

# Adaptación a la Regulación de Inteligencia Artificial

*Aspectos clave de la regulación y principales implicaciones para las empresas*

## Panorama normativo

---

## Principales implicaciones para las empresas

---

## Anexo

*Reacciones de los supervisores sobre la IA*

*AI Bill of Rights en EE.UU.*

*Supervisión e iniciativas de IA*

*Abreviaturas*

---

# 1 | Panorama normativo

## Visión general

Los recientes avances en las capacidades de la Inteligencia Artificial (IA) están empujando a los reguladores<sup>1</sup> de todo el mundo a **establecer normativas** y diferentes tipos de directrices para el uso adecuado de la IA



### Visión general

- 1 La Inteligencia Artificial (IA) es la capacidad de una máquina para mostrar **capacidades similares a las humanas**, como el razonamiento, el aprendizaje, la planificación y la creatividad<sup>2</sup>. Un sistema de IA es, por tanto, un software desarrollado con técnicas de IA, que puede, para un conjunto determinado de objetivos definidos por el ser humano, generar resultados como recomendaciones, contenidos, predicciones o decisiones, que influyen en los entornos con los que interactúan<sup>3</sup>.
- 2 Algunas tecnologías de IA existen desde hace más de 50 años, pero los avances en potencia de cálculo, la disponibilidad y capacidad de almacenamiento de enormes cantidades de datos y el desarrollo de nuevos algoritmos, han propiciado **grandes avances en IA en los últimos años**.
- 3 Aunque la mayoría de los sistemas de IA no plantean ningún riesgo y pueden contribuir a resolver muchos retos sociales, en algunos casos generan riesgos que **deben abordarse** para evitar **resultados indeseables**.



### Normativa mundial

- El AI Act en Europa es el primer marco jurídico sobre IA. EE.UU. también ha adoptado un enfoque hacia la IA mediante la publicación de la **Declaración de Derechos de la IA**. Estas dos publicaciones se consideran las normas de referencia más relevantes.
- **Otros países** también están tomando medidas para regular la IA (véase el mapa más abajo).
- A **nivel internacional**, también hay algunas iniciativas. Por ejemplo, la OCDE adoptó algunas recomendaciones sobre IA, la OICV adoptó orientaciones sobre el uso de la IA y la UNESCO adoptó recomendaciones sobre la ética de la IA.



(1) En el contexto de esta nota técnica, el término "regulador" incluye también a los supervisores y a los organismos de normalización reconocidos.









(2) Comisión Europea.

(3) AI Act de la UE. El conjunto de lo que pueden considerarse técnicas de IA también se describe en esta propuesta legislativa, e incluye enfoques de aprendizaje automático, enfoques lógicos y basados en el conocimiento, enfoques estadísticos, estimación bayesiana y métodos de búsqueda y optimización.

# 1 | Panorama normativo

## UE y EE.UU.: principales características de la regulación

Mientras que el **AI Act** establecerá el marco jurídico de IA en Europa, la **Carta de Derechos de la IA de EE.UU.** es un conjunto de principios para ayudar a orientar el diseño, el uso y la implantación de sistemas automatizados.

	 <b>AI Act (Europa)</b>	 <b>AI Bill of Rights (EE.UU.)</b>
 <b>Objetivo</b>	<ul style="list-style-type: none"> <li>Mejorar el funcionamiento del mercado interior estableciendo un <b>marco jurídico uniforme</b>, en particular para el desarrollo, la comercialización y el uso de la IA de conformidad con los valores de la UE.</li> </ul>	<ul style="list-style-type: none"> <li>Ayudar a <b>orientar</b> el <b>diseño</b>, el <b>uso</b> y la <b>implantación</b> de <b>sistemas automatizados</b> para proteger los derechos de los ciudadanos estadounidenses en la era de la inteligencia artificial.</li> </ul>
 <b>Fecha de publicación</b>	<ul style="list-style-type: none"> <li><b>Abril de 2021 (borrador)</b><sup>1</sup></li> </ul>	<ul style="list-style-type: none"> <li><b>Octubre de 2022 (borrador)</b></li> </ul>
 <b>Alcance</b>	<ul style="list-style-type: none"> <li>Proveedores de sistemas de IA</li> <li>Usuarios del sistema de IA</li> <li>Distribuidores, importadores y operadores de sistemas de IA y personas afectadas ubicadas en la UE</li> </ul>	<ul style="list-style-type: none"> <li>Diseñadores</li> <li>Desarrolladores</li> <li>Implantadores de sistemas automatizados</li> </ul>
 <b>Exigibilidad</b>	<ul style="list-style-type: none"> <li>Ley <b>obligatoria</b>. Cuando se apruebe, será inmediatamente aplicable sobre los 27 países de la Unión Europea.</li> </ul>	<ul style="list-style-type: none"> <li>Libro Blanco <b>voluntario</b> (+ Orden Ejecutiva del Presidente de EE.UU.)</li> </ul>
 <b>Contenido principal</b>	Clasifica las aplicaciones de IA en niveles de riesgo: <ul style="list-style-type: none"> <li><b>Riesgo inaceptable</b> (prácticas prohibidas)</li> <li><b>Alto riesgo</b> (sujeto a una serie de requisitos y obligaciones para acceder al mercado de la UE)</li> <li><b>Riesgo bajo o mínimo</b></li> <li>Algunos sistemas de IA suponen un <b>riesgo de transparencia</b> y/o <b>riesgo sistémico</b> y están sujetos a requisitos adicionales</li> </ul>	<b>5 principios:</b> <ul style="list-style-type: none"> <li>Sistemas seguros y eficaces</li> <li>Protección contra la discriminación algorítmica</li> <li>Protección de datos</li> <li>Aviso y explicabilidad</li> <li>Alternativas humanas, consideración y fallback</li> </ul>
 <b>Próximos pasos</b>	<ul style="list-style-type: none"> <li>Trílogo acordado en <b>Dic'23</b>, aprobación en el Parlamento en Mar'24, entrada en vigor, en general, 2 años después.</li> </ul>	<ul style="list-style-type: none"> <li>No hay calendario oficial para las próximas publicaciones</li> </ul>

(1) El primer borrador fue publicado por la CE en abril de 2021. La propuesta ha sido debatida por los legisladores (PE, CE, Consejo). Hay 3 versiones diferentes que deben consolidarse: [comparison of versions](#). La versión definitiva se ha publicado por el Parlamento Europeo en Marzo de 2024, tras su aprobación.

# 1 | Panorama normativo

## Otras publicaciones

Los avances en las capacidades de la IA impulsan los esfuerzos reguladores mundiales para establecer directrices éticas y procedimientos de seguridad, con iniciativas como la **Declaración de Bletchley** y una **Orden Ejecutiva del Presidente de EE.UU.**

*Declaración Bletchley (02/11/2023)*



### Los cinco objetivos fijados en la cumbre son:

1. **Comprensión compartida** de los riesgos que plantea la IA en las fronteras y la necesidad de actuar
2. Proceso de avance para la **colaboración internacional** en materia de seguridad de la IA en las fronteras, incluyendo la mejor manera de apoyar los marcos nacionales e internacionales
3. **Definir las medidas adecuadas** que deben adoptar las organizaciones individuales para **aumentar la seguridad de la IA en las fronteras**
4. Áreas de posible **colaboración** en la **investigación sobre la seguridad** de la IA, incluyendo la **evaluación de las capacidades de los modelos** y el desarrollo de **nuevas normas para apoyar la gobernanza**.
5. Mostrar cómo garantizar el desarrollo seguro de la IA permitirá que ésta **se utilice para el bien en todo el mundo**.

