmo de control de seguridad integrado en trenes automatizados bajo normativa ferroviaria.
Evaluación de conformidad por tercero	El producto no puede comercializarse sin certificación CE de un organismo notificado.	IA que controla la dosificación automática en una bomba de insulina. (producto sanitario Clase IIb bajo MDR)	Sistema de IA de control de apertura de puertas en metros automatizados.

Tabla 2: Ejemplo de clasificación de riesgo alto

6.1.2.2. Sistemas recogidos en el Anexo III:

El Anexo III del RIA identifica ámbitos de especial sensibilidad en los que el uso de IA se presume de alto riesgo debido a su potencial impacto en los derechos fundamentales, la seguridad o la equidad social.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

En el **ámbito sanitario**, los sistemas de IA destinados a la salud y seguridad de personas físicas, incluyendo aquellos que determinan el acceso a prestaciones sanitarias, priorizan tratamientos o gestionan infraestructuras hospitalarias esenciales, quedan comprendidos en el Anexo III del RIA bajo la categoría de servicios esenciales e infraestructuras críticas. Cuando estos sistemas constituyan además productos sanitarios (MDR) o productos de diagnóstico in vitro (IVDR), la clasificación de alto riesgo se deriva tanto del Anexo I como del Anexo III del RIA, reforzando su carácter crítico y la necesidad de aplicación convergente de ambos marcos regulatorios.

En **Smart Cities**, se incluye todos los sistemas de IA utilizados para el control de tráfico, gestión energética, reconocimiento facial en espacios públicos o asignación de recursos urbanos.

Lista completa del Anexo III:

- Identificación biométrica remota y categorización de personas físicas.
- Gestión y funcionamiento de infraestructuras críticas.
- Educación y formación profesional.
- Empleo, gestión laboral y acceso al autoempleo.
- Acceso y disfrute de servicios esenciales (públicos y privados).
- Aplicación de la ley.
- Migración, asilo y gestión de fronteras.
- Administración de justicia y procesos democráticos.

Área crítica	Descripción	Ejemplo en Salud	Ejemplo en Smart Cities
Biometría	Identificación remota, categorización y reconocimiento de emociones.	Sistema de reconocimiento facial remoto para detección de estados emocionales en pacientes con enfermedades mentales.	Identificación biométrica remota en tiempo real en espacios públicos
Infraestructuras críticas	Sistemas que controlan servicios esenciales.	IA que controla el suministro eléctrico y oxígeno en hospitales (puede constituir además producto sanitario bajo MDR si	Algoritmo que regula el tráfico en túneles urbanos.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

Área crítica	Descripción	Ejemplo en Salud	Ejemplo en Smart Cities
		cumple finalidad médica).	
Educación y formación	Evaluación y asignación educativa.	IA que califica exámenes de medicina.	Plataforma que selecciona candidatos a formación municipal.
Empleo y gestión laboral	Procesos de contratación, asignación o despido.	Motor de cribado automático de CV de enfermería.	Algoritmo que asigna turnos a operarios de transporte.
Servicios esenciales	Acceso a sanidad, vivienda, energía o ayudas sociales.	IA que prioriza pacientes en listas de trasplante (sistema de alto riesgo bajo Anexo III; si incorpora funcionalidad de diagnóstico o pronóstico, puede constituir además producto sanitario bajo MDR).	Sistema que asigna acceso a viviendas sociales.
Aplicación de la ley	Investigación, detección y prevención del delito.	Herramienta que analiza historiales médicos en procesos judiciales.	Sistema de videovigilancia con IA en espacios urbanos.
Migración y asilo	Procesos de residencia o control fronterizo.	IA que analiza documentación sanitaria en solicitudes de asilo.	Algoritmo que gestiona flujos migratorios en aeropuertos.
Justicia y democracia	Decisiones judiciales o gestión electoral.	Asistente que apoya jueces en valoración de pruebas médicas.	Herramienta de recuento y validación de votos electrónicos.

Tabla 3: Anexo III del RIA

6.1.2.3. Exclusiones y modulaciones (artículo 6.3)

Dentro de los ámbitos del Anexo III, un sistema puede quedar **excluido de la categoría de alto riesgo** cuando no plantea un riesgo significativo ni influencia material en las decisiones que afecten a personas físicas. La exclusión solo es válida si se documenta y justifica previamente (art. 49.2 del RIA).

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

Situación	Explicación	Ejemplo
Tareas de procedimiento limitadas	El sistema solo ejecuta funciones instrumentales sin impacto decisivo. Es decir, no se toma una decisión para llevar a cabo una acción.	Clasificador automático de documentos administrativos en un hospital.
Mejora de eficiencia sin sustituir decisiones humanas	El sistema optimiza procesos, pero la decisión sustantiva la mantiene el humano.	Asistente de rutas para la gestión de tráfico urbano.
Detección de patrones sin influencia directa	El sistema detecta anomalías, pero la decisión final recae en un profesional.	Detección de picos de consumo eléctrico en un hospital bajo supervisión técnica.
Función preparatoria sin impacto directo*	IA que prepara datos sin afectar directamente a personas.	Preprocesamiento de imágenes médicas para mejorar calidad visual.

Tabla 4: Exclusiones y modulaciones de riesgo alto

***Nota:** la exclusión por carácter instrumental/preparatorio en el RIA no exime de MDR/IVDR si el sistema es producto sanitario/IVD. Además, los sistemas que realicen **elaboración de perfiles de personas físicas** no pueden beneficiarse de esta modulación.*

6.1.2.4. Interpretaciones claves para la clasificación de alto riesgo:

- **Evaluación caso por caso:** la inclusión en el Anexo III no implica automáticamente alto riesgo; es necesario comprobar si concurren las exclusiones del artículo 6.3. del RIA.
- **Impacto material:** el criterio decisivo es si el sistema puede afectar de forma significativa la salud, la seguridad o los derechos de las personas.
- **Documentación de exclusiones:** toda decisión de no clasificar un sistema como de alto riesgo debe justificarse y conservar trazabilidad documental.
- **Enfoque sectorial:** en salud y Smart Cities la mayoría de los sistemas recaen, directa o indirectamente, en ámbitos del Anexo III. Se recomienda una evaluación exhaustiva del caso de uso de cada sistema en uso para evitar riesgos regulatorios y sanciones.

Financiado por la Unión Europea NextGenerationEU

- **Coherencia normativa:** en el ámbito sanitario, la clasificación de alto riesgo debe aplicarse de forma coordinada con las obligaciones establecidas en el MDR y en el IVDR. Esto implica que la evaluación de conformidad y el marcado CE constituyen pasos obligatorios previos, mientras que el RIA introduce requisitos adicionales orientados a la protección de los derechos fundamentales y a la gestión proporcional del riesgo. En el ámbito sanitario, la clasificación de alto riesgo debe aplicarse de forma coordinada con las obligaciones establecidas en el MDR y en el IVDR.
- **Acumulación de obligaciones:** Un sistema clasificado como de alto riesgo puede estar simultáneamente sujeto a las obligaciones de transparencia del artículo 50, cuando interactúe con personas físicas, genere contenido sintético o aplique reconocimiento de emociones o categorización biométrica. Estas obligaciones son acumulativas y deben consignarse en la documentación técnica.

