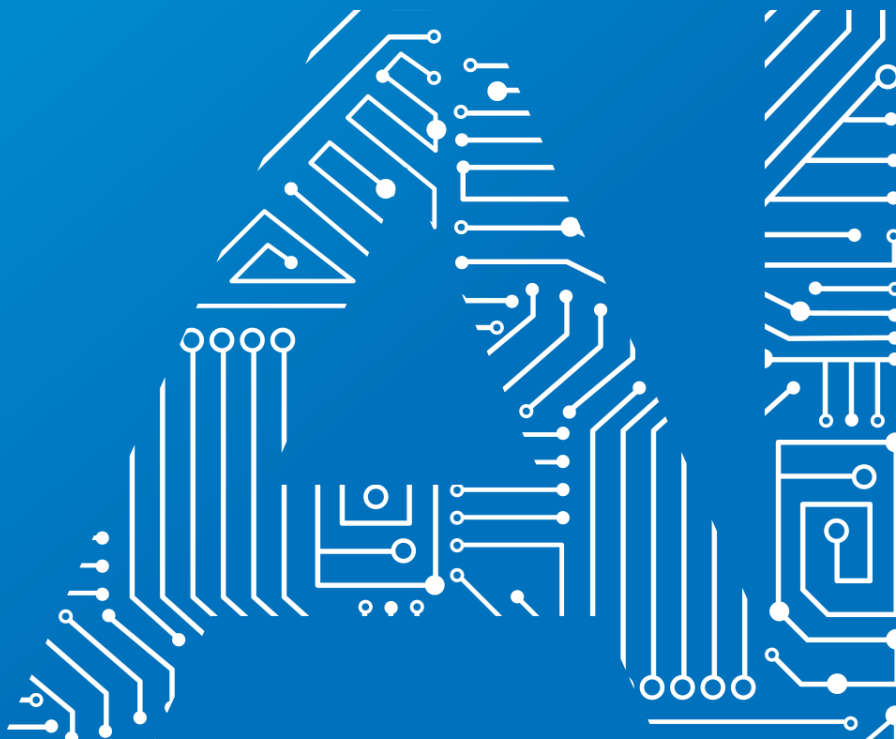
Ética y normativa de la IA – Guía para una IA responsable

BASQUE ARTIFICIAL INTELLIGENCE CENTER



Contenido

Ilustraciones

BASQUE ARTIFICIAL INTELLIGENCE CENTER

Contexto

Dentro del ecosistema de la IA en Euskadi, BAIC se configura como referente, punto de encuentro y colaboración tanto de agentes públicos como privados para el impulso de la IA en Euskadi. La misión de BAIC se centra en aumentar la competitividad vasca acelerando el desarrollo e implantación de la IA de un modo ético, colaborativo y soberano.

En cuanto a los valores de BAIC, estos tienen que ver con que sus contribuciones y actuaciones se alineen con el desarrollo e implementación sostenible, equitativa y éticamente responsable de la IA. De la misma forma, BAIC actúa como el faro que guía y apoya a los diversos agentes del ecosistema de IA en Euskadi hacia el desarrollo y aplicación de una IA confiable y responsable, asegurando que las innovaciones en IA no solo sean técnicamente avanzadas y seguras, sino también legalmente sólidas y éticamente intachables, fomentando así un ambiente de confianza y colaboración que impulse un progreso tecnológico al servicio de la sociedad.

BAIC cumple con su misión a través del despliegue de distintos ejes estratégicos, que, en conjunto, permite construir soluciones completas a los retos complejos que se plantean en materia de IA. Como parte del eje estratégico de Observatorio, BAIC persigue fortalecer y desarrollar el ecosistema de la IA en Euskadi con un enfoque en la vigilancia, el conocimiento, la conexión y el posicionamiento entre distintos agentes del territorio. Se enmarcan dentro de este eje aquellas iniciativas orientadas a guiar y apoyar a los agentes del ecosistema hacia el alineamiento con los pilares de robustez, legalidad y ética, así como con los requisitos para una IA confiable y responsable en su desarrollo e implementación.

En este contexto, en una primera definición de un marco ético para el uso, implementación y desarrollo de la IA en el ecosistema de Euskadi, BAIC presentó a principios del 2024 el Código ético para el desarrollo, uso e implementación de IA en Euskadi[1], con un enfoque basado en el apoyo voluntario y unilateral de este código ético por parte de las organizaciones del ecosistema de IA en Euskadi.

Por su parte, a través del presente documento, **BAIC presenta una Guía de ética y normativa con el objetivo ayudar a los agentes del ecosistema de la IA en Euskadi a comprender y alinearse con la nueva normativa de IA, facilitando la adaptación de sus estrategias y procesos para cumplir tanto con aspectos legislativos como éticos**. Esta nueva guía proporciona directrices detalladas y ejemplos prácticos para garantizar el cumplimiento normativo y promover prácticas responsables y transparentes en el desarrollo y uso de la IA.

[1] BAIC – Código ético para el desarrollo, uso e implementación de la IA en Euskadi

IA ética y legal

IA ética y legal

¿Qué es un Sistema de IA?

En esta Guía tomamos la definición proporcionada por el Reglamento Europeo de Inteligencia Artificial[2] (RIA o AI ACT por sus siglas en inglés):

> *"Un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales".*

¿Cuáles son los requisitos para una IA FIABLE?

Una IA fiable, se apoya sobre tres pilares: debe ser legal (cumplir la legislación aplicable), ética (respetar los principios éticos), y robusta (funcionar de manera segura, para no provocar daños accidentales).

Cada uno de estos pilares es necesario, pero no suficiente.

Así, para que exista una IA fiable, no solo es necesario cumplir la ley. Debemos tener en cuenta que normalmente el desarrollo tecnológico va por delante de las leyes, por lo que es necesario que los sistemas de IA respeten los principios éticos.

El uso de sistemas de IA en nuestra sociedad plantea algunos desafíos éticos, relacionados, por ejemplo, con sus efectos sobre las personas y la sociedad, las capacidades de adopción de decisiones y la seguridad.