*Orden Ejecutiva del Presidente Biden (30/10/2023)*



### La Orden Ejecutiva ordena las siguientes acciones:

1. **Nuevas normas de seguridad** y protección de la IA (USAISA, requisito de aprobación de la IA en las fronteras)
2. **Protección de la privacidad de las personas**
3. **Fomento de la equidad** y los **derechos civiles**
4. **Defender a los consumidores, los pacientes y los estudiantes**
5. Apoyo a los trabajadores
6. **Promover la innovación y la competencia**
7. Promover el **liderazgo estadounidense en el extranjero**
8. Garantizar un **uso responsable y eficaz de la IA por parte del Gobierno**

- **Especial atención a la IA en las fronteras** → fuerte **riesgo de uso indebido** o **problemas de control** que pueden no estar en consonancia con la intención humana original
- **Los principales riesgos** reconocidos están relacionados con **la ética, la parcialidad, la explicabilidad, la necesidad actual de regulación y la necesidad de supervisión humana**.
- **En conclusión**, los principales puntos planteados **confirman los principios establecidos en el AI Act de la UE**.

# 1 | Panorama normativo

## AI Act de la UE - Principales cambios en la versión consolidada

El AI Act de la UE ha sido sometida al escrutinio de todas las partes interesadas pertinentes, lo que ha dado lugar a la identificación de 14 modificaciones clave, que se han dado a conocer al público en general en 2024

"Sistema de IA es un sistema basado en máquinas diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras su despliegue y que, para objetivos explícitos o implícitos, infiere, a partir de la entrada que recibe, cómo generar salidas tales como predicciones, contenidos, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales" (Art. 3.1)

### Definición de sistema de IA



Obligación de que quienes desplieguen IA de alto riesgo realicen una **evaluación de impacto sobre los derechos fundamentales** (art. 27)

Definición de IA de alto riesgo basada en su uso (anexo III) + sistemas de IA utilizados como componentes de productos (vehículos, aviación...)

Desplazamiento de las responsabilidades de los "proveedores" a lo largo de la cadena de la IA (art. 25)

### IA de alto riesgo



Obligaciones de los **proveedores** de GPAI (art. 53)

Obligaciones adicionales para los **proveedores** de modelos GPAI con **riesgo sistémico** (art. 55, anexo XI)

Obligaciones adicionales para los **proveedores e implantadores de modelos GPAI** (o sistemas de IA) con **riesgo de transparencia** (art. 50)

**Modelos GPAI<sup>(1)</sup>**

### Identificación biométrica

La identificación biométrica como práctica prohibida (Art. 5.1 + Anexo III)

Salvaguardias adicionales a la identificación biométrica a distancia (art. 26)

### Pruebas de IA de alto riesgo en condiciones reales

Definición de salvaguardias para las pruebas de IA de alto riesgo fuera de los entornos aislados (art. 60)

Excepciones definidas para las fuerzas del orden

### Entrada en vigor y sanciones

El acuerdo de compromiso prevé 24 meses con excepciones como 6, 12 y 36 meses en determinados casos. (Art. 111)

Sanciones fijadas en 35 millones de euros o el 7% del volumen de negocios anual (art. 99)

# 1

## Panorama normativo

### AI Act de la UE - Definición de la IA

El AI Act impone **obligaciones importantes a los sistemas de IA de alto riesgo**, que se definen como aquellos que suponen un riesgo elevado para la salud, la seguridad, los derechos fundamentales o el medio ambiente

#### Definiciones del AI Act



#### Definición de Sistema de IA

“Los sistemas de IA y los modelos fundacionales” se definen<sup>1,2</sup> como:

**Art. 3:** “un sistema basado en máquina que está diseñado para operar con distintos niveles de autonomía y que puede, para objetivos explícitos o implícitos, **generar resultados como predicciones, recomendaciones o decisiones que influyen en entornos físicos o virtuales**”.

**Art. 3 (63) “modelo de IA de propósito general”:** un modelo de IA, incluso cuando se ha entrenado con una gran cantidad de datos utilizando la autosupervisión a escala, que muestra una generalidad significativa y es capaz de realizar de forma competente una amplia gama de tareas distintas, independientemente de la forma en que se comercialice el modelo, y que puede integrarse en una variedad de sistemas o aplicaciones posteriores. Esto no incluye los modelos de IA que se utilizan antes de su comercialización para actividades de investigación, desarrollo y creación de prototipos”.



#### Definición de Sistema de IA de riesgo alto

**Art. 3 (2) “Riesgo”** significa “la **combinación de la probabilidad de que ocurra un daño y la gravedad de ese daño**”;

**Art 6 (1)** “un sistema de IA se considerará de alto riesgo cuando reúna las dos condiciones que se indican a continuación:

- a) que el sistema de IA esté destinado a ser utilizado como componente de seguridad de un producto que entre en el ámbito de aplicación de los actos legislativos de armonización de la Unión enumerados en el anexo I, o que el propio sistema de IA sea uno de dichos productos;
- b) que el producto del que el sistema de IA sea componente de seguridad con arreglo a la letra a), o el propio sistema de IA como producto, deba someterse a una evaluación de la conformidad

**Art 6 (2)** “se considerarán de alto riesgo los sistemas de IA contemplados en el anexo III.”

**Art. 6 (3)**, un sistema de IA 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

1. La definición se ha alineado con la acordada por la Organización para la Cooperación y el Desarrollo Económico (OCDE) en la versión del documento del Parlamento Europeo. Obsérvese que los restantes borradores del AI Act sí contemplan en su ámbito de aplicación los enfoques lógicos y basados en el conocimiento.
2. Excepciones dentro del AI Act: Sistemas de IA utilizados con fines militares, autoridades públicas de terceros países u organizaciones internacionales.

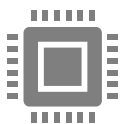
# 1

## Panorama normativo

### AI Act de la UE - Definición de la IA

El AI Act impone algunas obligaciones a los modelos de propósito general, que plantean riesgo sistemático o riesgo de transparencia

#### Definiciones del AI Act



#### Definición de riesgo Sistemático

**Art 51** “Un modelo de IA de uso general se clasificará como modelo de IA de uso general con riesgo sistémico si reúne alguno de los siguientes requisitos:

- a) tiene capacidades de gran impacto evaluadas a partir de herramientas y metodologías técnicas adecuadas, como indicadores y parámetros de referencia;
- b) con arreglo a una decisión de la Comisión, adoptada de oficio o a raíz de una alerta cualificada del grupo de expertos científicos, tiene capacidades o un impacto equivalente a los establecidos en la letra a), teniendo en cuenta los criterios establecidos en el anexo XIII

Se presumirá que un modelo de IA de uso general tiene capacidades de gran impacto con arreglo al apartado 1, letra a), cuando la cantidad acumulada de cálculo utilizada para su entrenamiento, medida en FLOP, sea superior a  $10^{25}$ .



#### Obligaciones de transparencia

**Art 50** “Los proveedores garantizarán que los sistemas de IA destinados a interactuar directamente con personas físicas se diseñen y desarrollen de forma que las personas físicas de que se trate **estén informadas de que están interactuando con un sistema de IA**”.

Los proveedores de sistemas de IA, entre los que se incluyen los sistemas de IA de uso general, que generen contenido sintético de audio, imagen, vídeo o texto, velarán por que la información de salida del sistema de IA esté marcada en un formato legible por máquina y que sea posible detectar que ha sido generada o manipulada de manera artificial



# 1 | Panorama normativo

## AI Act de la UE - Supervisión de la IA

El AI Act de la UE establece **requisitos estrictos de supervisión** para los sistemas de IA de alto riesgo con el fin de salvaguardar los derechos y la seguridad, con la participación de los actuales supervisores de servicios financieros.

Comisión  
Europea

Parlamento  
Europeo

Consejo  
Europeo

*Acuerdo final*

**General:** ...cada Estado miembro debe designar al menos una **autoridad nacional de supervisión** con el fin de supervisar la aplicación y ejecución del presente Reglamento. [...]

**Servicios financieros:** Para garantizar la aplicación y ejecución coherentes de las obligaciones previstas en el presente Reglamento, así como de las normas y los requisitos oportunos de los actos jurídicos de la Unión relativos a los servicios financieros, se ha de designar a las autoridades competentes de supervisar y ejecutar dichos actos jurídicos, en particular las autoridades competentes definidas en el Reglamento (UE) n.º 575/2013 del Parlamento Europeo y del Consejo<sup>49</sup> y las Directivas 2008/48/CE<sup>50</sup>, 2009/138/EC<sup>51</sup>, 2013/36/UE<sup>52</sup>, 2014/17/UE<sup>53</sup> y (UE) 2016/9754 del Parlamento Europeo y del Consejo, dentro de sus respectivas competencias, como las autoridades competentes encargadas de supervisar la aplicación del presente Reglamento, [...]