6.1.3. Sistemas de riesgo limitado

El Reglamento (UE) 2024/1689 reconoce la existencia de determinados sistemas de inteligencia artificial cuyo potencial de riesgo es moderado, en la medida en que no afectan directamente a la seguridad ni se encuadran en ámbitos críticos, pero sí pueden generar riesgos de desinformación, manipulación o afectación indirecta de derechos fundamentales si no se gestionan con transparencia.

Esta categoría no está asociada a sectores críticos como los enumerados en el Anexo III, ni a productos regulados por actos de armonización de la Unión como los mencionados en el Anexo I, sino a situaciones de interacción con personas físicas o de generación de contenidos sintéticos en las que la falta de identificación clara podría inducir a error.

Esta clasificación no se determina por la complejidad tecnológica del sistema, sino por su capacidad de inducir a error o de alterar la percepción de los usuarios cuando no se comunica adecuadamente su carácter artificial.

Por ello, el Reglamento europeo establece en su artículo 50 una serie de obligaciones específicas de transparencia destinadas a garantizar que los usuarios puedan identificar con claridad cuándo interactúan con un sistema de inteligencia artificial o cuándo están expuestos a contenidos generados o modificados artificialmente.

El **artículo 50** del Reglamento establece que un sistema de IA es de riesgo limitado cuando:

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

- Un sistema interactúa con personas físicas, estas deben ser informadas de que interactúan con una IA, salvo que ello sea evidente para una persona razonablemente informada, atenta y perspicaz, atendiendo al contexto de uso.

- Los sistemas que generan contenido sintético de audio, imagen, vídeo o texto, salvo cuando tiene funciones de edición estándar o que no alteren sustancialmente la semántica del contenido facilitado por el usuario.

- Un sistema cuyo objetivo es el reconocimiento de emociones o categorización biométrica basados en el tratamiento de datos biométricos.

- Un sistema que genere o manipule suplantaciones (deepfakes) de imagen, audio o vídeo. Para texto destinado a informar al público sobre asuntos de interés público, debe divulgarse que ha sido generado o manipulado artificialmente, salvo que haya revisión humana o control editorial con responsabilidad editorial.

Ejemplos de caso de uso según las casuísticas en los entornos de Salud y Smart Cities:

Criterio	Ejemplo en Salud	Ejemplo en Smart Cities
Interacción con personas	Chatbot hospitalario que atiende consultas iniciales de pacientes.	Asistente virtual en la web municipal que resuelve dudas sobre impuestos o transporte.
Reconocimiento de emociones	Sistema que analiza expresiones faciales en terapias psicológicas con fines médicos.	Herramienta que detecta frustración en usuarios de oficinas municipales para redirigirlos a un canal de atención más adecuado.
Categorización biométrica	Aplicación que utiliza reconocimiento facial para agilizar el registro voluntario de pacientes en centros de vacunación, sin clasificar ni priorizar por rasgos biométricos.	Algoritmo que categoriza peatones por edad aproximada para mejorar la planificación de semáforos inteligentes.
Generación de contenidos sintéticos	Creación de vídeos de formación médica utilizando modelos sintéticos en lugar de casos reales.	Producción de simulaciones urbanísticas en 3D generadas por IA para presentar proyectos de movilidad.

Tabla 5: Casos de uso de riesgo limitado

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

6.1.4. Sistemas de riesgo mínimo

El Reglamento (UE) 2024/1689 reconoce la existencia de sistemas de inteligencia artificial cuyo nivel de riesgo es mínimo o inexistente. Se trata de aplicaciones que, por su naturaleza, no afectan a la seguridad, la salud ni a los derechos fundamentales de las personas, y cuyo impacto se limita a funciones auxiliares, recreativas o administrativas.

A efectos prácticos, la presente guía incorpora esta categoría con finalidad orientativa, destinada a identificar aquellas aplicaciones de uso generalista, auxiliar o recreativo que no generan consecuencias jurídicas ni materiales significativas y pueden desplegarse sin requisitos regulatorios específicos adicionales.

Los sistemas que se consideren como riesgo mínimo deben cumplir con al menos una de las siguientes características:

- **No generan consecuencias significativas**: su funcionamiento no produce efectos jurídicos ni prácticos relevantes sobre individuos o colectivos.
- **Uso generalista y cotidiano:** predominan aplicaciones de apoyo, administrativos, entretenimiento o mejora de experiencia de usuario.

Nota: un sistema inicialmente considerado de riesgo mínimo puede reclasificarse como riesgo limitado o alto riesgo si se amplía su funcionalidad hacia ámbitos sensibles.

A continuación, se proporciona unos ejemplos de casos de uso en los entornos de Salud y Smart Cities que serán clasificados como **riesgo mínimo**:

Ejemplo en Salud	Ejemplo en Smart Cities
Aplicación móvil que sugiere rutinas de bienestar sin efectos médicos vinculantes.	Asistente de recomendación turística que sugiere rutas culturales.
Herramienta de IA que corrige automáticamente la ortografía en historiales médicos sin alterar su contenido sustantivo.	Sistema que traduce automáticamente avisos municipales a diferentes idiomas.
Aplicación de recordatorios de medicación basada en temporizadores, sin acceso a datos clínicos ni decisiones automatizadas.	Chatbot informativo que responde preguntas frecuentes sobre horarios de transporte.

Tabla 6: Caso de uso riesgo mínimo

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
Plan de Recuperación, Transformación y Resiliencia
incibe_ INSTITUTO NACIONAL DE CIBERSEGURIDAD
Junta de Andalucía
Agencia Digital de Andalucía

6.2. Modelos y sistemas de IA de propósito general (GPAI)

El Reglamento (UE) 2024/1689 introduce la categoría de modelos de inteligencia artificial de propósito general (GPAI, por sus siglas en inglés), definida en el artículo 3, como aquellos modelos capaces de desempeñar múltiples funciones en una amplia variedad de aplicaciones. A diferencia de los sistemas diseñados para un caso de uso específico, los GPAI se caracterizan por su versatilidad y transversalidad, lo que les permite ser reutilizados en sectores muy diversos, desde la salud hasta la gestión urbana.

Estos modelos no constituyen, por sí mismos, aplicaciones finales, sino que representan una base tecnológica generalista que puede integrarse en distintos sistemas de IA. En función de esa integración y del contexto de uso, los sistemas resultantes podrán clasificarse en cualquiera de las categorías de riesgo previstas en el Reglamento: inaceptable, alto, limitado o mínimo.

Un modelo de IA de propósito general (GPAI) se reconoce por las siguientes características:

- **Generalidad funcional:** capacidad para ejecutar de manera competente una amplia variedad de tareas diferenciadas.
- **Entrenamiento a gran escala:** habitualmente entrenados con volúmenes masivos de datos utilizando técnicas como aprendizaje supervisado, no supervisado o por refuerzo.
- **Disponibilidad reutilizable**: se ponen a disposición de terceros mediante bibliotecas de software, interfaces de programación de aplicaciones (API), descargas digitales o copias físicas, lo que facilita su integración en aplicaciones heterogéneas.