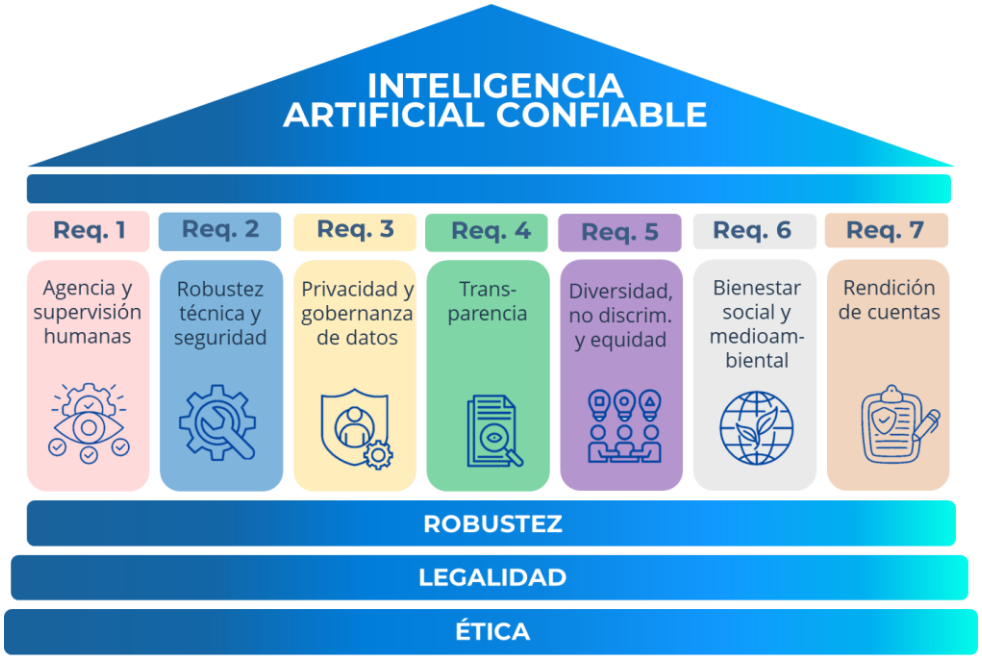
Los principios éticos que deben inspirar el desarrollo y uso de la IA están basados en los derechos fundamentales (el respeto de la dignidad humana es la base común), y son: i) respeto de la autonomía humana, ii) prevención del daño, iii) equidad, iv) explicabilidad.

Las directrices éticas para una IA fiable elaboradas por el grupo de personas expertas de alto nivel creado por la Comisión Europea[3] establece siete requisitos clave para una IA confiable.

[2] https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:L_202401689

[3] https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1

Ilustración 1: Comisión Europea - Requisitos para una IA confiable



BAIC persigue orientar y guiar a los distintos agentes del ecosistema de la IA en Euskadi hacia el alineamiento con los pilares de robustez, legalidad y ética, así como con los requisitos para el desarrollo y uso de una IA confiable y responsable a través del *Código Ético para el desarrollo, uso e implementación de Inteligencia Artificial en Euskadi*[4]. En este, se propone un marco ético con el objetivo de promover prácticas éticas y responsables que maximicen los beneficios de la IA, aseguren su fiabilidad y minimicen los riesgos potenciales, fomentando la confianza de la sociedad y agentes del ecosistema de la IA y expresando nuestros valores compartidos.

Los siete requisitos para una IA confiable han quedado incorporados, en mayor o menor medida, en el articulado del Reglamento de Inteligencia Artificial[5]:

Requisito	Descripción	Artículo RIA
Acción y supervisión humanas	Los sistemas de IA se desarrollan y utilizan como herramienta al servicio de las personas, que respeta la dignidad humana y la autonomía personal, y que funciona de manera que pueda ser controlada y vigilada adecuadamente por seres humanos.	Art. 14. Supervisión humana
Solidez técnica y seguridad	Los sistemas de IA se desarrollan y utilizan de manera que sean sólidos en caso de problemas y resilientes frente a los	Art. 15. Precisión, solidez y ciberseguridad

[4] BAIC – Código ético para el desarrollo, uso e implementación de la IA en Euskadi

[5] https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:L_202401689

	intentos de alterar el uso o el funcionamiento del sistema de IA para permitir su uso ilícito por terceros y reducir al mínimo los daños no deseados.	
Gestión de la privacidad y de los datos	Los sistemas de IA se desarrollan y utilizan de conformidad con las normas en materia de protección de la intimidad y de los datos, y los datos que trata cumplen normas estrictas en términos de calidad e integridad.	Art. 10. Datos y gobernanza de datos
Transparencia	Los sistemas de IA se desarrollan y utilizan de un modo que permita una trazabilidad y explicabilidad adecuadas, y que, al mismo tiempo, haga que las personas sean conscientes de que se comunican o interactúan con un sistema de IA e informe debidamente a los responsables del despliegue acerca de las capacidades y limitaciones de dicho sistema de IA y a las personas afectadas acerca de sus derechos.	Art. 11. Documentación técnica (transparencia del proveedor frente a las autoridades) Art. 13. Transparencia y comunicación de información a los responsables del despliegue (transparencia del proveedor frente a los responsables del despliegue) Art. 50. Obligaciones de transparencia de los proveedores y responsables del despliegue de determinados sistemas de IA (transparencia frente a las personas).
Diversidad, no discriminación y equidad	Los sistemas de IA se desarrollan y utilizan de un modo que incluya a diversos agentes y promueve la igualdad de acceso, la igualdad de género y la diversidad cultural, al tiempo que se evitan los efectos discriminatorios y los sesgos injustos prohibidos por el Derecho nacional o de la Unión.	Art. 5. Prácticas de IA prohibidas Art. 9. Sistema de gestión de riesgos (debe tenerse en cuenta el contexto) Art. 14. Supervisión humana Art. 27. Evaluación de impacto relativa a los derechos fundamentales para los sistemas de IA de alto riesgo Anexo III. Sistemas de IA de alto riesgo incluídos en la lista de casos de uso del Anexo III.

Bienestar social y ambiental	Los sistemas de IA se desarrollan y utilizan de manera sostenible y respetuosa con el medio ambiente, así como en beneficio de todos los seres humanos, al tiempo que se supervisan y evalúan los efectos a largo plazo en las personas, la sociedad y la democracia.	Art. 1, art. 26 (obligación de informar de incidentes graves, que incluyen daños al medio ambiente).
Rendición de cuentas	Los sistemas de IA deben ser auditables, minimizar los efectos negativos, y si estos existen deben notificarse y compensarse.	Art. 17. Sistema de gestión de la calidad (que incluye marco de rendición de cuentas que defina las responsabilidades del personal, directivo y no directivo). Capítulo IX, Sección 4. Vías de recurso.

En definitiva, se promueve la adopción de una IA antropocéntrica, es decir, una IA centrada en el ser humano.

En cualquier caso, desde la Unión Europea se anima a los proveedores y a los responsables del despliegue de todos los sistemas de IA, sean o no de alto riesgo, y de los modelos de IA, a aplicar, con carácter voluntario, requisitos adicionales establecidos en las Directrices éticas de la Unión para una IA fiable[6].

¿Cuál es la relación entre IA FIABLE e IA RESPONSABLE?

Una IA responsable que, cuando se diseña, desarrolla y aplica, respeta los principios éticos y la normativa aplicable, y es auditable, dará como resultado una IA fiable frente a la sociedad y frente a las personas expuestas a esa IA. Es decir, una IA sin riesgo para la salud, la seguridad y los derechos fundamentales y respetuosa con el medio ambiente y la democracia.

¿Cuál es la diferencia entre ÉTICA y NORMATIVA?

Los principios éticos no obligan desde un punto de vista jurídico, y por ello su incumplimiento no genera consecuencias jurídicas, sin perjuicio de las consecuencias reputacionales que pueda sufrir una organización que presenta comportamientos no éticos.

Sin embargo, las obligaciones que se establecen en la normativa son obligaciones jurídicas y, por tanto, su incumplimiento constituye una infracción legal de la cual pueden derivarse consecuencias jurídicas para la organización incumplidora en forma de reclamaciones, multas, etc.