Supervisión  
de la IA  
(AI Act)

# 1

## Panorama normativo

### AI Act de la UE - Posición del BCE

En el sector bancario, el BCE como supervisor prudencial de las entidades de crédito, publicó en diciembre de 2021 su posición sobre los requisitos del AI Act



#### Observaciones generales

- El BCE celebra el objetivo del reglamento propuesto y **reconoce la importancia de establecer requisitos armonizados para los sistemas de IA**, especialmente en el sector bancario.
- El reglamento propuesto debería entenderse sin perjuicio de las obligaciones prudenciales más específicas o estrictas de las entidades de crédito establecidas en la regulación sectorial y complementadas por orientaciones de supervisión (por ejemplo, el control efectivo de la externalización, tal como se especifica en la DRC).
- **El papel del BCE en virtud del reglamento propuesto debería aclararse:** (1) las competencias de supervisión prudencial del BCE en general, y en relación con la vigilancia del mercado y la evaluación de la conformidad; y 2) la aplicación del reglamento propuesto al desempeño de las funciones que el Tratado atribuye al BCE.
- Con arreglo al reglamento propuesto, las actividades de calificación crediticia que utilicen sistemas de inteligencia artificial estarían sujetas a los requisitos mínimos aplicables a los sistemas de inteligencia artificial de alto riesgo.
- En cuanto a la calificación crediticia, **el BCE propone retrasar la entrada en vigor hasta que se especifiquen las condiciones para verificar la conformidad con los requisitos aplicables** y definir cuándo debe considerarse que los sistemas de IA "son puestos en servicio por proveedores a pequeña escala para su propio uso".

3.2. Under the proposed regulation <sup>(43)</sup>, the vast majority of credit scoring activities making use of AI systems would be automatically subjected to the horizontal minimum requirements imposed on high-risk AI systems. Consequently, several activities, including acquisition targeting for marketing, collections modelling and standard credit scoring models (e.g. a scorecard using logistical regression), would need to comply with the same requirements. To enhance clarity in supervisory expectations and in line with the ECB's technology-neutral approach, it is suggested that AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score and which leverage on the standalone use of linear or logistic regression or decision trees under human supervision should not be classified as high-risk AI systems, provided that the impact of such approaches to the assessment of natural persons' creditworthiness or credit score is minor.

# 1

## Panorama normativo

### Principios de la regulación sobre IA

Las propuestas de regulación de la IA pretenden abordar los riesgos potenciales, las limitaciones y los problemas éticos asociados a los modelos de IA, al tiempo que promueven su **desarrollo, aplicación y uso responsables**

#### Resumen de requerimientos regulatorios de IA

##### Transparencia e interpretabilidad

- Comunicación a los clientes
- Explicabilidad del modelo
- Derecho a una explicación

##### Equidad y prevención de sesgos

- Atributos protegidos
- Identificación de sesgos
- Corrección de sesgos
- Retos: atributos no observables, correlaciones

##### Robustez y fiabilidad

- Evaluación del rendimiento
- Robustez (ataques de adversarios, sensibilidad...)
- Human-in-the-loop
- Human fallback



##### Responsabilidad

La responsabilidad recae en el usuario de IA

##### Privacidad y seguridad

Cómo la IA gestiona y protege los datos de los usuarios

- Riesgos de divulgación no autorizada o uso indebido
- Protección de datos, consentimiento, anonimización
- Requisitos de ciberseguridad

##### Ética y sostenibilidad

Evaluar las implicaciones éticas de los modelos de IA

- Contar con expertos en ética de la IA
- Medir el impacto en el medio ambiente (por ejemplo, emisiones)

##### Riesgos y gobierno

Evaluación de todos los riesgos de IA

- Auditoría y supervisión reglamentaria

# 1 | Panorama normativo

## Una evaluación de riesgos de 360 grados

La evaluación de riesgos exigida por el AI Act abarca todos los riesgos incluidos en un sistema de IA de alto riesgo


Evaluación de riesgos	Contenido / Requisitos de presentación
<b>Riesgo del Modelo</b>	<p>(AI Act, Art., 8, 9, 10, 13, 14, 15, 29) Informe de validación del modelo que contenga una evaluación de la solidez técnica (por ejemplo, pruebas de rendimiento según los procedimientos de MV), incluidos tanto los requisitos tradicionales del modelo como los requisitos específicos de la IA sobre interpretabilidad y explicabilidad del modelo, imparcialidad y sesgo, y supervisión humana.</p> <p>(Art. 4a) Opinión sobre la clasificación de riesgos de la IA y la clasificación de riesgos de los modelos.</p>
<b>Riesgo operacional</b>	<p>(AI Act, Art. 8, 9, 12) Evaluación de los riesgos operativos y de aplicación, incluido el mantenimiento automatizado de registros</p>
<b>Riesgo del modelos de terceros</b>	<p>(AI Act, Art. 8 y 9) Evaluación del riesgo de proveedores para modelos de terceros, como parte de la evaluación de riesgos establecida, aplicada, documentada y a lo largo de todo el ciclo de vida.</p>
<b>Calidad de los datos</b>	<p>(AI Act, Art., 8 y 10) Evaluación de la gobernanza de los datos apropiada para la finalidad prevista del modelo (transparencia, opciones de diseño, hipótesis, limpieza de datos, disponibilidad y adecuación de los datos, sesgos, lagunas de datos...), posibles usos indebidos del modelo, prácticas de gestión de datos.</p>
<b>Protección y privacidad de los datos</b>	<p>(AI Act, Art. 8 y 10) Evaluación del cumplimiento de las normas vigentes sobre privacidad y protección de datos.</p>
<b>Legal y cumplimiento</b>	<p>(AI Act, Art. 8, 9) Evaluación de los riesgos jurídicos (por ejemplo, derechos de autor), como parte de la evaluación de riesgos.</p>
<b>Ética</b>	<p>(AI Act, Art. 8 y 29a) Evaluación de la ética (uso ético del modelo), seguimiento y evaluación de las repercusiones a largo plazo sobre el individuo, la sociedad y la democracia.</p>
<b>Ciberseguridad / Riesgo de IT</b>	<p>(AI Act, Art 8 y 15) Evaluación de la seguridad de los datos y de la infraestructura informática.</p>
<b>Documentación del modelo</b>	<p>(AI Act, Art. 8, 11, 52, Annex IV) Documentación técnica del modelo, incluida su descripción, diseño y desarrollo, seguimiento, cambios de modelo, normas legales, entre otros.</p>

Requisitos para todos los sistemas de IA y modelos fundacionales

- Solidez Técnica, robustez y ciberseguridad
- Transparencia y trazabilidad
- Agencia humana y supervisión
- Privacidad y protección de datos
- No discriminación y equidad
- Derechos fundamentales

Requisitos adicionales para los sistemas de IA de alto riesgo

- Documentación técnica
- Sistema de gestión de riesgos
- Datos y gobernanza de datos
- Mantenimiento de registros



Panorama normativo

---

Principales implicaciones para las empresas

---

Anexo

*Reacciones de los supervisores sobre la IA*

*AI Bill of Rights en EE.UU.*

*Supervisión e iniciativas de IA*

*Abreviaturas*

---

# 2 Principales implicaciones para las empresas

## Implicaciones de la regulación de la IA

Adoptar la inteligencia artificial en una organización y garantizar el cumplimiento de la normativa sobre IA conlleva una serie de implicaciones en varias dimensiones

### Contexto, regulación y supervisión

1. **Control de la normativa:** Seguimiento de la normativa aplicable (GDPR, EU AI Act, US AI Bill of Rights, etc.) y de las actividades de supervisión (BCE, AESIA, etc.).
2. **Gap analysis normativo:** Análisis gap con respecto a la normativa local y mundial sobre IA, y diseño y aplicación de planes de corrección.
3. **Presentación de informes y relaciones de supervisión:** Coordinación de la relación con los supervisores, requisitos de información e inspecciones reglamentarias, y preparación de la documentación del sistema de IA.