Además, conviene diferenciar claramente entre modelo y sistema:

- ✓ Un modelo GPAI constituye la tecnología de base.
  Ejemplo: un modelo multimodal entrenado en texto, imagen y sonido, con capacidad de generar y analizar contenidos en diferentes formatos.
- ✓ Un sistema de IA basado en GPAI surge cuando ese modelo se combina con otros componentes funcionales (interfaces de usuario, módulos de control, infraestructura de despliegue, conexión con datos de un sector específico) para dar lugar a una aplicación concreta.
  Ejemplo: un chatbot hospitalario que emplea un modelo lingüístico generalista para responder preguntas médicas de pacientes.

Esta diferenciación entre modelo y sistema resulta crítica para la clasificación regulatoria, ya que un mismo modelo GPAI puede generar sistemas con clasificaciones de riesgo muy distintas según su integración:

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

- En entornos críticos como hospitales o infraestructuras urbanas, puede situarse en la categoría de alto riesgo.
- En interacciones simples con usuarios, puede quedar en la categoría de riesgo limitado.
- En usos generales sin impacto relevante, puede considerarse incluso de riesgo mínimo.

6.2.1. Indicadores para su clasificación como GPAI

Para identificar y clasificar un modelo como GPAI, es necesario analizar ciertos rasgos distintivos:

- **Amplitud funcional:** el modelo ejecuta con solvencia tareas heterogéneas (procesamiento de lenguaje natural, análisis de imágenes, generación de contenidos, razonamiento básico, etc.).
- **Versatilidad de dominios:** puede aplicarse en contextos muy diversos y transferirse a nuevos entornos con ajustes mínimos.
- **Escala de entrenamiento**: se entrena en grandes corpus de datos, lo que refuerza su carácter generalista (aunque no lo determina de manera exclusiva).
- **Transferibilidad**: presenta capacidad de adaptación a tareas no previstas originalmente, sin necesidad de rediseño sustantivo.
- Forma de entrega: se distribuye como recurso reutilizable para terceros, en lugar de estar limitado a un único caso de uso cerrado.

A continuación, se proporciona unos ejemplos en los entornos de Salud y Smart Cities que serán clasificados como **GPAI**:

	GPAI	NO GPAI
Ejemplo en Salud	Modelo fundacional biomédico entrenado en textos médicos, imágenes y señales fisiológicas que puede redactar informes, responder consultas clínicas y realizar tareas de apoyo administrativo.	Red neuronal entrenada exclusivamente para segmentar imágenes de un órgano concreto en radiología.
Ejemplo en Smart Cities	Modelo multimodal capaz de analizar patrones de movilidad, redactar comunicaciones municipales, generar simulaciones urbanísticas y responder consultas de los ciudadanos.	Algoritmo entrenado únicamente para detectar ocupación en plazas de aparcamiento de una calle concreta.

Tabla 7: Clasificación GPAI

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

Nota: la clasificación como modelo GPAI no implica por sí misma un nivel de riesgo determinado. Este se evaluará conforme al uso específico y al entorno de integración del sistema resultante, siguiendo la metodología descrita en el apartado 6.1.

6.2.2. Modelos y sistemas de IA de propósito general (GPAI) con riesgo sistémico

Dentro de la categoría de modelos GPAI, el Reglamento introduce una subcategoría específica: los modelos de propósito general con riesgo sistémico, definidos en el artículo 52 y en el Anexo XIII. Se consideran así aquellos modelos cuyas capacidades o efectos potenciales alcanzan un nivel que puede afectar de forma significativa y transversal a múltiples sectores o a la sociedad en su conjunto.

La diferencia esencial respecto a un GPAI "no sistémico" es que, además de ser generalista, el modelo alcanza un umbral de capacidades de gran impacto que incrementa de manera sustancial la probabilidad de generar riesgos extendidos.

Existen dos mecanismos principales para determinar si un GPAI debe clasificarse como de **riesgo sistémico**:

1. **Capacidades de gran impacto**: cuando evaluaciones técnicas y metodologías reconocidas muestran que el modelo presenta un nivel de rendimiento o de comportamiento emergente que puede ocasionar repercusiones significativas.
2. **Designación por parte de la Comisión Europea**: tras análisis especializado y siguiendo los criterios establecidos en el anexo XIII del Reglamento, la Comisión puede declarar un modelo como GPAI con riesgo sistémico.

Adicionalmente, el artículo 51 establece una **presunción automática** en la que se presume que un GPAI tiene capacidades de gran impacto y, por tanto, riesgo sistémico, cuando la cantidad total de cálculo utilizada en su entrenamiento supera 10^25 operaciones de coma flotante (FLOPs). Este umbral podrá actualizarse periódicamente mediante actos delegados de la Comisión para reflejar la evolución tecnológica.

Los principales elementos prácticos de diferenciación:

- Un **GPAI estándar** es generalista y reutilizable, pero sus capacidades no alcanzan el umbral de gran impacto.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

- Un **GPAI con riesgo sistémico** es generalista y, además, cumple las condiciones técnicas o de designación que lo sitúan como potencialmente transformador y eleva el riesgo a escala de ecosistema.

Implicaciones en salud y Smart Cities:

- Si el modelo GPAI con riesgo sistémico se integra en aplicaciones críticas (por ejemplo, triaje sanitario, control de infraestructuras hospitalarias o gestión inteligente del tráfico urbano), el sistema resultante deberá evaluarse conforme a los criterios de alto riesgo del artículo 6 y Anexo III.
- En interacciones no críticas o de generación de contenido sintético (información ciudadana o apoyo administrativo), podrá situarse en riesgo limitado, aplicando las obligaciones de transparencia del artículo 50.
- En usos residuales o auxiliares, podrá considerarse de riesgo mínimo, sin perjuicio de que el proveedor del modelo deba cumplir las obligaciones reforzadas de documentación, pruebas de seguridad y notificación de incidentes previstas en los artículos 53 y 55.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

7. METODOLOGÍA PARA LA CLASIFICACIÓN DE SISTEMAS DE IA EN LA UE

Este apartado detalla una metodología completa de cómo se debe ejecutar una correcta clasificación de sistemas de IA según el Reglamento. Este capítulo detalla la metodología completa para realizar la correcta clasificación de los sistemas de inteligencia artificial de acuerdo con lo establecido en el Reglamento (UE) 2024/1689 sobre inteligencia artificial (RIA).

La metodología tiene como finalidad a garantizar que los proveedores, usuarios profesionales y autoridades competentes puedan aplicar un enfoque trazable, coherente y verificable, que permita demostrar el cumplimiento de las obligaciones previstas en el RIA.

7.1. Realizar un inventario de sistemas

Como primer paso del proceso de clasificación, se debe realizar una revisión de todos los sistemas y casos de uso, categorizándolos en un inventario que incluya: descripción de funcionalidades, propósito previsto, contexto de uso y tecnología subyacente. Esta información determina si el sistema constituye IA según el Reglamento y requiere cumplimiento con el RIA.

7.1.1. Entendimiento y análisis del propósito y contexto

Antes de clasificar un sistema, debe confirmarse que se encuentra dentro del ámbito material y territorial de aplicación del RIA, conforme a su artículo 2. El análisis debe quedar registrado en un inventario interno de sistemas de IA, incluyendo como mínimo:

- **Propósito funcional y** decisiones que habilita o influye (qué tareas, qué efectos produce).
- **GPAI (modelo/sistema de propósito general) o no:**
    - ¿Es un modelo generalista reutilizable para tareas variadas o un sistema específico de caso de uso? (véase definición en el apartado 6.2 y criterios de identificación en el apartado 6.2.1)
- **Usuarios previstos:** operadores profesionales, personal sanitario, policía local, técnicos de movilidad, ciudadanos/pacientes.
- **Entorno de uso:** hospital, centro de salud, red de transporte, gestión de residuos, alumbrado, centros de control urbano.
- **Tecnología implicada:** fuentes de datos (clínicos, sensores, cámaras, IoT), técnicas (ML clásico, DL, RL, LLM, modelos multimodales), canal (API, app, dispositivo médico, plataforma nube).