En la práctica nos encontramos con principios éticos que están integrados en la normativa aplicable, lo que les dota de fuerza jurídica (como hemos visto en la tabla anterior).

[6] https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1

Marco normativo de la IA

Marco normativo de la IA

La inteligencia artificial, en la medida que implica el uso de datos, debe respetar la normativa aplicable a los datos, como la normativa de protección de datos y la normativa de propiedad intelectual.

Además, también debe tenerse en cuenta el Reglamento Europeo de Inteligencia Artificial (RIA o AI ACT)[7], recientemente publicado, cuyas claves fundamentales abordamos en el siguiente apartado de esta Guía.

Reglamento General de Protección de Datos

Cuando se entrena un modelo de IA con conjuntos de datos que incluyan datos personales, cuando la información de entrada en el sistema de IA incorpora datos personales, o cuando los resultados de salida del sistema de IA afectan a las personas de un modo significativo (por ejemplo, por tener efectos jurídicos sobre ellas) y, en general, si se tratan datos personales, tenemos que aplicar la normativa de protección de datos (fundamentalmente el Reglamento General de Protección de Datos o RGPD).

De modo resumido, deben cumplirse los principios de protección de datos, que nos obligan a:

1. Tratar los datos solo cuando tengamos una base de legitimación para poder hacerlo; y solo hay seis bases: consentimiento, contrato, cumplimiento legal, interés vital, interés legítimo e interés público (principio de licitud).
2. Informar a las personas de lo que vamos a hacer con sus datos (principio de transparencia).
3. Usar los datos solo para las finalidades que hemos informado (principio de limitación de la finalidad).
4. Usar solo aquellos datos que sean necesarios para esas finalidades (principio de minimización de datos).
5. Usar datos actualizados (principio de exactitud).
6. Suprimir los datos cuando hayan dejado de ser necesarios para esas finalidades (principio de limitación del plazo de conservación).
7. Usar los datos garantizando una seguridad adecuada (principio de integridad y confidencialidad).

Por ejemplo, en virtud del principio de minimización de datos, es necesario analizar si el modelo de IA puede ser entrenado con datos seudonimizados o incluso anonimizados sin que ello afecte al objetivo perseguido. La anonimización requiere aplicar una serie de técnicas específicas por personal especializado para minimizar el riesgo de reidentificación.

Asimismo, en los proyectos basados en IA que traten datos personales, muy probablemente será necesario realizar una evaluación de impacto de protección de datos, para analizar si el tratamiento es necesario y proporcionado, y su impacto en los derechos y libertades de las personas afectadas.

Si los resultados de salida de la IA suponen decisiones individuales automatizadas sobre personas físicas, deberá verificarse si se da alguno de los supuestos permitidos, e informarles (i) de que van a existir esas

[7] https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:L_202401689

decisiones automatizadas, (ii) de la lógica aplicada, y (iii) sobre la importancia y las consecuencias previstas para el interesado.

En esta misma línea, el Estatuto de los Trabajadores establece el derecho del Comité de Empresa a ser informado de las reglas en las que se basan los algoritmos (sean o no de IA) que afectan a la toma de decisiones sobre las condiciones de trabajo.

Normativa de propiedad intelectual

Los datos preexistentes utilizados para el entrenamiento de modelos o como información de entrada al sistema de IA, pueden estar sujetos a derechos de propiedad intelectual. En ocasiones, y tratándose de organismos de investigación sin ánimo de lucro, será posible ampararse en la excepción de minería de textos y datos. Para el resto de los casos, es decir, obras distintas de textos y datos y para finalidades distintas a la investigación, se plantea la cuestión de la autorización para poder usar obras protegidas.

Asimismo, se planten cuestiones relativas a la protección, tanto de la inteligencia artificial en sí misma, como de los resultados de salida obtenidos.

Reglamento europeo de IA

Reglamento europeo de IA

Custiones generales

El Reglamento de Inteligencia Artificial (RIA)[8] se publicó el 12/07/2024 en el Diario Oficial de la Unión Europea (DOUE) y entró en vigor el 01/08/2024.

Sin embargo, no será de aplicación hasta el 02/08/2026, con algunos plazos distintos para ciertas previsiones, a los que nos referimos más adelante.

¿Cuál es el enfoque del RIA?

El Reglamento de Inteligencia Artificial (RIA) es una normativa de seguridad de producto, como lo son las normativas que regulan los juguetes, los ascensores, los vehículos o los dispositivos sanitarios.

Las normativas de seguridad de producto establecen requisitos regulatorios para garantizar que aquellos productos con potencial de causar un daño para la salud o la seguridad de las personas hayan sido probados y validados antes de su lanzamiento al mercado, de modo que solo lleguen al mercado productos que sean seguros.

Así, este tipo de normativa exige que el fabricante realice evaluaciones de conformidad del producto[9] (incluso por terceros independientes), incluya el marcado CE en el producto, retire el producto del mercado si presenta un riesgo, facilite instrucciones de uso para evitar que un uso incorrecto del producto provoque daños, etc.

Estos requisitos (evaluación de la conformidad, marcado CE, instrucciones de uso, vigilancia poscomercialización) también los vamos a encontrar en el RIA, que pretende garantizar que los sistemas de IA en la Unión Europea respeten los valores de la Unión y sean seguros, y no causen daños a la seguridad y salud de las personas ni a sus derechos fundamentales.

Se suele decir que el RIA tiene un enfoque basado en riesgos porque de su articulado pueden deducirse cuatro.

[8] Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo de 13 de junio de 2024 por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) nº 300/2008, (UE) nº 167/2013, (UE) nº 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828.

[9] Procedimiento para demostrar que el producto cumple todos los requisitos obligatorios del producto antes de introducirlo en el mercado de la UE. En ocasiones, se exige que la evaluación de la conformidad la realice un organismo independiente.

Ilustración 2: Niveles de riesgo y distintas exigencias regulatorias en función del caso de uso o finalidad del sistema de IA

	Motivo	**Ejemplos** (Entre otros)
Riesgo inaceptable	Por vulneración de derechos fundamentales y valores de la Unión Europea.	• Sistemas de puntuación social • Identificación biométrica (Real-time) • Manipulación del comportamiento
Alto riesgo	Por riesgo para la salud, seguridad o derechos fundamentales. Sujetos a requisitos regulatorios.	• Acceso a empleo y servicios públicos • Aplicación de la ley • Identificación biométrica (Posteriori)
Riesgo limitado	Por riesgo de suplantación, manipulación o engaño.	• Suplantación de personalidad • Chatbots • Reconocimiento de emociones
Riesgo mínimo	Quedan fuera de la RIA. Pueden aplicar Códigos de Conducta.	• Resto de usos

¿Cuáles son los conceptos relevantes?