### Estrategia de IA

1. **Estrategia:** Posicionamiento de la IA, apetito por el riesgo de la IA (modelo, legal, reputacional...).
2. **Adopción de la IA:** Análisis de la situación actual, Plan de adopción de la IA.

### Marco de la IA


1. **Gobernanza:** Comités y grupos de trabajo para adaptarse a la normativa sobre IA y supervisar el riesgo. PMO para coordinar esfuerzos y adaptarse cambios normativos.
2. **Funciones y responsabilidades:** Nuevas figuras creadas o adaptadas dentro de la organización, como el Director de IA / Director de IA y Tecnología, que cubran todas las etapas del ciclo de vida de un modelo de IA.
3. **Políticas y procedimientos:** Revisión o desarrollo de políticas y procedimientos internos que regulen los datos, el uso y la definición de sistemas de IA, incl. marcos de modelización y validación, directrices de transparencia y ética, procedimientos de datos, etc., y medidas de supervisión.
4. **Gestión de riesgos:** Identificación inicial, inventario y evaluaciones de riesgos y clasificación de los usos de los modelos de IA (por ejemplo, riesgo inaceptable / alto / limitado / mínimo, siguiendo el AI Act de la UE).
5. **(Re)capacitación en IA:** Potenciación de las competencias operativas en IA a todos los niveles de la organización.

### Desarrollo y uso de la IA

1. **Selección de casos de uso empresarial:** Selección de casos de uso de IA y procedimiento de inversión en pilotos (MVPs).
2. **Desarrollo y validación de sistemas de IA:** Desarrollo y validación de modelos de IA cumpliendo los requisitos normativos (aprendizaje supervisado y no supervisado, GenAI y NLP...), automatización de informes y herramientas de modelos de IA, definición de metodologías de trabajo (por ejemplo, Agile) y perfiles (por ejemplo, data scientist, data engineer, data analyst...).
3. **Supervisión:** Definición y aplicación de medidas de control y supervisión humana adecuadas a los sistemas de IA. Implantación y revisión del mantenimiento de registros y diarios de los sistemas de IA. Revisión del uso correcto de los sistemas de IA desplegados .
4. **Interpretabilidad y equidad:** Implementación de marcos de IA explicable (XAI), tratamiento de posibles sesgos y diseño de métricas de equidad, sesgos de modelo, y no discriminación.

### Datos, infraestructura y arquitectura

1. **Roadmap tecnológico:** Elaboración de un roadmap tecnológico para el desarrollo, despliegue y uso de la IA en función del nivel de adopción de la entidad.
2. **Infraestructura:** Despliegue en la nube para permitir la computación paralela y la escalabilidad, y el uso de soluciones Big Data. Requisitos de proveedores externos de IA.
3. **Arquitectura:** Diseño de arquitecturas que permitan el uso eficiente de la infraestructura por los diferentes departamentos / usuarios, gestión de recursos (por ejemplo, uso de GPUs, almacenamiento), uso de arquitecturas innovadoras (por ejemplo, aprendizaje federado).
4. **Datos y seguridad informática:** Revisión del marco de datos (BCBS239, GDPR) para incorporar requisitos de IA (p.ej., transparencia, privacidad, seguridad, sesgo) y facilitar el uso de datos no estructurados (p.ej., creación de lagos de datos), requisitos de seguridad informática.
5. **MLOps y código abierto:** Adopción de la filosofía MLOps y migración hacia lenguajes de código abierto (p.ej., Python), uso de modelos de código abierto (p.ej., HuggingFace, Meta).
6. **Sandbox:** Creación de un sandbox para el desarrollo de prototipos de IA.
7. **LLMs:** Adaptación de los modelos y selección de infraestructuras donde hacer fine-tuning de un LLM, privacidad, trazabilidad, integración con otros sistemas, tiempo de respuesta, etc.



Panorama normativo

---

Principales implicaciones para las empresas

---

**Anexo**

*Reacciones de los supervisores sobre la IA  
AI Bill of Rights en EE.UU.  
Supervisión e iniciativas de IA  
Abreviaturas*

---

# A Anexo 1

## Reacciones de los supervisores sobre la IA: Global

A escala mundial, las recomendaciones de la OCDE destacan como los primeros principios suscritos por los gobiernos, y otros organismos también han emitido orientaciones y recomendaciones



Global

### Recomendación del Consejo sobre Inteligencia Artificial | OCDE | Mayo. 2019

G FV NB

Esta recomendación se centra en dos pilares. Por un lado, establece una serie de principios para la gestión responsable de una IA fiable: i) crecimiento integrador, desarrollo sostenible y bienestar; ii) valores centrados en el ser humano y equidad; iii) transparencia y explicabilidad; iv) solidez, seguridad y protección; y v) rendición de cuentas. Por otro lado, establece recomendaciones para la integración de la IA en las políticas nacionales y fomenta la cooperación internacional de los gobiernos para una IA segura.

### Uso de IA y ML por intermediarios del mercado y gestores de activos | IOSCO | Sep. 2021

I FV NB

Las guías propuestas pretenden ayudar a sus miembros a regular y supervisar el uso de la IA y el ML por parte de los intermediarios del mercado y los gestores de activos. También describe cómo los reguladores están abordando los retos creados por la IA y el blanqueo de capitales y las orientaciones publicadas por los organismos supranacionales en este ámbito.

### Ética de la Inteligencia Artificial | UNESCO | Nov. 2021

G FV NB

Un conjunto de valores, principios y recomendaciones sobre políticas de actuación, con el objetivo de sentar las bases para que los sistemas de IA trabajen por el bien de la humanidad, las personas, las sociedades y el medio ambiente y los ecosistemas, y para prevenir daños, garantizando que contribuyan a un mundo más inclusivo, sostenible y pacífico.