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

- **Interacciones relevantes:** contacto directo con personas, generación de contenidos sintéticos, biometría, reconocimiento de emociones, funciones de seguridad de productos.

7.1.2. Verificación de excepciones de aplicación del RIA

Una vez recopilada la información relevante del sistema, se debe comprobar si este se encuentra dentro del ámbito de aplicación del Reglamento, teniendo en cuenta las excepciones en el artículo 2. Con el fin de determinar si un sistema de IA entre en las excepciones del RIA, responde a las siguientes preguntas:

N.º	Criterio de verificación	Explicación técnica	Respuesta (Sí/No)	Observaciones
1	¿El sistema se utiliza exclusivamente para fines de defensa, seguridad nacional o actividades militares?	El RIA excluye de su aplicación los sistemas de IA destinados a operaciones militares o de seguridad nacional.		Si: el sistema queda excluido del RIA y no procede su clasificación. No: continúa en el proceso de clasificación.
2	¿El sistema está en fase de investigación o desarrollo, sin ser puesto en servicio ni comercializado?	Los prototipos experimentales o proyectos de I+D no están sujetos al RIA hasta que se despliegan o introducen en el mercado.		Si: el sistema queda excluido del RIA hasta su puesta en servicio o comercialización. No: procede la clasificación regulatoria.
3	¿El sistema se utiliza exclusivamente en el marco de cooperación policial o judicial internacional bajo acuerdos específicos?	Determinadas actividades de cooperación internacional quedan fuera del ámbito del RIA si se regulan mediante tratados o convenios.		Si: el sistema queda excluido del RIA bajo el marco del acuerdo internacional. No: sigue sujeto al RIA y se debe clasificar.

Tabla 8: Cuestionario de verificación de aplicación

En el caso de responder “No” a todas las preguntas anteriores, el sistema resulta sujeta a la clasificación de nivel de riesgo por lo que se debería seguir al punto 7.2.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

7.2. Clasificación en dos fases

La clasificación de los sistemas de IA requiere un doble enfoque:

- Una primera fase, destinada a la determinación del nivel de riesgo (prohibido, alto, limitado o mínimo, además de GPAI/GPAI sistémico);
- Una segunda fase, enfocada en las obligaciones de transparencia aplicables.

7.2.1. Evaluación del nivel del riesgo

Los sistemas o prácticas prohibidas descritos en el artículo 5 generan un riesgo incompatible con los derechos fundamentales de la Unión y no pueden ser comercializados ni utilizados en la UE, salvo las excepciones expresas. Si el sistema encaja en alguno de estos supuestos, queda prohibida su comercialización, puesta en servicio o utilización en la Unión Europea Prácticas prohibidas.

Con el fin de garantizar una verificación exhaustiva, se propone el siguiente cuestionario estructurado, que cubre todas las tipologías de prohibición previstas:

N.º	Pregunta de verificación	Referencia legal	Si "Sí"	Si "No"
1	¿El sistema utiliza técnicas subliminales o engañosas que alteran significativamente el comportamiento de una persona de modo que pueda causarle daño físico o psicológico considerable?	Art. 5. 1.a	Clasificar como prohibido. Detener implantación.	Continuar al punto 2.
2	¿El sistema explota vulnerabilidades de personas por edad, discapacidad o situación social para influir en sus decisiones de manera que pueda causarles perjuicio?	Art. 5.1. b	Clasificar como prohibido.	Continuar.
3	¿El sistema aplica puntuación social que derive en trato perjudicial o desproporcionado en un contexto no relacionado?	Art. 5.1.c	Clasificar como prohibido.	Continuar.
4	¿Realiza identificación biométrica remota en tiempo real en espacios de acceso público fuera de las excepciones legales?	Art. 5.1. d	Clasificar como prohibido.	Continuar.
5	¿Realiza identificación biométrica remota "a posteriori" en espacios	Art. 5.1. d	Clasificar como prohibido.	Continuar.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

N.º	Pregunta de verificación	Referencia legal	Si "Sí"	Si "No"
	públicos sin base legal estricta (ej. autorización judicial)?			
6	¿Crea o amplía bases de datos de reconocimiento facial mediante ¨scraping¨ o captura indiscriminada de imágenes?	Art. 5.1. e	Clasificar como prohibido.	Continuar.
7	¿Emplea reconocimiento de emociones en entornos laborales o educativos para evaluar a trabajadores, estudiantes o candidatos?	Art. 5.1. f	Clasificar como prohibido.	Continuar.
8	¿Clasifica datos biométricos con el fin de inferir orientación sexual, raza, opiniones políticas, creencias religiosas o filosóficas?	Art. 5.1. g	Clasificar como prohibido.	Continuar.
9	¿Realiza predicción de probabilidad de cometer delitos por personas físicas?	Art. 5.1.h	Clasificar como prohibido.	Continuar.
10	¿Se utiliza para manipular el comportamiento electoral de votantes a través de campañas de desinformación dirigida a gran escala?	Art. 5.1. i	Clasificar como prohibido.	Continuar.
11	¿Genera "deepfakes" o contenidos sintéticos que puedan confundirse con material auténtico, sin una divulgación clara de su carácter artificial?	Art. 5.1.j (y art. 50)	Clasificar como prohibido.	Continuar.
12	¿El sistema explota vulnerabilidades de personas por edad, discapacidad o situación social para influir en sus decisiones de manera que pueda causarles perjuicio?	Art. 5.1. b	Clasificar como prohibido.	Continuar al 7.2.1.1.

Tabla 9: Cuestionario riesgo prohibido

7.2.1.1. Sistemas de alto riesgo

Este apartado explicará cómo identificar y evaluar un sistema de IA para determinar si debe clasificarse como de alto riesgo según los criterios establecidos en el Reglamento. Se detallarán los factores clave a analizar, incluyendo el sector de aplicación, el propósito del sistema y su potencial impacto en la seguridad, salud y derechos fundamentales.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

7.2.1.1.1. EN FUNCIÓN DE LOS ANEXOS

El Reglamento (UE) 2024/1689 define en sus anexos I y III los supuestos en los que un sistema de IA debe clasificarse como de alto riesgo.

En estos casos, la clasificación como alto riesgo es automática, salvo que pueda aplicarse la excepción del artículo 6.3, que excluye sistemas de carácter meramente instrumental, preparatorio o de eficiencia, siempre que no produzcan efectos sustantivos en las personas y que esta exclusión se documente con rigor.