A efectos del Reglamento de Inteligencia Artificial, es necesario conocer los conceptos de "sistema de IA", "modelo de IA de uso general", "sistema de IA de uso general" y "modelo de IA de uso general de riesgo sistémico":

Concepto	Características
Sistema de IA	1. basado en máquina 2. autonomía 3. capacidad de adaptación 4. inferencia de resultados de salida (predicciones, contenidos, recomendaciones o decisiones) a partir de la información de entrada, 5. influencia de los resultados de salida en el entorno
Modelo de IA de uso general (MIAUG)	1. generalidad y capacidad de realizar una amplia variedad de tareas distintas de manera competente, 2. puede integrarse en diversos sistemas o aplicaciones posteriores. Por ejemplo, los grandes modelos de lenguaje natural son modelos de IA de uso general. Suelen entrenarse usando grandes volúmenes de datos y a través de diversos métodos, como el aprendizaje autosupervisado, no supervisado o por refuerzo. Pueden introducirse en el mercado de diversas maneras, por ejemplo, a través de bibliotecas, interfaces de programación de aplicaciones (API), como descarga directa o como copia física. Estos modelos pueden modificarse o perfeccionarse y transformarse en nuevos modelos. Aunque los modelos de IA son componentes esenciales de los sistemas de IA, no constituyen por sí mismos sistemas de IA. Los modelos de IA requieren que se les añadan otros componentes, como, por ejemplo, una interfaz de

	usuario, para convertirse en sistemas de IA. Los modelos de IA suelen estar integrados en los sistemas de IA y formar parte de dichos sistemas.
Sistema de IA de uso general	Sistema de IA basado en un MIAUG, que puede servir para diversos fines, tanto para su uso directo como para su integración en otros sistemas de IA. Por ejemplo, Chat GPT.
MIAUG de riesgo sistémico	Modelo de IA de uso general que tiene capacidades de gran impacto. Presunción: cuando la cantidad acumulada de cálculo utilizada para su entrenamiento, medida en operaciones de coma flotante, sea superior a 10^{25}.

¿Qué ámbitos quedan fuera del RIA?

Existen algunos ámbitos en los que el Reglamento de Inteligencia Artificial no se aplica:

a) Sistemas de IA con fines militares, de defensa o de seguridad nacional.

b) Sistemas o modelos de IA desarrollados y puestos en servicio con la única finalidad específica de la investigación y el desarrollo científicos.

c) Actividades de investigación, prueba o desarrollo relativas a sistemas de IA o modelos de IA antes de su introducción en el mercado o puesta en servicio. No están cubiertas por esta exclusión las pruebas en condiciones reales.

d) Actividades puramente personales (no profesionales) realizadas por personas físicas.

¿Qué organizaciones deben cumplir el RIA?

El RIA establece obligaciones tanto para los proveedores que diseñan y desarrollan sistemas de IA incluidos dentro de su ámbito de aplicación, como para las organizaciones que utilizan dichos sistemas, que en el RIA se les denomina responsables del despliegue.

> En una versión anterior se les llamaba usuarios, pero este término se ha descartado, suponemos que para evitar confusiones con la terminología empleada en la normativa de consumidores y usuarios, dirigida a personas físicas que adquieren productos o servicios con fines domésticos o no profesionales.

También establece obligaciones para los proveedores de modelos de IA de uso general y de riesgo sistémico.

¿Cuáles son las autoridades competentes en IA en los Estados Miembros y la UE?

El RIA pretende regular de manera armonizada el uso, desarrollo y supervisión de la IA en la Unión Europea. Para ello, establece un marco de gobernanza a nivel europeo y a nivel de cada Estado Miembro:

1) Unión Europea:

- Oficina de IA (adscrita a la Comisión): asesorará sobre los modelos de IA de uso general, y facilitará la elaboración de códigos de buenas prácticas para este tipo de IA.
- Consejo[10] de IA: asesorará a la Comisión y a los Estados miembros para facilitar la aplicación coherente y eficaz del RIA.

2) Autoridades nacionales competentes:

Cada uno de los Estados Miembros debe designar al menos una autoridad de vigilancia del mercado, encargada de velar por el cumplimiento del RIA, y aplicar las sanciones en caso de infracción.

La Agencia Española de Supervisión de la Inteligencia Artificial (AESIA) será la autoridad de vigilancia del mercado en España.

¿Se prevé la publicación de estándares para acreditar el cumplimiento del RIA?

En la actualidad los organismos de normalización están desarrollando, a petición de la Comisión Europea, una serie de estándares que permitirán acreditar (por medio de presunción) la conformidad con el Reglamento de Inteligencia Artificial en relación con los siguientes diez ámbitos:

1) Sistema de gestión de riesgos para sistemas de IA
2) Gobernanza y calidad de los conjuntos de datos utilizados para crear sistemas de IA
3) Mantenimiento de registros mediante la capacidad de registro de los sistemas de IA
4) Transparencia e información a los usuarios de los sistemas de IA
5) Supervisión humana de los sistemas de IA
6) Especificaciones de precisión para los sistemas de IA
7) Especificaciones de robustez de los sistemas de IA
8) Especificaciones de ciberseguridad para los sistemas de IA
9) Sistemas de gestión de la calidad para proveedores de sistemas de inteligencia artificial, incluidos los procesos de seguimiento postcomercialización
10) Evaluación de la conformidad de los sistemas de IA

Prácticas de IA prohibidas - Riesgo inaceptable

El RIA prohíbe comercializar, instalar y usar los siguientes sistemas de IA:

10 En borradores anteriores del RIA, denominado “Comité de IA”

1) Sistema de IA que **usa técnicas subliminales, manipuladoras o engañosas** para alterar el comportamiento de una persona o colectivo.

2) Sistema de IA que **aprovecha vulnerabilidades** (edad, discapacidad, situación social o económica) para alterar el comportamiento de una persona o colectivo.

3) Sistemas de IA para **puntuación social**.

4) Sistemas de IA para **evaluar el riesgo de que una persona cometa delitos**, sobre la única **base** de **su perfil o** de su **personalidad**.

   Excepción: sistemas de IA para apoyar la valoración humana que ya se base en hechos objetivos directamente relacionados con una actividad delictiva.

5) Sistemas de IA para crear **bases de datos de reconocimiento facial mediante la extracción no selectiva de imágenes faciales** de internet.

6) Sistemas de IA para **inferir las emociones de una persona en los lugares de trabajo y en los centros educativos.**

   Excepción: sistema de IA destinado a ser instalado o introducido en el mercado por motivos médicos o de seguridad**.**

7) Sistemas de **categorización biométrica** que clasifiquen individualmente a las personas sobre la base de sus datos biométricos para deducir o inferir su raza, opiniones políticas, afiliación sindical, convicciones religiosas o filosóficas, vida sexual u orientación sexual (categorías especiales de datos).

   Esta prohibición no incluye el etiquetado o filtrado de conjuntos de datos biométricos adquiridos lícitamente, como imágenes, basado en datos biométricos ni la categorización de datos biométricos con fines policiales.

El RIA prohíbe usar:

8) Sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público con fines policiales, con algunas excepciones definidas y sujetas a una serie de salvaguardias.

Estas prácticas quedarán prohibidas a partir del 02/02/2025.

Sistemas de IA de alto riesgo - Riesgo alto

Como recordatorio, corresponden a sistemas de IA de riesgo alto para la salud, la seguridad, o los derechos fundamentales (estos sistemas de IA quedan sujetos a requisitos regulatorios).