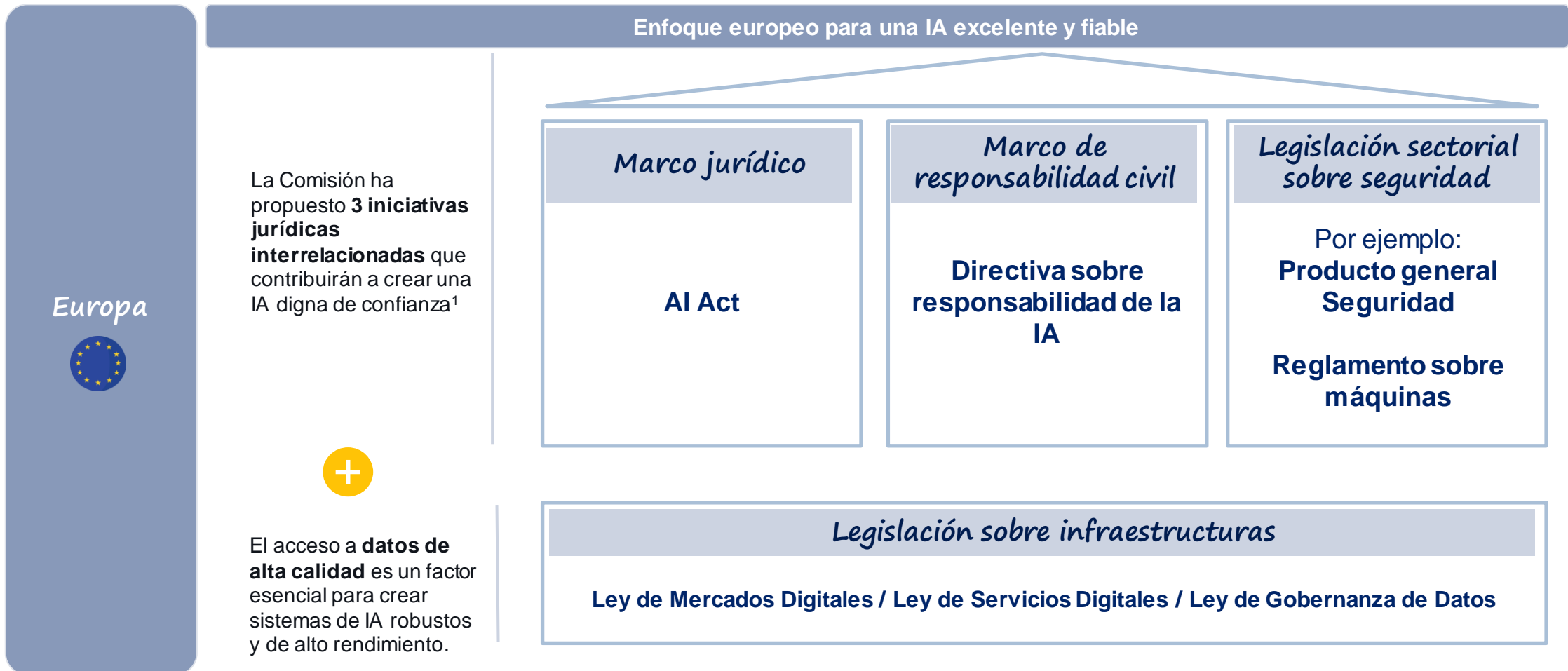


# A

## Anexo 1

### Reacciones de los supervisores sobre la IA: Europa (incl. Reino Unido)

La Estrategia Europea de Inteligencia Artificial pretende convertir a la UE en un centro mundial de IA y garantizar que la IA esté centrada en el ser humano y sea digna de confianza.



(1) Además, en enero de 2023 el Comité sobre Inteligencia Artificial del Consejo de Europa ha publicado un proyecto de Convenio sobre Inteligencia Artificial, derechos humanos, democracia y Estado de Derecho para establecer principios, normas y derechos destinados a garantizar que el diseño, el desarrollo y la aplicación de sistemas de inteligencia artificial sean plenamente coherentes con el respeto de los derechos humanos, el funcionamiento de la democracia y la observancia del Estado de Derecho. Este documento deberá negociarse en la Comisión.

# A

## Anexo 1

### Reacciones de los supervisores sobre la IA: Europa (incl. Reino Unido)

En Europa, la Comisión ha propuesto algunas iniciativas jurídicas interrelacionadas que contribuirán a crear una IA digna de confianza y a abordar los derechos fundamentales

Europa



#### AI Act<sup>1</sup> | CE | Abr. 2021

G D B

El borrador del Reglamento sobre Inteligencia Artificial pretende **garantizar un alto nivel de confianza en la IA y sus aplicaciones**, al tiempo que sienta las bases para la innovación. Propone una clasificación de las prácticas de IA en los siguientes niveles: i) prácticas prohibidas; ii) sistemas de IA de alto riesgo; iii) sistemas de IA de riesgo bajo o mínimo. Además, incluye obligaciones de transparencia para determinados sistemas de IA que i) interactúan con seres humanos; ii) se utilizan para detectar emociones; o iii) generan o manipulan contenidos.

#### Informe de seguimiento sobre Machine Learning en el enfoque IRB | EBA | Ago. 2023

I FV NB

El Informe de seguimiento de la EBA ofrece una visión general sobre los casos actuales del uso de técnicas de Machine Learning en el contexto de los modelos basados en calificaciones internas. Además, analiza la interacción del uso de estas técnicas en los modelos de riesgo de crédito con otros dos marcos jurídicos: el Reglamento General de Protección de Datos y la Ley de Inteligencia Artificial.

#### Directiva sobre responsabilidad civil por IA | CE | Sep. 2022

G D NB

La propuesta de Directiva sobre la adaptación de las normas de responsabilidad civil **extracontractual** a la IA tiene los siguientes objetivos: i) **adaptar las normas de responsabilidad civil extracontractual** a la IA; ii) **promover la adopción de la IA** y abordar los riesgos asociados a su uso; iii) **identificar y abordar los retos específicos que plantea la IA** a las normas de responsabilidad civil existentes; iv) establecer un **régimen de responsabilidad civil** para la IA que sea adecuado y eficaz; y v) garantizar que las víctimas de daños causados por productos y servicios basados en la IA tengan acceso a **una indemnización justa y eficaz**.

#### Reglamento relativo a la seguridad general de los productos (RSPG) | CE | Mayo 2023

I FV B

Pretende abordar los **retos en materia de seguridad de los productos que plantean las tecnologías emergentes**, incluido el uso de la IA y los dispositivos conectados, y establecer obligaciones claras para los mercados online, que los consumidores utilizan cada vez más para sus compras en Internet.

(1) El BCE ha publicado su posición respecto al AI Act. Véase [el Anexo 3](#).

# A

## Anexo 1

### Reacciones de los supervisores sobre la IA: Europa (incl. Reino Unido)

Esta normativa se complementa con otras leyes para abordar la infraestructura y el acceso a datos y sistemas

Europa



#### Reglamento sobre máquinas | PE y Consejo | Jun. 2023

I FV B

Reglamento que introduce un nuevo **marco jurídico en la industria europea de maquinaria y equipos industriales**. Los fabricantes, importadores y distribuidores de todo tipo de maquinaria tendrán que cumplir nuevas y amplias obligaciones. La normativa introduce nuevos requisitos de seguridad para las máquinas autónomas, la colaboración entre personas y máquinas y, por primera vez, el uso seguro de sistemas de IA en las máquinas.

#### Ley de Mercados Digitales (LMD) | CE | Oct. 2022

G FV B

La LMD establece una serie de **criterios objetivos claramente definidos** para identificar a los "gatekeepers". Los "gatekeepers" son las grandes plataformas digitales que prestan servicios básicos de Plataforma: motores de búsqueda en línea, tiendas de aplicaciones o servicios de mensajería. Los "gatekeepers" tendrán que cumplir las obligaciones y prohibiciones enunciadas en la LMD. Su impacto en la esfera digital podría tener implicaciones en el uso y regulación de las tecnologías de IA en el contexto de las plataformas en línea y la economía digital en la UE.

#### Ley de Servicios Digitales (LSD) | CE | Oct. 2022

G FV B

La LSD establece **normas jurídicas para las plataformas en línea que operan en la UE**, incluidas las plataformas de redes sociales, los mercados en línea y los motores de búsqueda. Su objetivo es hacer a las plataformas en línea más responsables de los contenidos que albergan y reforzar los derechos y protecciones de los usuarios ante la posible amenaza de la IA.

#### Ley de Gobernanza de Datos | PE y Consejo | Mayo 2022

G FV B

Apoya la creación y el desarrollo de **espacios comunes de datos en Europa**, es decir, un mercado interior de datos en el que éstos puedan utilizarse con independencia de su ubicación física de almacenamiento en la UE, en conformidad con la legislación aplicable, pudiendo llegar a ser fundamental para el rápido **desarrollo de las tecnologías de IA**.

# A

## Anexo 1

### Reacciones de los supervisores sobre la IA: Europa (incl. Reino Unido)

En España, un Real Decreto estableció un Sandbox para el cumplimiento de la ley de IA, mientras que Alemania y Reino Unido han publicado libros blancos, e Italia ha regulado el uso de ChatGPT



#### Real Decreto por el que se crea un Sandbox para el cumplimiento de la ley de IA | MINECO | Mayo 2023

G D B

El objetivo es **crear un entorno de pruebas controlado** para ensayar las medidas previstas en la propuesta del Reglamento de la UE sobre IA (AI Act). Se cribarán los sistemas de inteligencia artificial que impliquen riesgos a la salud, la seguridad y los derechos fundamentales de las personas, con el fin de diseñar los principios que rijan su diseño, validación y monitorización para mitigar tales riesgos. Véase [el Anexo 2](#).

#### Real Decreto por el que se aprueba la constitución de la Agencia Española de Supervisión de la IA | Gob. Español | Ago. 2023

I FV B

El objetivo de esta Agencia es **supervisar el uso de los sistemas de IA** para proteger los derechos fundamentales y minimizar riesgos, en colaboración con las autoridades nacionales y europeas. Véase [el Anexo 2](#).