A continuación, se presenta un cuestionario de clasificación por riesgo alto:

N.º	Pregunta de verificación	Referencia legal	Si "Sí"	Si "No"
1	¿Es un componente de seguridad de un producto, o es él mismo un producto, cubierto por la legislación de armonización de la Unión enumerada en el Anexo I?	Art. 6.1 + Anexo I	Alto riesgo vinculado al producto.	Continuar.
2	¿El sistema está destinado a ser utilizado en uno de los ámbitos específicos enumerados en el Anexo III?	Art. 6.2 + Anexo III	Alto riesgo. Continuar al punto 3 para confirmar.	Evaluar derechos fundamentales en 7.2.1.1.2.
3	A pesar de estar en el Anexo III, ¿el sistema cumple **TODAS** estas condiciones: (a) realiza una tarea procedimental limitada; (b) mejora el resultado de una actividad humana previa; y (c) no sustituye ni influye sustancialmente en la evaluación humana?	Art. 6.3	Puede excluirse de alto riesgo con justificación documentada y continuar a punto 7.2.1.1.2.	Mantener alto riesgo y continuar.

Tabla 10: Cuestionario riesgo alto

7.2.1.1.2. EN BASE A LOS DERECHOS FUNDAMENTALES

Más allá de la referencia formal a los anexos, resulta imprescindible que las organizaciones integren un análisis de derechos fundamentales en el proceso de clasificación. Aunque este análisis no aparece como un artículo independiente en el RIA, constituye una buena práctica metodológica y garantiza la proporcionalidad en la aplicación del Reglamento.

A continuación, se presenta un cuestionario de clasificación por riesgo alto:

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

N.º	Pregunta de verificación	Fundamento metodológico	Si "Sí"	Si "No"
1	¿El sistema puede condicionar el acceso, la denegación o la calidad de servicios esenciales (sanidad, vivienda, energía, prestaciones sociales)?	Principio de acceso a derechos básicos	Clasificar como alto riesgo.	Continuar.
2	¿Influye en procesos de contratación, promoción, evaluación o despido de personas, o en la asignación de tareas?	Derecho al trabajo y no discriminación	Clasificar como alto riesgo.	Continuar.
3	¿Puede comprometer la seguridad física o la salud de las personas en caso de fallo o de un resultado incorrecto?	Derecho a la vida e integridad física	Clasificar como alto riesgo.	Continuar.
4	¿El sistema procesa datos sensibles o puede llevar a una vigilancia que afecte a la privacidad, la libertad de expresión o de reunión?	Derecho a la vida e integridad	Clasificar como alto riesgo.	Continuar.
5	¿Existe riesgo de impacto negativo en privacidad, igualdad de trato u otros derechos fundamentales no previstos expresamente en los anexos?	Principio de igualdad y no discriminación	Clasificar como alto riesgo.	Puede considerarse de riesgo inferior con justificación documentada.

Tabla 11: Cuestionario de impacto a derechos fundamentales

Una vez completada la evaluación de alto riesgo, debe continuarse obligatoriamente con la verificación de obligaciones de transparencia (apartado 7.2.2), dado que algunos sistemas de alto riesgo también pueden requerir medidas de transparencia explícita a los usuarios conforme al artículo 50 del RIA.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

7.2.1.2. Sistemas de riesgo limitado

En el caso de no clasificarse como riesgo prohibido ni riesgo alto un sistema de IA, se procede a identificarse como sistema con riesgo limitado cuando su utilización pueda generar confusión, manipulación o falta de transparencia en la interacción con personas físicas. El objetivo es garantizar que las personas sepan cuándo interactúan con una IA, cuándo consumen contenidos sintéticos o cuándo se aplican técnicas sensibles como el reconocimiento de emociones o la categorización biométrica.

A continuación, se presenta un cuestionario de clasificación por riesgo limitado:

N.º	Pregunta de verificación	Referencia legal	Si "Sí"	Si "No"
1	¿El sistema interactúa con personas físicas sin que sea evidente que se trata de una IA?	Art. 50.1	Riesgo limitado. Exige informar de forma clara al inicio de la interacción.	Continuar.
2	¿Genera o manipula contenidos sintéticos (texto, audio, vídeo, imagen) que puedan confundirse con materiales auténticos?	Art. 50.2	Riesgo limitado. Exige etiquetado o marca visible.	Continuar.
3	¿Realiza reconocimiento de emociones o categorización biométrica sobre personas físicas?	Art. 50.3	Riesgo limitado. Exige informar a los afectados.	Continuar.
4	¿Genera "deepfakes" sin divulgación clara y visible de su carácter artificial?	Art. 50.2 + 5.1. j	Si hay divulgación → Riesgo limitado; si no → Prohibido.	Continuar.

Tabla 12: Cuestionario riesgo limitado

7.2.1.3. Sistemas de riesgo mínimo

La categoría de **riesgo mínimo** corresponde a los sistemas de IA que, tras la aplicación de los pasos anteriores, **no encajan en las categorías de prohibido, alto riesgo o riesgo limitado**. Estos sistemas no suponen amenazas relevantes para la salud, la seguridad ni los derechos fundamentales, y por ello no están sujetos a obligaciones regulatorias específicas en el RIA.

No obstante, se recomienda documentar su clasificación y aplicar las **directrices de buenas prácticas** (seguridad técnica, calidad de datos, transparencia básica) con el fin de fomentar la confianza y anticipar futuras auditorías.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

A continuación, se presenta un cuestionario de verificación:

N.º	Pregunta de verificación	Fundamento	Si "Sí"	Si "No"
1	¿El sistema realiza funciones rutinarias de filtrado o clasificación sin impacto en derechos fundamentales (ej. filtros de spam, detección de virus)?	Considerandos RIA	Clasificar como riesgo mínimo.	Continuar.
2	¿Se limita a recomendar contenidos o productos sin condicionar el acceso a servicios esenciales?	Principio de proporcionalidad	Clasificar como riesgo mínimo.	Continuar.
3	¿Se aplica en entornos de soporte administrativo o eficiencia interna, sin efectos directos sobre personas?	Evaluación residual	Clasificar como riesgo mínimo.	Reevaluar el sistema o ámbito de aplicación

Tabla 13: Cuestionario riesgo mínimo

7.2.1.4. Clasificación de modelo o sistema GPAI

Los **modelos de uso general (GPAI)** son sistemas de IA entrenados con grandes volúmenes de datos y diseñados para tareas diversas en múltiples dominios, más allá del caso de uso inicial. Por su potencial de reutilización, el RIA impone obligaciones específicas de **documentación, trazabilidad y transparencia**.

La clasificación como GPAI se confirma a través de preguntas que evalúan el alcance de los datos, el tipo de entrenamiento y el grado de generalidad de las capacidades.

A continuación, se presenta un cuestionario de clasificación:

N.º	Pregunta de verificación	Referencia legal	Si "Sí"	Si "No"
1	¿El sistema ha sido entrenado con un gran volumen de datos (multilingües, multimodales o de alta escala)?	Arts. 51–52	Considerar GPAI. Continuar verificaciones.	Continuar.
2	¿Los datos utilizados abarcan múltiples dominios o contextos?	Arts. 51–52	Reforzar clasificación como GPAI.	Continuar.
3	¿Se ha utilizado autosupervisión o técnicas	Arts. 51–52	Indicio fuerte de GPAI.	Continuar.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

N.º	Pregunta de verificación	Referencia legal	Si "Sí"	Si "No"
	de entrenamiento de gran escala?			
4	¿El sistema presenta un alto grado de generalidad en sus capacidades (ej. responder preguntas, generar texto, analizar datos en diferentes contextos sin reentrenamiento)?	Arts. 51–52	Clasificar como GPAI.	Continuar.
5	¿Puede realizar de manera competente una gran variedad de tareas distintas?	Arts. 51–52	Confirmar GPAI. Aplicar obligaciones específicas.	Si todas las respuestas previas son "No", no es GPAI.