Los sistemas de IA de alto riesgo se clasifican en **dos tipos**:

a) Por un lado, sistemas de IA asociados a productos regulados (es decir, productos que están sujetos a normativa de seguridad de producto), que pueden tener un alto impacto en la salud o en la seguridad de las personas.

b) Y, por otro lado, sistemas de IA que se encuentran incluidos en una lista de casos de uso que se consideran de alto riesgo por su alto impacto en los derechos fundamentales de las personas.

a) Sistemas de IA asociados a productos regulados:

A su vez, existen dos tipos:

1. **Sistema de IA que sea un producto regulado** y este producto deba someterse a una **evaluación de la conformidad de terceros**. Por un ejemplo, un robot autónomo, o un sistema de diagnóstico.
2. **Sistema de IA que sea componente de seguridad de un producto regulado** y este producto deba someterse a una **evaluación de la conformidad de terceros**. Por ejemplo, un sistema de freno de un ascensor.

Los productos regulados afectados son los siguientes:

- Máquinas | Juguetes | Motos acuáticas | Ascensores y componentes de seguridad para ascensores | Aparatos y sistemas de protección para uso en atmósferas potencialmente explosivas | Equipos radioeléctricos | Equipos a presión | Instalaciones de transporte por cable | Equipos de protección individual | Aparatos que queman combustibles gaseosos | Productos sanitarios | Productos sanitarios para diagnóstico in vitro.

- Aviación civil | Vehículos | Equipos marinos | Sistema ferroviario | Aeronaves no tripuladas.

Por lo tanto, habrá que analizar si de acuerdo con la normativa reguladora del producto en cuestión (del cual el sistema de IA es un componente de seguridad o es el propio producto), el producto debe someterse o no a una evaluación de la conformidad de terceros, de tal modo, que si la respuesta es afirmativa, el sistema de IA se considera de alto riesgo.

Las previsiones del RIA en relación con este tipo de sistemas de IA, serán aplicables a partir del 02/08/2027.

b) Sistemas de IA incluidos en la lista de casos de uso de alto riesgo

Se establecen 8 ámbitos, en los cuales se identifican una serie de casos de uso o aplicaciones del sistema de IA, que se consideran de alto riesgo:

1. Biometría
2. Infraestructuras críticas
3. Educación y formación profesional
4. Empleo, gestión de los trabajadores y acceso al autoempleo
5. Acceso a servicios y prestaciones esenciales
6. Aplicación de la ley
7. Migración, asilo y gestión del control fronterizo
8. Administración de justicia y procesos democráticos

Las previsiones del RIA en relación con este tipo de sistemas de IA, serán aplicables a partir del 02/08/2026.

A continuación, detallamos los casos de uso de sistemas de IA que se consideran de alto riesgo en cada uno de los 8 ámbitos:

(1) Biometría

a) Sistemas de identificación biométrica remota (se excluye la verificación biométrica).

b) Sistemas de IA para categorización biométrica en función de características sensibles basada en la inferencia de dichas características.

c) Sistemas de IA para el reconocimiento de emociones (fuera de los lugares de trabajo o centros educativos).

(2) Infraestructuras críticas

Sistemas de IA que sean componentes de seguridad en la gestión y el funcionamiento de las infraestructuras digitales críticas, del tráfico rodado o del suministro de agua, gas, calefacción o electricidad.

(3) Educación y formación profesional

a) Sistemas de IA para determinar el acceso o la admisión de personas físicas a centros educativos y de formación profesional.

b) Sistemas de IA para evaluar los resultados del aprendizaje.

c) Sistemas de IA para evaluar el nivel de educación adecuado que recibirá una persona o al que podrá acceder.

d) Sistemas de IA para el seguimiento y la detección de comportamientos prohibidos de los estudiantes durante los exámenes.

(4) Empleo, gestión de los trabajadores y acceso al autoempleo

a) Sistemas de IA para contratar o seleccionar personal.

b) Sistemas de IA para tomar decisiones sobre las condiciones laborales, la promoción o terminación de contratos laborales, la asignación de tareas a partir de comportamientos individuales o rasgos o características personales o para supervisar y evaluar el rendimiento y el comportamiento de las personas en el marco de relaciones laborales.

(5) Acceso y disfrute de servicios y prestaciones esenciales

a) Sistemas de IA para evaluar la admisibilidad de las personas físicas para beneficiarse de servicios y prestaciones esenciales de asistencia pública, incluidos los servicios de asistencia sanitaria, así como para conceder, reducir o retirar dichos servicios y prestaciones o reclamar su devolución.

b) Sistemas de IA para evaluar la solvencia de personas físicas o establecer su calificación crediticia, salvo los sistemas de IA utilizados para detectar fraudes financieros.

c) Sistemas de IA para la evaluación de riesgos y la fijación de precios de seguros de vida y de salud.

d) Sistemas de IA para la evaluación y la clasificación de las llamadas de emergencia realizadas por personas físicas o para el envío o el establecimiento de prioridades en el envío de servicios de primera intervención en situaciones de emergencia, por ejemplo, policía, bomberos y servicios de

asistencia médica, y en sistemas de triaje de pacientes en el contexto de la asistencia sanitaria de urgencia.

(6) **Aplicación de la ley (fines policiales)**

a) Sistemas de IA para evaluar el riesgo de que una persona física sea víctima de delitos.

b) Sistemas de IA como polígrafos o similares.

c) Sistemas de IA para evaluar la fiabilidad de pruebas durante la investigación o el enjuiciamiento de delitos.

d) Sistemas de IA para evaluar el riesgo de que una persona física cometa un delito o reincida o para evaluar rasgos y características de la personalidad o comportamientos delictivos pasados de personas físicas o colectivos.

e) Sistemas de IA para elaborar perfiles de personas físicas, durante la detección, la investigación o el enjuiciamiento de delitos.

(7) Migración, asilo y gestión del control fronterizo

a) Sistemas de IA como polígrafos o herramientas similares.

b) Sistemas de IA para evaluar el riesgo que plantee una persona física que entre en el territorio de un Estado miembro.

c) Sistemas de IA para ayudar a examinar las solicitudes de asilo, visado o permiso de residencia.

d) Sistemas de IA para detectar, reconocer o identificar a personas físicas, con excepción de la verificación de documentos de viaje.

(8) Administración de justicia y procesos democráticos

a) Sistemas de IA para ayudar a una autoridad judicial en la investigación e interpretación de hechos y de la ley, así como en la aplicación de la ley a un conjunto concreto de hechos.

b) Sistemas de IA para influir en el resultado de una elección o referéndum o en el comportamiento electoral de personas físicas que ejerzan su derecho de voto en elecciones o referendos.

No obstante, un sistema de IA de los incluidos en la lista de casos de uso **no se considerará de alto riesgo si no plantea un riesgo importante** de causar un perjuicio a la salud, la seguridad o los derechos fundamentales de las personas físicas, por ejemplo, si no influye de manera sustancial **en el resultado de la toma de decisiones**. Así, **esta excepción se aplica** a los siguientes sistemas de IA:

a) Para hacer una **tarea de procedimiento limitada**.

b) Para mejorar el **resultado de una actividad humana previa.**

c) Para **detectar patrones de toma de decisiones o desviaciones** de patrones de toma de decisiones y no sustituye la valoración humana previa, ni influye en ella.

d) Para hacer una **tarea preparatoria**.