#### Libro Blanco sobre Ética e Inteligencia Artificial en Alemania | Ministerio de Economía y Energía | Sep. 2020

G FV NB

El objetivo del proyecto era **presentar la situación actual de normalización** en el campo interdisciplinario de la IA y la ética e identificar posibles campos de acción futuros para la normalización. El proyecto examina las interrelaciones entre ética e IA y el papel que pueden desempeñar en este contexto las normas y reglas técnicas. El proyecto se centra en los ámbitos de las máquinas y los vehículos autónomos.



#### Libro blanco de IA para impulsar el crecimiento | Gobierno del Reino Unido | Mar. 2023

G D NB

El libro blanco esboza **5 principios claros** que los reguladores, como la Comisión de Igualdad y Derechos Humanos y la Autoridad de Competencia y Mercados, deben tener en cuenta para facilitar el uso seguro e innovador de la IA en las industrias que supervisan: i) seguridad; ii) protección y robustez; iii) transparencia y explicabilidad; iv) equidad; v) responsabilidad y gobernanza; y vi) contestabilidad y reparación.



#### Medidas para la gestión de los servicios de IA generativa | GPDP | Mar. 2023

G D B

El Garante para la Protección de Datos Personales (GPDP) prohibió el uso de Chat GPT el 30 de Marzo de 2023. Fue emitido para **garantizar la protección de datos personales** y solicitaba la aplicación de una serie de medidas concretas a OpenAI con plazo límite el 30 de abril. Poco más de un mes después, la empresa había aplicado las medidas solicitadas y Chat GPT volvió a estar disponible en Italia.

# A Anexo 1

## Reacciones de los supervisores sobre la IA: América

En EE.UU. destacan dos iniciativas no vinculantes: el AI Bill of Rights y el AI Risk Management Framework, mientras que Canadá y México cuentan también con leyes vinculantes para regular los sistemas de IA



### AI Bill of Rights | WH | Oct. 2022

G FV NB

Establece **cinco principios o derechos de los ciudadanos en relación con la IA**, que incluyen sistemas seguros y eficaces, protección contra la discriminación de algoritmos, privacidad de los datos, notificación y explicación, y evaluación y corrección por un humano en caso de fallo de la IA (fallback). Estos principios incluyen la explicabilidad de los modelos de IA, que requiere una documentación en lenguaje sencillo, además de explicaciones técnicas válidas, significativas y útiles, y notificaciones de uso claras, oportunas, comprensibles y accesibles.

### AI Risk Management Framework | NIST | Ene. 2023

G FV NB

Este marco pretende **ofrecer un medio** a las organizaciones que diseñan, desarrollan, despliegan o utilizan sistemas de IA para **ayudarles a gestionar los numerosos riesgos de la IA** y promover un desarrollo y uso fiables y responsables de estos sistemas. Este marco es específico del riesgo e independiente del caso de uso, proporcionando flexibilidad a organizaciones de todos los tamaños y sectores. Según los expertos, este marco de gestión de riesgos del NIST debe complementarse con la Declaración de Derechos de la IA para proteger eficazmente a los ciudadanos.



### AI and Data Act (AIDA) | Gobierno de Canadá | Nov. 2022

G D B

La AIDA propuesta pretende **regular y normalizar el comercio internacional e interprovincial de sistemas de IA** exigiendo a determinadas personas que tomen medidas para reducir el riesgo de daños y resultados sesgados asociados a los sistemas de IA de alto impacto. La AIDA propone: i) garantizar que los sistemas de IA de alto rendimiento cumplan las mismas expectativas en materia de seguridad y derechos humanos a las que están acostumbrados los canadienses; y ii) prohibir los usos imprudentes y malintencionados de la IA que causen graves daños a los canadienses y a sus intereses mediante la creación de nuevas disposiciones penales.



### Ley de Regulación Ética de la Inteligencia Artificial y la Robótica | La Cámara de Diputados | Mayo 2023

G FV B

Su objetivo es regular el uso de la IA y la robótica con fines gubernamentales, económicos, comerciales, administrativos, comunicativos y financieros, de modo que su uso se base siempre en la ética y la legalidad.

# A Anexo 1

## Reacciones de los supervisores sobre la IA: América

Perú y Chile han elaborado normas no vinculantes,  
y Brasil ha publicado un borrador para la regulación de la IA



### Ley que promueve el uso de la IA para el desarrollo económico y social del país | Congreso | Jul. 2023

G FV NB

Su objetivo es promover el uso de la IA en el marco del proceso nacional de transformación digital, potenciando al individuo y el respeto a los derechos humanos para promover el desarrollo económico y social del país en un entorno seguro que garantice su uso ético, sostenible, transparente, replicable y responsable.



### Política nacional de IA | MinCiencia | Oct. 2021

G FV NB

Esta política contiene las directrices estratégicas que el país debe seguir en esta materia durante los próximos 10 años, con el objetivo de empoderar a las personas en el uso y desarrollo de herramientas de IA, y participar en el debate sobre sus consecuencias legales, éticas, sociales y económicas. Esta hoja de ruta se construye en torno a **tres ejes: factores habilitantes, uso y desarrollo** de la IA en Chile, y **aspectos éticos y de seguridad**.



### Proyecto para la regulación de la IA | Senado brasileño | Mayo 2023

G D B

El proyecto de ley pretende **mitigar los riesgos que entraña el sistema teniendo en cuenta la libertad**, la igualdad y el libre desarrollo de la personalidad. La regulación se basa en **tres pilares centrales**: i) garantizar los **derechos de las personas** afectadas por el sistema; ii) clasificar el **nivel de riesgo**; y iii) prever medidas de **gobernanza** para las empresas que proporcionen o exploten el sistema de IA.

# A

## Anexo 1

### Reacciones de los supervisores sobre la IA: Asia y Oceanía

China ha publicado un proyecto de normas vinculantes sobre la IA, India ha elaborado un proyecto de ley sobre protección de datos, y Japón ha publicado directrices con principios sociales para la IA



#### Medidas de IA Generativa | ACC | Jul. 2023

G FV B

Con el fin de promover el **desarrollo saludable y la aplicación normalizada** de la tecnología de IA generativa, la Administración del Ciberespacio de China (ACC), junto con otras seis agencias, colaboró para publicar las Medidas Administrativas Provisionales oficiales para los Servicios de Inteligencia Artificial Generativa. Estas medidas, conocidas como las Medidas de IA Generativa, constituyen la primera normativa integral sobre IA en China y abarcan una amplia gama de temas relacionados con el desarrollo y la prestación de servicios de IA generativa. Esta normativa afectará a las exportaciones tecnológicas chinas y a las redes mundiales de investigación en IA.



#### Proyecto de Ley de Protección de Datos Personales Digitales | Gobierno de la India | Jul. 2023

G D B

El Gabinete ha aprobado el Proyecto de Ley de Protección de Datos Personales Digitales. Sus disposiciones afectan a la IA y cuestionan directamente el tratamiento de datos personales que ésta permite.



#### Directrices de gobernanza para la aplicación de los principios de la IA | METI | Ene. 2022

G FV NB

El documento establece **siete principios sociales** para la IA que deben **aplicarse en toda la sociedad**: i) centrado en el ser humano; ii) educación/alfabetización; iii) protección de la privacidad; iv) garantía de seguridad; v) competencia leal; vi) equidad, responsabilidad y transparencia; e vii) innovación.

# A | Anexo 1

## Reacciones de los supervisores sobre la IA: Asia y Oceanía

### Singapur y Australia han creado un marco voluntario para la IA



#### Marco de gobernanza del modelo de IA | PDPC | Ene. 2020

G FV NB

Se centra principalmente en cuatro grandes áreas: i) **estructuras y medidas de gobernanza interna**; ii) **participación humana** en la toma de decisiones aumentada por IA; iii) **gestión de operaciones**; e iv) **interacción** y comunicación **con las partes interesadas**.