Tabla 14: Cuestionario sistema GPAI

7.2.1.4.1. EVALUACIÓN DE GPAI CON RIESGO SISTÉMICO

En caso de disponer de un sistema GPAI tras el análisis realizado en el punto 7.1.1, se procede a la evaluación y clasificación como GPAI con riesgo sistémico.

El RIA prevé dos vías para su identificación:

- **Criterio objetivo**: superar el umbral de 10^{25} FLOPs en entrenamiento acumulado.
- **Criterios cualitativos**: designación por la Comisión Europea u otro organismo competente, así como evidencias de que el modelo tiene un **impacto sistémico transversal** (derechos fundamentales, economía, democracia, seguridad pública).

N.º	Pregunta de verificación	Referencia legal	Si "Sí"	Si "No"
1	¿El entrenamiento acumulado supera el umbral de 10^{25} FLOPs?	Art. 52 + Anexo XIII	Clasificar como GPAI con riesgo sistémico.	Continuar.
2	¿Ha sido designado por la Comisión Europea u otra autoridad como modelo de alto impacto?	Art. 52	Clasificar como GPAI con riesgo sistémico.	Continuar.
3	¿El modelo tiene un alcance masivo en usuarios o dominios críticos, con riesgo de amplificar desigualdades, manipular	Art. 52 + Considerandos	Clasificar como GPAI con riesgo sistémico.	Continuar.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

N.º	Pregunta de verificación	Referencia legal	Si "Sí"	Si "No"
	opinión pública o alterar procesos democráticos?			
4	¿El sistema es estructuralmente interconectado en múltiples aplicaciones, generando riesgos de cascada o dependencia masiva en caso de fallo?	Art. 52	Clasificar como GPAI con riesgo sistémico.	Continuar.
5	¿Existen evidencias de falta de control humano efectivo o de que un error del modelo podría producir daños irreversibles a gran escala?	Art. 52	Clasificar como GPAI con riesgo sistémico.	Fin del cuestionario.

Tabla 15:Cuestionario GPAI con riesgo sistémico

7.2.2. Evaluación de obligaciones de transparencia

El Reglamento (UE) 2024/1689 establece en su artículo 50 un conjunto de obligaciones de transparencia aplicables a determinados sistemas de IA, independientemente de su clasificación de riesgo. Esto significa que incluso los sistemas de alto riesgo pueden tener que cumplir con estos requisitos si interactúan directamente con usuarios o generan contenidos sintéticos susceptibles de confundir.

El objetivo de este paso es asegurar que las personas afectadas sepan cuándo están interactuando con una IA, cuándo están expuestas a reconocimiento biométrico o de emociones, o cuándo consumen contenidos artificiales. Estas medidas no alteran la clasificación de riesgo, pero sí añaden obligaciones adicionales que deben documentarse en el expediente de cumplimiento.

Para determinar si su sistema de alto riesgo también tiene un riesgo de interacción:

N.º	Pregunta de verificación	Referencia legal	Si "Sí"	Si "No"
1	¿El sistema interactúa con personas físicas sin que sea evidente que se trata de una IA?	Art. 50.1	Debe informar clara y previamente a los usuarios.	Continuar.
2	¿El sistema genera o manipula contenidos	Art. 50.2	Debe etiquetar o marcar de forma	Continuar.

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
Plan de Recuperación, Transformación y Resiliencia
incibe_ INSTITUTO NACIONAL DE CIBERSEGURIDAD
Junta de Andalucía
Agencia Digital de Andalucía

N.º	Pregunta de verificación	Referencia legal	Si "Sí"	Si "No"
	sintéticos (texto, imagen, audio, vídeo) que puedan confundirse con material auténtico?		visible dichos contenidos.	
3	¿El sistema realiza reconocimiento de emociones o categorización biométrica sobre personas físicas?	Art. 50.3	Debe informar explícitamente a los afectados.	Continuar.
4	¿El sistema genera deepfakes sin divulgación clara?	Art. 5.1.j + 50.2	Si hay divulgación → se aplica obligación de transparencia; si no hay → prohibido.	Fin de cuestionario.

Tabla 16: Cuestionario de riesgo de interacción

7.3. Justificación de la clasificación

Se hace referencia a la documentación y trazabilidad necesaria para verificar y demostrar que la clasificación de riesgo asignada al sistema de IA se ha realizado correctamente conforme al Reglamento. Esta documentación constituye un conjunto de las evidencias que permite justificar la categoría de riesgo evaluado, asegurar la trazabilidad del proceso de evaluación.

La clasificación de un sistema de IA conforme al Reglamento (UE) 2024/1689 no puede quedar en una simple decisión final; debe estar respaldada por un expediente documentado que asegure la trazabilidad y permita demostrar, ante auditorías internas o autoridades competentes, que la evaluación se realizó de forma completa y conforme al marco legal.

Este expediente constituye la evidencia central de cumplimiento, y debe conservarse actualizado durante todo el ciclo de vida del sistema.

Elementos esenciales del expediente de clasificación

I. **Identificación y descripción del sistema**

- Nombre, versión y responsable del sistema.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

- Propósito funcional y resultados esperados.
- Contexto de uso (sector, entorno operativo, usuarios previstos).
- Tecnología subyacente y componentes clave.

II. Inventario de información de base

- Resumen de los datos recopilados en el análisis de propósito y contexto (7.1.1).
- Resultados de la verificación de excepciones del artículo 2.
- En el caso de sistemas de IA que constituyan productos sanitarios (MDR) o productos de diagnóstico in vitro (IVDR), deberá incluirse la documentación de la evaluación de conformidad, el certificado del organismo notificado, el marcado CE y el identificador único de producto (UDI), como evidencias regulatorias que deben presentarse inicialmente.

III. Cuestionarios de clasificación cumplimentados

- Respuestas “sí/no” en los apartados de prohibido, alto riesgo, limitado, mínimo y GPAI.
- Observaciones justificativas en cada respuesta, con ejemplos o evidencias.
- Referencias explícitas a artículos o anexos del RIA aplicados.

IV. Decisión final de clasificación

- Categoría asignada (riesgo prohibido, alto, limitado, mínimo, GPAI/GPAI con riesgo sistémico).
- Motivación breve, pero clara, del razonamiento seguido.
- En caso de aplicación de exclusiones (art. 2 o 6.3), justificación técnica y jurídica detallada.

V. Metadatos de la evaluación

- Fecha de la evaluación.
- Equipo y roles de las personas responsables.
- Versión de la metodología empleada (para asegurar consistencia en el tiempo).

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

8. SANCIONES

El Reglamento (UE) 2024/1689 establece un régimen sancionador específico para garantizar el cumplimiento de sus disposiciones. Las sanciones tienen como finalidad disuadir conductas de incumplimiento, corregir desviaciones en la clasificación de sistemas de inteligencia artificial y proteger los derechos fundamentales de los ciudadanos de la Unión.