No obstante, existe una **excepción a la excepción**, ya que los sistemas de IA incluidos en la lista de casos de uso **siempre se considerarán de alto riesgo cuando el sistema de IA elabore perfiles** de personas físicas.

¿Qué agentes de la cadena de valor de la IA están sujetos al RIA?

Los agentes que se encuentran sujetos al RIA son aquellos que formen parte de la cadena de valor de sistemas de IA de alto riesgo y riesgo limitado y de modelos de IA de uso general:

- Proveedores de sistemas de IA
- Proveedores de MIAUGs
- Representantes autorizados de proveedores no establecidos en la Unión
- Importadores de sistemas de IA
- Distribuidores de sistemas de IA
- Responsables del despliegue de sistemas de IA
- Proveedores de herramientas y componentes que se integran en los sistemas de IA
- Fabricantes de productos que incluyan sistemas de IA

¿Qué requisitos deben cumplir los sistemas de IA de alto riesgo?

Los sistemas de IA de alto riesgo deben cumplir una serie de requisitos regulatorios (y al proveedor le corresponde velar por que se cumplan):

1. **Sistema de gestión de riesgos**

   Se debe establecer, implantar, documentar y mantener un sistema de gestión de riesgos, y ejecutarse durante todo el ciclo de vida del sistema, con revisiones periódicas.

   Entre otros aspectos, deben analizarse los riesgos que el sistema pueda plantear para la salud, la seguridad o los derechos fundamentales cuando se utilice de conformidad con su **finalidad prevista**, y también cuando se le dé un **uso indebido** que sea **previsible.**

   Los riesgos a gestionar son aquellos que pueden mitigarse o eliminarse mediante el diseño o el desarrollo del sistema o el suministro de información técnica adecuada.

   Los sistemas de IA de alto riesgo deben **someterse a pruebas** antes de su introducción en el mercado o puesta en servicio, para comprobar que funcionan de manera coherente con su finalidad prevista y cumplen los requisitos regulatorios.

2. **Datos y gobernanza de los datos**

   Los sistemas de IA de alto riesgo que implican el entrenamiento de modelos de IA con datos deben desarrollarse a partir de conjuntos de datos (de entrenamiento, validación y prueba) que cumplan determinados criterios de calidad:

1. **Gobernanza y gestión de datos** adecuada para la finalidad prevista del sistema de IA de alto riesgo. Por ejemplo: la **evaluación** de la disponibilidad, la cantidad y la adecuación de los conjuntos de datos necesarios, o el **examen de posibles sesgos** que puedan afectar a la salud y la seguridad de las personas, afectar negativamente a los derechos fundamentales o dar lugar a algún tipo de discriminación prohibida, o la detección de lagunas o deficiencias en los datos.

2. Deben ser **pertinentes, suficientemente representativos y**, en la mayor medida posible, **carecer de errores y estar completos** en vista de su finalidad prevista. Asimismo, deben tener las **propiedades estadísticas adecuadas**.

3. Deben tener en cuenta las **características particulares del entorno** (geográfico, contextual, conductual o funcional) específico en el que está previsto que se utilice el sistema de IA de alto riesgo.

3. Documentación técnica

Se trata de información que está pensada para ser facilitada a las autoridades nacionales competentes y, en su caso, a los organismos notificados[11].

La **documentación técnica de un sistema de IA de alto riesgo**:

1. Debe **elaborarse antes** de su introducción en el mercado o puesta en servicio, y mantenerse actualizada.

2. Debe **redactarse** de modo que (i) **demuestre que** el sistema **cumple los requisitos regulatorios** y (ii) **proporcione** la información necesaria **para que las autoridades nacionales competentes y los organismos notificados** evalúen la conformidad del sistema de IA con dichos requisitos.

3. Debe incluir el contenido mínimo que establece el RIA.

4. Conservación de registros

Los sistemas de IA de alto riesgo deben permitir técnicamente el **registro automático de eventos** (archivos de registro) a lo largo de todo el ciclo de vida del sistema. Por ejemplo, para **detectar situaciones** que puedan dar lugar a que el sistema presente un **riesgo** para la salud, la seguridad o los derechos fundamentales de las personas.

5. Transparencia y comunicación de información a los responsables del despliegue

Los sistemas de IA de alto riesgo deben diseñarse y desarrollarse con la **transparencia** suficiente para que

(i) los **responsables del despliegue interpreten y usen correctamente sus resultados de salida**, y

(ii) el **proveedor y el responsable del despliegue cumplan sus obligaciones** establecidas en el RIA.

Los sistemas de IA de alto riesgo deben adjuntar las instrucciones de uso, con información clara y comprensible para los responsables del despliegue; incluyendo entre otros aspectos:

- la finalidad prevista;

11 Los organismos notificados son las entidades de certificación que realizan la evaluación de la conformidad del sistema de IA cuando se exige que la realice un tercero independiente

- las **características, capacidades y limitaciones del funcionamiento** del sistema de IA de alto riesgo;
- la **información que permita a los responsables del despliegue interpretar los resultados de salida** del sistema de IA de alto riesgo y utilizarlos adecuadamente;
- las **medidas de supervisión humana** (incluidas las medidas técnicas) para facilitar que los responsables del despliegue interpreten los resultados de salida.

6. Supervisión humana

De acuerdo con la premisa de "humano al mando", los sistemas de IA de alto riesgo deben diseñarse y desarrollarse de modo que puedan **ser vigilados por personas físicas durante su uso**, lo que exige una **interfaz humano-máquina** adecuada.

Las medidas de supervisión deben ser proporcionales a los riesgos, al nivel de autonomía y al contexto de uso del sistema de IA de alto riesgo.

La supervisión se garantiza mediante medidas técnicas integradas en el sistema y medidas organizativas que aplica el responsable del despliegue.

El sistema de IA de alto riesgo debe proporcionarse al responsable del despliegue de tal modo **que las personas físicas a quienes se encomiende la supervisión humana** puedan:

a) **entender las capacidades y limitaciones del sistema y vigilar su funcionamiento**;

b) **ser conscientes del "sesgo de automatización"**, para evitar la tendencia a confiar automáticamente en los resultados de salida generados por el sistema;

c) **interpretar correctamente los resultados de salida** del sistema;

d) **no utilizar el sistema en una situación concreta o descartar los resultados de salida** que genere;

e) **detener el funcionamiento** del sistema.

7. Precisión, solidez y ciberseguridad

Los sistemas de IA de alto riesgo deben diseñarse y desarrollarse de modo que alcancen un **nivel adecuado de precisión, solidez y ciberseguridad**.

Antes de introducir el sistema de IA en el mercado o ponerlo en servicio, el proveedor debe demostrar que el sistema cumple los requisitos anteriores, del siguiente modo:

a) Sistemas de IA de los casos de uso de biometría: por la aplicación de estándares de normalización junto con la evaluación de la conformidad por control interno, o bien a través de la evaluación de la conformidad realizada por organismo notificado[12].

b) Sistemas de IA de los casos de uso restantes: a través de la evaluación de la conformidad por control interno.