#### Principios éticos de la IA en Australia | Departamento de Industria, Ciencia y Recursos | Nov. 2019

G FV NB

**Estos principios ayudarán a conseguir resultados más seguros, fiables y justos** para todos los australianos. También ayudarán a reducir el riesgo de impacto negativo en los afectados por las aplicaciones de la IA; y estimularán a empresas y gobiernos a practicar normas éticas más estrictas a la hora de diseñar, desarrollar e implantar la IA. Los principios son **voluntarios** y pretenden ser **una aspiración y un complemento** a la normativa vigente sobre IA. Se clasifican en: i) bienestar humano, social y medioambiental; ii) valores centrados en el ser humano; iii) equidad; iv) protección de la privacidad y protección; v) fiabilidad y seguridad; vi) transparencia y explicabilidad; vii) contestabilidad; y viii) responsabilidad.



# A Anexo 2

## AI Bill of Rights en EE.UU.

El AI Bill of Rights es un conjunto de cinco principios y prácticas asociadas para ayudar a guiar el diseño, uso y despliegue de sistemas automatizados para proteger los derechos del público de EE.UU.

### Contexto

En octubre de 2022, la Oficina de Política Científica y Tecnológica de la Casa Blanca publicó el Plan para una Declaración de Derechos de la IA, un ejercicio para imaginar un futuro en el que el público estadounidense esté protegido de los daños potenciales y pueda disfrutar plenamente de los beneficios de los sistemas automatizados. Describe principios que pueden ayudar a garantizar estas protecciones. Algunas de estas protecciones ya se exigen en la Constitución de EE.UU. o se aplican en virtud de la legislación estadounidense vigente.

### Principios

1

Sistemas seguros y eficaces



Los sistemas automatizados deben desarrollarse **consultando a diversas comunidades, partes interesadas y expertos en la materia** para identificar preocupaciones, riesgos y posibles repercusiones del sistema.

2

Protecciones contra la discriminación algorítmica



Los diseñadores, desarrolladores e implantadores de sistemas automatizados deben tomar **medidas proactivas y continuas** para **proteger a las personas y comunidades de la discriminación algorítmica** y para utilizar y diseñar los sistemas de forma equitativa.

3

Protección de datos



Los diseñadores, desarrolladores e implantadores de sistemas automatizados deben pedir permiso y **respetar las decisiones de las personas sobre la** recogida, uso, acceso, transferencia y supresión de **sus datos de forma** adecuada.

4

Aviso y explicación



Los diseñadores, desarrolladores e implantadores deben proporcionar una descripción clara de: i) el funcionamiento general del sistema y el papel que desempeña la automatización; ii) el aviso de que tales sistemas están en uso; iii) la persona u organización responsable del sistema.

5

Alternativas humanas, consideración y repliegue



Optar por sistemas automatizados en favor de una alternativa humana, **cuando proceda**. La idoneidad debe determinarse en función de las expectativas razonables en un **contexto dado** y centrándose en garantizar **una amplia accesibilidad** y proteger al público de impactos especialmente perjudiciales.

Los sistemas automatizados deben desarrollarse consultando a diversas comunidades, partes interesadas y expertos en la materia para determinar las preocupaciones, los riesgos y las posibles repercusiones del sistema.

1

### Proteger al público de los daños de forma proactiva y continua

Consulta

El público debe ser consultado en las fases de **diseño, implantación, despliegue**, adquisición y mantenimiento del desarrollo de **sistemas automatizados**.

Pruebas

Someterse a **pruebas exhaustivas antes de la implantación**. Estas pruebas deben seguir las **mejores prácticas específicas del sector**.

Identificación y mitigación de riesgos

Antes del despliegue, y de forma proactiva y continua, **deben identificarse y mitigarse los riesgos potenciales**.

Seguimiento continuo

Procedimientos de supervisión continua para **garantizar que el rendimiento no caiga por debajo de un nivel aceptable con el paso del tiempo**, en función de los cambios en las condiciones externas o de despliegue, o modificaciones posteriores.

Supervisión clara de la organización

Incluir **procedimientos de gobernanza claramente establecidos** antes de desplegar el sistema, así como la **responsabilidad** de personas o entidades específicas para supervisar la evaluación y mitigación continuas.

2

### Evitar el uso inadecuado, de baja calidad o irrelevante de los datos y el daño compuesto de su reutilización.

Datos pertinentes y de alta calidad

**Los datos utilizados en la** creación, evaluación o despliegue de **cualquier sistema automatizado** deben ser pertinentes, de alta calidad y adaptados a la tarea en cuestión.

Seguimiento y revisión cuidadosos de las fuentes de datos derivadas

**Los datos derivados de otros datos mediante el** uso de algoritmos, como los datos derivados o inferidos de resultados de modelos anteriores, deben identificarse y rastrearse.

Límites a la reutilización de datos en ámbitos sensibles

**La reutilización de datos**, especialmente en un nuevo contexto, puede dar lugar a la **propagación y ampliación de daños**. En consecuencia, estos datos deben someterse a una supervisión adicional para garantizar su seguridad y eficacia.

3

### Demostrar la seguridad y eficacia del sistema

Evaluación independiente

Los evaluadores independientes deben tener **acceso al sistema y a muestras** de los datos asociados, de forma coherente con la privacidad, la seguridad, la ley o la normativa, para realizar dichas evaluaciones.

Informes

Proporcionar **informes actualizados periódicamente** que incluyan: i) una visión general del sistema; ii) los objetivos del sistema; iii) los procedimientos de gestión humana.

### Los algoritmos no deben ser discriminatorios y los sistemas deben utilizarse y diseñarse de forma equitativa

1

*Proteger al público de la discriminación algorítmica de forma proactiva y continua*

Evaluación proactiva de la equidad en el diseño

**Revisar los datos de entrada**, el contexto histórico asociado, la accesibilidad para las personas con discapacidad y los objetivos sociales para identificar posibles discriminaciones y efectos sobre la equidad derivados de la tecnología.

Datos representativos y sólidos

Todos los datos utilizados deben ser **representativos de las comunidades locales**, revisados para detectar sesgos basados en el contexto histórico y social de los datos, y suficientemente sólidos para identificar y mitigar sesgos y daños potenciales.

Protección contra proxies

**Identifique los proxies** comprobando la correlación entre la información demográfica y los atributos de los datos utilizados.

Garantizar la accesibilidad durante el diseño, el desarrollo e implantación

Consideración de **diversas discapacidades**, cumplimiento de las normas de accesibilidad pertinentes e investigación de la experiencia del usuario para abordar cualquier barrera de accesibilidad que impida el uso del sistema automatizado.

Evaluación de las disparidades

Probar los sistemas utilizando **medidas de rendimiento demográfico**, evaluación de la paridad global y de subgrupos, y medidas de calibración para evaluar si los componentes del sistema producen disparidades.

Mitigación de las disparidades

Evaluar varios modelos y seleccionar el que tenga el **menor impacto adverso**, modificar las opciones de introducción de datos o identificar un sistema con menos disparidades. Si no es posible, habrá que reconsiderar el uso del Sistema.

Supervisión y mitigación continuas

**Supervisar periódicamente los sistemas automatizados** para evaluar la discriminación algorítmica que podría surgir de interacciones imprevistas del sistema con desigualdades no contabilizadas.

2

*Demostrar que el sistema protege contra la discriminación algorítmica*

Evaluación independiente

**Permitir la evaluación independiente** de la posible discriminación algorítmica causada por los sistemas automatizados que utilizan o supervisan.

Informes

Informar de una evaluación del impacto algorítmico adecuadamente diseñada, con una **especificación clara** de quién realiza la evaluación, quién evalúa el sistema y cómo se toman **las medidas correctivas** en respuesta a la evaluación.

## Los usuarios deben estar protegidos de las prácticas abusivas mediante protecciones integradas y poder decidir cómo se utilizan los datos sobre el usuario

1

### Proteger la privacidad por diseño y por defecto

Privacidad por diseño y por defecto

Los sistemas automatizados deben **diseñarse** y construirse con la privacidad protegida por defecto.

Recogida de datos y límites del ámbito de uso

**La recopilación de datos** debe tener **un alcance limitado**, con objetivos específicos y estrechamente identificados.

Identificación y mitigación de riesgos.

**Intentar identificar** proactivamente los **daños** y tratar de gestionarlos cuando se recojan, utilicen o almacenen datos sensibles.