En el contexto de los entornos de salud y de Smart Cities, el riesgo de sanciones cobra especial relevancia, ya que una clasificación incorrecta puede afectar directamente a pacientes, comunidades urbanas y servicios públicos esenciales.

8.1. Reevaluación de sistemas de IA clasificados incorrectamente (artículo 80)

El artículo 80 del RIA prevé la posibilidad de que una autoridad competente revise la clasificación de un sistema de IA cuando existan indicios de que la categoría de riesgo asignada es incorrecta.

- **Corrección de clasificación.** Si un sistema ha sido identificado como de riesgo no alto, pero la autoridad determina que sí reúne las condiciones para ser considerado de alto riesgo conforme al artículo 6 y los anexos del Reglamento, podrá iniciar un procedimiento de reevaluación.
- **Obligación de subsanación.** El proveedor estará obligado a corregir la clasificación y a cumplir con las obligaciones reforzadas aplicables a los sistemas de alto riesgo (por ejemplo, gestión de riesgos, gobernanza de datos, supervisión humana, documentación técnica).
- **Elusión deliberada**. En caso de detectarse que la clasificación errónea fue intencionada o que se buscó eludir las obligaciones de forma consciente, la autoridad podrá imponer sanciones agravadas.

La correcta **documentación y trazabilidad del proceso de clasificación**, especialmente en la aplicación del artículo 6.3 (criterios de alto riesgo) y en la justificación de exclusiones, constituye la principal herramienta de defensa frente a una reevaluación desfavorable.

Financiado por la Unión Europea NextGenerationEU
Plan de Recuperación, Transformación y Resiliencia
Junta de Andalucía
Agencia Digital de Andalucía

8.2. Consecuencias económicas por incumplimiento de las obligaciones de clasificación (artículo 99)

El artículo 99 del Reglamento establece un régimen de sanciones económicas graduadas en función de la gravedad del incumplimiento:

- Infracciones más graves. Incluyen la utilización de prácticas prohibidas (riesgo inaceptable) o la comercialización de sistemas de alto riesgo sin cumplir las obligaciones esenciales. Estas conductas pueden dar lugar a multas administrativas de hasta 35 millones de euros o el 7 % del volumen de negocio anual total mundial de la empresa, el que resulte mayor.

- Infracciones relativas a sistemas de alto riesgo. La omisión en el cumplimiento de obligaciones específicas aplicables a sistemas de alto riesgo (gestión de riesgos, gobernanza de datos, documentación, supervisión humana, ciberseguridad) puede acarrear sanciones de hasta 15 millones de euros o el 3 % del volumen de negocio anual total mundial.

- Infracciones de carácter menos grave. El suministro de información incorrecta, incompleta o engañosa a autoridades competentes puede sancionarse con multas de hasta 7,5 millones de euros o el 1 % del volumen de negocio anual total mundial.

- Atenuación y proporcionalidad. Las sanciones deben ser proporcionales y tener en cuenta factores como la naturaleza, gravedad y duración de la infracción, así como el grado de cooperación del infractor con las autoridades.

En los sectores de salud y Smart Cities, la materialización de estas sanciones puede suponer no solo un impacto económico significativo, sino también una pérdida reputacional y de confianza institucional.

9. ANEXOS

9.1. Anexo 1: Casos de uso de IA en Salud y Smart Cities clasificados como riesgo prohibido

Casos de uso en Salud:

Tipo de sistema	Descripción
Manipulación subliminal con perjuicio considerable	Un asistente digital hospitalario que introduce estímulos para convencer a pacientes a renunciar a pruebas diagnósticas costosas, alterando sustancialmente su decisión informada y con impacto potencial en su salud.
Puntuación social sanitario	Sistema que asigna una "puntuación cívica" al paciente y limita su acceso a tratamientos públicos por comportamientos fuera del contexto clínico original de los datos.
Predicción de criminalidad	Motor que valora el riesgo delictivo de visitantes en un hospital basándose únicamente en su perfil o rasgos de personalidad, sin hechos objetivos.
Scraping facial	Proyecto para ampliar una base de reconocimiento facial de seguridad hospitalaria mediante extracción indiscriminada de rostros de redes sociales.
Categorización biométrica sensible	Clasificador que etiqueta a pacientes por orientación sexual a partir de su biometría para "optimizar" flujos asistenciales.

Casos de uso en Smart Cities:

Tipo de sistema	Descripción
Puntuación social cívica	Sistema que asigna una puntuación de "buen ciudadano" que condiciona acceso a vivienda pública, permisos o servicios por conductas no relacionadas (p. ej., publicaciones en redes, historial de compras).
Reconocimiento biométrico remoto en espacios públicos en tiempo real	Proyecto para crear una red de cámaras urbanas que identifica y rastrea peatones en directo para fines generales de seguridad o gestión urbana.
Scraping facial a gran escala	Proyecto municipal que compila una base de datos de rostros desde redes sociales y cámaras públicas para alimentar sistemas de reconocimiento facial de ciudad.
Categorización biométrica por atributos sensibles	Sistema que infiere etnia, religión u orientación sexual de viandantes a partir de CCTV para "optimizar" control de accesos o recursos.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

Tipo de sistema	Descripción
Predicción de criminalidad basada en perfiles	Motor que etiqueta a individuos o zonas "de alto riesgo" exclusivamente por rasgos demográficos, personalidad o presencia física, sin hechos objetivos.
Sistemas de voz/rasgos para detectar "intenciones" delictivas	Detección de estrés/emoción en altavoces urbanos o transporte para identificar supuestas intenciones criminales basadas en perfiles.

9.1. Anexo 2: Casos de uso de Salud y Smart Cities clasificados como riesgo alto

Casos de uso en Salud:

Tipo de sistema	Descripción
Triaje de urgencias asistido por IA	Sistema que prioriza pacientes y recursos en base a signos vitales y síntomas. Afecta al acceso a la atención y seguridad del paciente.
Soporte al diagnóstico por imagen (radiología, cardio, etc.)	Herramienta que detecta o clasifica lesiones para ayudar al facultativo. Se introduce como componente de un producto sanitario y puede influir en decisiones clínicas críticas.
Recomendación de dosis/ajuste terapéutico (p. ej., insulina, anticoagulantes)	Herramienta que sugiere parámetros de medicación; requiere supervisión humana y gestión de riesgos rigurosa.
Planificación y control asistido de robots/árboles de decisión quirúrgicos	Sistema que apoya la planificación de incisiones, trayectorias o selección de instrumental; componente de seguridad de un producto sanitario.
Verificación biométrica de identidad del paciente para medicación de alto riesgo o acceso a historia clínica	Sistema de verificación (no identificación masiva) que evita errores de identidad en los pacientes

Casos de uso en Smart Cities:

Tipo de sistema	Descripción
Control adaptativo de semáforos con prioridad a emergencias	Herramienta de IA gestiona fases y coordina corredores verdes. Apoyo a la gestión de infraestructuras críticas con impacto en la seguridad vial.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

Tipo de sistema	Descripción
Automatización segura en metro/tranvía (regulación de velocidad y separación)	Componente de seguridad de sistemas ferroviarios urbanos; cualquier fallo puede comprometer la seguridad.
Verificación biométrica de acceso a instalaciones críticas	Sistema que autentica a operarios mediante biometría de verificación
Motor de decisión para adjudicar ayudas sociales o bonos energéticos	Sistema que determina el acceso a servicios públicos esenciales; requiere gobernanza de datos, explicabilidad y supervisión humana.
Gestión inteligente de red de agua	IA que opera infraestructura crítica con riesgo para salud pública si falla
Priorización y despacho en emergencias (112/911)	Sistema que clasifica llamadas y sugiere recursos/tiempos de respuesta; afecta directamente a la seguridad y acceso a servicios de emergencia.