12 Tercero independiente autorizado oficialmente para realizar esa evaluación de la conformidad

c) Sistemas de IA asociados a productos regulados: a través de la evaluación de la conformidad por 3º independiente, según su legislación aplicable.

¿Cuáles son las obligaciones en el desarrollo de sistemas de IA de alto riesgo?

Los proveedores de sistemas de IA de alto riesgo tienen las siguientes obligaciones:

1) Velar por que sus sistemas cumplan los requisitos regulatorios de los sistemas de IA de alto riesgo (indicados en el apartado anterior).
2) Indicar su nombre, su nombre comercial registrado o marca registrada y su dirección de contacto.
3) Contar con un sistema de gestión de la calidad que cumpla los requisitos establecidos en el RIA.
4) Conservar la documentación relativa al sistema establecida en el RIA durante un período de 10 años.
5) Conservar al menos 6 meses los archivos de registro generados automáticamente por sus sistemas.
6) Someter sus sistemas al procedimiento pertinente de evaluación de la conformidad[13] antes de su introducción en el mercado o puesta en servicio.
7) Elaborar la declaración UE de conformidad[14].
8) Colocar el marcado CE en el sistema.
9) Registrar su sistema y a ellos mismos en la base de datos de la UE.
10) Adoptar las medidas correctoras necesarias cuando consideren que su sistema no es conforme con el RIA o presenta un riesgo para la salud, la seguridad o los derechos fundamentales e informar a los operadores y a las autoridades de vigilancia del mercado.
11) Demostrar, a solicitud de la autoridad nacional competente, la conformidad del sistema de IA de alto riesgo con los requisitos regulatorios.
12) Velar por que el sistema de IA de alto riesgo cumpla requisitos de accesibilidad de conformidad con las Directivas aplicables.
13) Adoptar medidas para que su personal y demás personas que se encarguen del funcionamiento y la utilización de sistemas de IA tengan un nivel suficiente de alfabetización en materia de IA, de manera que conozcan las oportunidades y riesgos que plantea la IA y los perjuicios que puede causar.

13 Por control interno o por evaluación por tercero independiente, según los casos

14 Si el sistema implica el tratamiento de datos personales, debe contener la declaración de que el sistema de IA de ajusta al RGPD

¿Cuáles son las obligaciones en el uso de sistemas de IA de alto riesgo?

Los responsables del despliegue de sistemas de IA de alto riesgo deben cumplir las siguientes obligaciones:

1) Adoptar medidas técnicas y organizativas adecuadas para utilizar el sistema con arreglo a las instrucciones de uso.

2) Encomendar la supervisión humana a personas físicas que tengan la competencia, la formación y la autoridad necesarias.

3) Asegurarse de que los datos de entrada sean pertinentes y suficientemente representativos teniendo en cuenta la finalidad prevista del sistema (en la medida en que tenga el control sobre dichos datos).

4) Vigilar el funcionamiento del sistema basándose en las instrucciones de uso.

5) Informar al proveedor y a la autoridad de vigilancia del mercado si el sistema presenta un riesgo para la salud, la seguridad o los derechos fundamentales, y suspender el uso de ese sistema.

6) Informar al proveedor y a la autoridad de vigilancia del mercado si detecta un incidente grave.

7) Conservar al menos 6 meses los archivos de registro que el sistema genera automáticamente (en la medida en que dichos archivos estén bajo su control).

8) En el caso de los empleadores, informar a los representantes de los trabajadores y a los trabajadores afectados de que estarán expuestos a la utilización del sistema de IA de alto riesgo.

9) Utilizar la información facilitada por el proveedor en las instrucciones de uso para realizar, cuando proceda, la evaluación de impacto relativa a la protección de datos.

10) En el caso de sistemas de IA de alto riesgo incluidos en la lista de casos de uso, que tomen decisiones relacionadas con personas físicas, informar a las personas físicas de que están expuestas a la utilización de los sistemas de IA de alto riesgo.

11) Adoptar medidas para que su personal y demás personas que se encarguen del funcionamiento y la utilización de sistemas de IA tengan un nivel suficiente de alfabetización en materia de IA, de manera que conozcan las oportunidades y riesgos que plantea la IA y los perjuicios que puede causar.

12) Realizar, en algunos casos*, una evaluación del impacto que la utilización del sistema puede tener en los derechos fundamentales, y notificar sus resultados a la autoridad de vigilancia del mercado.

*Responsables del despliegue (i) que sean organismos de Derecho público, o entidades privadas que prestan servicios públicos, o (ii) que usen sistemas de IA para establecer la calificación crediticia de personas físicas o para fijar los precios de seguros de vida y salud.

¿Cuáles son las obligaciones cuando se suministran componentes que se integran en un sistema de IA de alto riesgo?

El proveedor de un sistema de IA de alto riesgo y el tercero que le suministre herramientas o componentes que se integren en dicho sistema, deben suscribir un acuerdo escrito en el que se establezca:

- la información,
- las capacidades,
- el acceso técnico y
- otra asistencia

que sean necesarios, para que el proveedor del sistema de IA de alto riesgo pueda cumplir las obligaciones establecidas en el RIA. Esta obligación no se aplica a terceros que pongan a disposición del público herramientas, o componentes distintos de modelos de IA de uso general, en el marco de una licencia libre y de código abierto.

Obligaciones de transparencia de algunos sistemas de IA - Riesgo de manipulación, suplantación o engaño

Por los riesgos que plantean de manipulación, suplantación o engaño, determinados sistemas de IA están sujetos a una serie de obligaciones de transparencia. Si estos sistemas son de alto riesgo deben cumplir, además, las obligaciones establecidas para los sistemas de alto riesgo.

Algunas obligaciones de transparencia recaen sobre los proveedores del sistema de IA, y otras sobre los responsables del despliegue:

Sistemas de IA	Obligaciones de transparencia	Agente obligado
Destinados a interactuar directamente con personas físicas	Las personas físicas tienen que ser informadas de que están interactuando con sistema de IA (salvo que sea evidente).	Proveedor
Que generen contenido sintético de audio, imagen, vídeo o texto	Los resultados de salida tienen que estar marcados y debe poder detectarse que son sintéticos.	Proveedor
De reconocimiento de emociones De categorización biométrica	Informar del funcionamiento del sistema a las personas expuestas Tratar sus datos personales conforme a la normativa de protección de datos.	Responsable del despliegue
Que generan contenido deep fake (ultrafalsificación) de imagen, audio o vídeo	Hacer público que el contenido es sintético.	Responsable del despliegue

Que generen texto que para informar al público sobre asuntos de interés público	Divulgar que el texto se ha generado de manera artificial.	Responsable del despliegue

Resto de sistemas de IA - Riesgo mínimo

Los proveedores de sistemas basados en inteligencia artificial que quedan fuera del RIA (por no encajar en la definición legal de sistema de IA, por no ser de alto riesgo, o por no requerir medidas de transparencia), pueden aplicar voluntariamente Códigos de Conducta.