9.2. Anexo 3: Casos de uso en entornos de Salud y Smart Cities clasificados como riesgo limitado

Casos de uso en Salud:

Tipo de sistema	Descripción
Herramienta conversacional de información administrativa	Chatbot que responde dudas sobre horarios, ubicación de consultas, preparación para pruebas, o estado de trámites. Transparencia: avisa que es IA y ofrece canal humano.
Herramienta de lectura fácil de informes	Herramienta que simplifica el lenguaje de informes clínicos para comprensión del paciente sin alterar su contenido médico. Transparencia: marca como “texto generado/transformado por IA” y sugiere confirmar con el profesional.
Motor generativo de materiales educativos	Motor que crea folletos o vídeos explicativos sobre autocuidados y prevención, revisados por personal sanitario antes de publicar. Transparencia: marca de agua/etiqueta de contenido sintético.
Sistema de recordatorios personalizados de citas y preparación	Envía avisos de ayuno, hidratación, documentación, o transporte. No decide tratamientos ni prioriza a pacientes.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

Tipo de sistema	Descripción
Herramienta de traducción asistida para consultas	Traduce en tiempo real entre paciente y personal, con opción de revisión o corrección por el profesional. Transparencia: informa que la traducción es asistida por IA.
Sistema informativo de tiempos de espera y ocupación	Estima y muestra tiempos de espera en urgencias/consultas para gestionar expectativas. No determina triaje ni acceso.

Casos de uso en Smart Cities:

Tipo de sistema	Descripción
Herramienta conversacional municipal	Chatbot que atiende preguntas sobre trámites, eventos, reciclaje o normativas locales. Transparencia: identifica que es IA y permite escalar a una persona.
Motor de recomendación de rutas urbanas	Sugiere combinaciones de transporte (a pie, bici compartida, bus, metro) optimizando tiempo o coste; no controla el tráfico ni impone decisiones.
Sistema de predicción de ocupación de aparcamientos	Muestra probabilidad de encontrar plaza por zona para informar al conductor; no activa sanciones ni vigilancia.
Clasificador de incidencias ciudadanas (back-office)	Etiqueta y enruta reportes de baches, alumbrado o ruido al área correspondiente; un operador valida antes de actuar.
Motor generativo de borradores de comunicaciones públicas	Sistema que redacta notas de prensa, cartelería o mensajes en redes, con revisión humana y marcado de contenido sintético cuando aplique.
Sistema informativo de calidad del aire y alertas no coercitivas	Agrega sensores y notifica niveles de contaminación con recomendaciones voluntarias (p. ej., “evitar ejercicio intenso al aire libre”). No restringe movilidad ni derechos.

9.3. Anexo 3: Casos de uso de Salud y Smart Cities clasificados como riesgo mínimo

Casos de uso en Salud:

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA SECRETARÍA DE ESTADO DE TELECOMUNICACIONES E INFRAESTRUCTURAS DIGITALES
Plan de Recuperación, Transformación y Resiliencia
incibe_ INSTITUTO NACIONAL DE CIBERSEGURIDAD
Junta de Andalucía
Agencia Digital de Andalucía

Tipo de sistema	Descripción
Corrección de ortografía y estilo en documentación clínica	Herramienta de asistencia lingüística que sugiere mejoras de gramática y claridad en textos, revisadas siempre por el profesional. No decide ni modifica contenidos clínicos sin validación.
Transcripción de voz a texto para borradores de notas	Sistema de dictado que convierte audio en texto para acelerar la redacción; el clínico revisa y valida antes de incorporar al historial. Apoyo administrativo con revisión humana.
Personalización de interfaz en el portal del paciente	Sistema de aprendizaje de preferencias (tamaño de letra, tema oscuro, disposición) sin afectar contenidos clínicos.
Generación de textos divulgativos para prevención con revisión previa	Herramienta generativa que propone borradores de materiales de promoción de salud que son revisados y aprobados por comunicación/sanidad. Contenido no clínico y revisado.
Filtrado antispam en canales de contacto institucionales	Sistema de filtrado que detecta y aparta mensajes promocionales o maliciosos antes de la bandeja de entrada del hospital. Riesgo mínimo: seguridad básica sin impacto clínico.

Casos de uso en Smart Cities:

Tipo de sistema	Descripción
Autocompletado de búsquedas en la web municipal	Herramienta de sugerencia que completa términos en el buscador del portal ciudadano. Mejora de usabilidad sin efectos decisorios.
Recomendaciones de ocio y cultura en la agenda local	Motor que sugiere eventos según intereses declarados o consultas previas, sin personalización sensible ni efectos en derechos. Recomendaciones no vinculantes.
Traducción automática de contenidos informativos	Sistema que traduce páginas y avisos públicos a varios idiomas con aviso de posible imprecisión y opción humana. Uso como apoyo lingüístico.
Subtitulado automático de vídeos institucionales	Se trata de una herramienta que genera subtítulos para contenidos audiovisuales del ayuntamiento, revisados antes de publicar. Mejora de accesibilidad
Clasificación temática de archivo fotográfico municipal	Motor que etiqueta imágenes históricas por temática (p. ej., “parques”, “eventos”) para facilitar búsquedas internas.

Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Ministerio para la Transformación Digital y de la Función Pública · Secretaría de Estado de Telecomunicaciones e Infraestructuras Digitales
Plan de Recuperación, Transformación y Resiliencia
incibe_ Instituto Nacional de Ciberseguridad
Junta de Andalucía
Agencia Digital de Andalucía

10. REFERENCIAS

1. Unión Europea. (2024). 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 artificial (Reglamento de Inteligencia Artificial, RIA). Diario Oficial de la Unión Europea. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

2. Unión Europea. (2017). Reglamento (UE) 2017/745 del Parlamento Europeo y del Consejo, de 5 de abril de 2017, sobre los productos sanitarios (MDR). Diario Oficial de la Unión Europea. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32017R0745

3. Unión Europea. (2017). Reglamento (UE) 2017/746 del Parlamento Europeo y del Consejo, de 5 de abril de 2017, sobre los productos sanitarios para diagnóstico in vitro (IVDR). Diario Oficial de la Unión Europea. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32017R0746

4. Comisión Europea. (2025). Directrices sobre prácticas prohibidas de inteligencia artificial en el marco del Reglamento de IA. Bruselas: European Commission. https://ec.europa.eu/newsroom/repository/practices-prohibidas-ia

5. Comisión Europea. (2025). Código de buenas prácticas para modelos de IA de propósito general (GPAI) conforme al artículo 56 del Reglamento de Inteligencia Artificial (RIA). Bruselas: European Commission – Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai

Guía de clasificación de sistemas de inteligencia artificial en base al Reglamento de IA de la UE en los entornos de Smart Cities y Salud

Septiembre de 2025

Versión 1.0

Financiado por la Unión Europea NextGenerationEU