La Unión Europea alienta la creación de Códigos de Conducta para la aplicación voluntaria de todos o parte de los requisitos aplicables a los sistemas de IA de alto riesgo, adaptados teniendo en cuenta la finalidad prevista de los sistemas y el menor riesgo que plantean y teniendo en cuenta las soluciones técnicas disponibles y las mejores prácticas del sector[15].

¿Cuándo será de aplicación el RIA?

El RIA entró en vigor el pasado 1 de agosto de 2024. Sin embargo, los plazos de aplicación se darán de forma escalonada en función del sistema de IA y/o su finalidad.

Ilustración 3: Línea del tiempo para la aplicación del RIA

15 Por ejemplo, la tarjeta del modelo (model card).

¿Qué ocurre con los sistemas de IA de alto riesgo y los modelos de IA de uso general anteriores a la aplicación del RIA?

Si el sistema de IA ya introducido en el mercado o puesto en servicio está prohibido, debe ser retirado antes del 02/02/2025.

En el resto de los casos, el RIA se aplica a los operadores de sistemas de IA de alto riesgo que se hayan introducido en el mercado o se hayan puesto en servicio antes del 02/08/2026, si con posterioridad al 02/08/2026, su diseño se modifica de manera significativa.

En cualquier caso, los proveedores y los responsables del despliegue de los sistemas de IA de alto riesgo destinados a ser utilizados por las autoridades públicas deben adoptar las medidas necesarias para cumplir los requisitos y obligaciones del RIA a más tardar el 02/08/2030.

¿Y antes de la aplicación del RIA?: AI Pact

La Comisión está proponiendo el llamado AI Pact, con el fin de que las organizaciones se comprometan voluntariamente a empezar a implementar los requisitos del RIA antes de su aplicación obligatoria. Para ello, las organizaciones que participen en esta iniciativa, se comprometen a realizar las siguientes acciones fundamentales:

- Adoptar una estrategia de gobernanza de la IA para fomentar la adopción de la IA en la organización y trabajar para lograr el cumplimiento futuro del RIA;
- Realizar un mapeo de los sistemas de IA desarrollados o utilizados en áreas que se considerarían de alto riesgo según el RIA;
- Promover la concienciación y la alfabetización en IA del personal y otras personas que se ocupan del despliegue de sistemas de IA en su nombre, teniendo en cuenta su conocimiento técnico, experiencia, educación y capacitación y el contexto en el que se utilizarán los sistemas de IA, y considerando a las personas. o grupos de personas afectadas por el uso de los sistemas de IA.

Además de lo anterior, la AI Pact prevé compromisos adicionales que la organización puede asumir, que varían en función de si la organización desarrolla sistemas de IA o usa sistemas de IA.

Recomendaciones para abordar el cumplimiento del RIA

Recomendaciones para abordar el cumplimiento del RIA

Es recomendable que las organizaciones empiecen a dar pasos para adecuarse al Reglamento Europeo de Inteligencia Artificial.

¿Qué actuaciones pueden ir realizando las organizaciones para adecuarse al RIA?

1) Identificar la IA dentro de la organización.

Es muy posible que ya existan usos y/o desarrollos de inteligencia artificial dentro de la organización. Con el fin de poder gestionar y controlar esos usos y/o desarrollos, previamente es necesario identificarlos e inventariarlos, y para ello debe existir una persona encargada de esta tarea.

2) Analizar si los usos/desarrollos de IA identificados entran o no dentro del ámbito de aplicación del RIA.

Una vez identificados, debemos comprobar si los usos y/o desarrollos de IA en la organización, por su finalidad, están excluidos del ámbito de aplicación del RIA (fines militares, de defensa o de seguridad nacional, finalidad única y específica de la investigación y el desarrollo científicos, actividades de investigación, prueba o desarrollo relativas a sistemas de IA o modelos de IA antes de su introducción en el mercado o puesta en servicio).

3) Analizar si los usos/desarrollos de IA identificados se consideran sistemas de IA a los efectos del RIA.

En este punto analizaremos si los usos de IA entran o no dentro de la definición de sistemas de IA, y, en su caso, si se trata de modelos de IA, o de componentes o soluciones de IA que se integran dentro de sistemas de IA.

4) Analizar si la IA identificada es práctica prohibida / sistema de IA de alto riesgo

5) Analizar el rol de la organización respecto a la IA identificada (proveedor/adquirente/usuario de IA).

Una vez identificados los usos/desarrollos de IA existentes en la organización, debemos analizar cuál es el rol que adopta la organización respecto a los mismos (si usa la IA, si la desarrolla, si entrena modelos de IA, si adquiere elementos de IA, si provee de soluciones de IA), ya que en función de cuál sea el rol asumido, la organización deberá cumplir unas u otras obligaciones.

Por ejemplo, antes de utilizar o de poner en servicio un sistema de IA de alto riesgo en el lugar de trabajo, la organización debe informar a los representantes de los trabajadores y a los trabajadores afectados, de que estarán expuestos a los sistemas de IA de alto riesgo.

6) Establecer un modelo de gobernanza de IA en la organización (roles, funciones, etc.).

Para poder cumplir el marco normativo aplicable a la IA es necesario que la organización disponga de un modelo de gobernanza de la IA, en el que se establezcan los roles necesarios (por ejemplo, Comité de Ética de la IA, Responsable de IA, Responsable de Datos, personas encargadas de la supervisión humana, participación del Delegado de Protección de Datos, Comité/Responsable de Privacidad, etc.), y las funciones de cada uno de ellos.

7) Establecer un sistema de gestión de la IA (incluyendo la gestión de los riesgos de IA).

Es necesario establecer, implantar y mantener un sistema de gestión de la IA dentro de la organización, que incluya un sistema de gestión de los riesgos asociados a la IA y el resto de requisitos establecidos en el RIA. Debe aprobarse una política de alto nivel de IA, y desplegarse a través de procedimientos, políticas y protocolos aplicables en la organización. El modelo de gobernanza de la IA formará parte del sistema de gestión de la IA, así como todo lo relativo al proceso de contratación y cadena de valor de la IA.

8) Elaborar protocolos de uso de la IA (incluyendo la IA generativa).

En el marco del sistema de gestión de la IA, uno de los elementos fundamentales es la elaboración de protocolos de uso ético, legal y responsable de la inteligencia artificial, incluyendo el uso de la IA generativa. La organización tendrá que decidir qué usos están autorizados o no, en su caso, bajo qué condiciones, etc.

9) Formación y concienciación en IA al personal (alfabetización en IA).

Debe promoverse la concienciación y la alfabetización en IA del personal y otras personas que se ocupan del despliegue de sistemas de IA en su nombre, teniendo en cuenta su conocimiento técnico, experiencia, educación y capacitación y el contexto en el que se utilizarán los sistemas de IA, y considerando a las personas. o grupos de personas afectadas por el uso de los sistemas de IA.

10) Establecer hoja de ruta para el cumplimiento.

Partiendo del punto de partida en el que se encuentre cada organización, se trata de establecer un plan de acción para llegar al cumplimiento del RIA, con plazos y responsables.