Seguridad para preservar la intimidad

Las entidades que crean, utilizan o gobiernan sistemas automatizados deben **seguir** las mejores prácticas de **privacidad** y seguridad diseñadas para garantizar que los datos y metadatos no se filtren más allá del caso de uso específico consentido.

2

### Proteger al público de la vigilancia incontrolada

Mayor supervisión de la vigilancia

Los sistemas de vigilancia o control deben estar sujetos a una **supervisión reforzada** que incluya, como mínimo, la evaluación de los daños potenciales durante el diseño.

Vigilancia limitada y proporcionada

**La vigilancia debe evitarse** a menos que sea necesaria para lograr un fin legítimo y sea proporcional a la necesidad.

Límites a la vigilancia para proteger los derechos y los valores democráticos

**Las libertades civiles** y los **derechos civiles** no deben verse **limitados** por la **amenaza de vigilancia** o acoso facilitada o ayudada por un sistema automatizado.

## Los usuarios deben estar protegidos de las prácticas abusivas mediante protecciones integradas y poder decidir cómo se utilizan los datos sobre el usuario

### 3 Proporcionar al público mecanismos de consentimiento, acceso y control adecuados y significativos sobre sus datos

Consentimiento de uso específico.

Las prácticas de consentimiento **no** deben permitir prácticas de **vigilancia abusivas**.

Solicitudes de consentimiento breves y directas.

Deben utilizarse solicitudes de consentimiento breves y en lenguaje sencillo para que los usuarios entiendan para qué **contextos de uso, período de tiempo** y entidades están dando su consentimiento a los datos y metadatos.

Acceso a los datos y corrección.

Las personas cuyos datos son recogidos, utilizados, compartidos o almacenados por sistemas automatizados deben poder **acceder a los datos y metadatos** sobre sí mismas.

Retirada del consentimiento y supresión de datos.

Las entidades deben **permitir la retirada** del consentimiento de acceso a los datos.

Apoyo al sistema automatizado.

Las entidades que diseñan, desarrollan e implantan sistemas automatizados deben **establecer y mantener las capacidades** que permitan a los particulares utilizar sus propios sistemas automatizados.

### 4 Demostrar que la privacidad de los datos y el control de los usuarios están protegidos

Evaluación independiente.

Las entidades deben permitir la **evaluación independiente** de las afirmaciones realizadas sobre las políticas de datos.

Informes

Cuando los ciudadanos deseen saber qué datos sobre ellos se están utilizando en un sistema, la entidad responsable del desarrollo del sistema debe **responder rápidamente** con un informe sobre los datos que ha recopilado o almacenado.

### Los usuarios deben ser notificados del uso y entender cómo y por qué el sistema automatizado contribuye a los resultados que les afectan

#### 1 *Proporcionar notificaciones de uso y explicaciones claras, oportunas, comprensibles y accesibles.*

Documentación en lenguaje llano generalmente accesible

La entidad responsable de la utilización del sistema automatizado debe asegurarse de que **la documentación** describe el sistema en su conjunto.

Responsable

Los avisos deben identificar claramente la **entidad responsable** del diseño de cada componente del sistema y la entidad que lo utiliza.

Puntualidad y actualidad

Los usuarios deben **recibir una notificación sobre el uso de sistemas automatizados** antes de utilizar la tecnología o mientras se ven afectados por ella.

Breve y claro

Las notificaciones y las **explicaciones** deben **evaluarse**, por ejemplo mediante la investigación de las experiencias de los usuarios, para garantizar que puedan encontrarlos fácilmente, leerlos rápidamente y comprenderlos y actuar en consecuencia.

#### 2 *Explicar cómo y por qué se ha tomado una decisión o se ha realizado una acción mediante un sistema automatizado.*

A la medida del objetivo

Las explicaciones deben **adaptarse a la finalidad específica para la que se** espera que el usuario las utilice, y deben indicar claramente dicha finalidad.

Adaptado al objetivo de la explicación

Las explicaciones deben dirigirse a **un público específico** e indicarlo claramente. Una explicación dirigida al sujeto de una decisión puede diferir de otra dirigida a un defensor, a un experto en la materia o a un responsable de la toma de decisiones.

Adaptado al nivel de riesgo

Debe realizarse una **evaluación** para determinar el nivel de riesgo del sistema automatizado.

Válido

La explicación proporcionada por un sistema debe **reflejar con exactitud los factores** y las influencias que llevaron a una **decisión concreta**, y debe tener sentido para la personalización particular basada en el propósito, el objetivo y el riesgo.

#### 3 *Demostrar la protección de la notificación y la explicación*

Informes

**Documente** las decisiones tomadas basándose en las consideraciones anteriores.

Los usuarios deben poder excluirse voluntariamente, cuando proceda, y tener acceso a una persona que pueda estudiar y solucionar rápidamente los problemas que se le planteen

**1** *Proporcionar un mecanismo para renunciar a los sistemas automatizados en favor de una alternativa humana.*

Aviso e instrucciones breves, claros y accesibles

Las personas afectadas por un sistema automatizado **deben recibir un aviso breve** y claro de que tienen derecho a excluirse, junto con instrucciones claras sobre cómo hacerlo.

Alternativas humanas cuando proceda

Cuando los sistemas automatizados formen parte del proceso de consecución, deben **preverse procesos oportunos alternativos dirigidos por personas**.

Alternativa humana oportuna y no gravosa

**La exclusión voluntaria** debe ser **oportuna** y no suponer una carga excesiva.

**2** *Si falla un sistema automatizado, proporcionar una atención humana oportuna y poner remedio mediante un sistema de emergencia y escalada.*

Proporcional

La disponibilidad de la **consideración humana** y de la reserva debe ser proporcional al potencial del sistema automatizado.

Accesible

Los mecanismos para la consideración humana y la retirada deben ser **fáciles de encontrar**.

Cómodo

Los mecanismos de consideración humana y de retroceso **no deben ser excesivamente onerosos en** comparación con el equivalente del sistema automatizado.

Equitativo

Se debe considerar la posibilidad de **garantizar que los resultados del** sistema de **retroceso** y escalada sean equitativos.

A tiempo

Las consideraciones humanas y los contratiempos sólo son útiles si se llevan a cabo y concluyen a **tiempo**.

Eficaz

La estructura organizativa que rodea a los procesos debe diseñarse de modo que si el **responsable humano de la toma de decisiones** determina que debe anularse una, la nueva decisión **se promulgue de forma efectiva**.

Mantenido

La consideración humana y el proceso de emergencia, así como los procesos automatizados asociados, deben **mantenerse y apoyarse mientras se** siga utilizando el sistema automatizado pertinente.

# A Anexo 3

## Supervisión e iniciativas - Posición del BCE (1/2)



En el sector bancario, el BCE, como supervisor prudencial de las entidades de crédito, ha publicado su posición sobre los requisitos del AI Act

### Observaciones generales

- El BCE celebra el objetivo del reglamento propuesto y **reconoce la importancia de establecer requisitos armonizados para los sistemas de IA**, especialmente en el sector bancario.
- El reglamento propuesto **integra las obligaciones y procedimientos establecidos en la Directiva 2013/36/UE (DRC)** en materia de gestión de riesgos y gobernanza, sin embargo, **se solicitan más aclaraciones para precisar las expectativas supervisoras en materia de gobernanza interna.**
- El BCE considera que la obligación de que el proveedor disponga de **un sistema de gestión de la calidad y supervise los sistemas de IA ya se cumple mediante el cumplimiento de la DRC.**
- La regulación propuesta debe entenderse sin perjuicio de las obligaciones prudenciales más específicas o estrictas de las entidades de crédito establecidas en la regulación sectorial y complementadas por orientaciones de supervisión (por ejemplo, el control efectivo de la externalización, tal como se especifica en la DRC).
- **El BCE sigue un enfoque tecnológicamente neutro.**
- **Debería aclararse la función del BCE conforme al reglamento propuesto:** (1) las competencias de supervisión prudencial del BCE en general, y en relación con la vigilancia del mercado y la evaluación de la conformidad; y 2) la aplicación del reglamento propuesto al desempeño de las funciones que el Tratado atribuye al BCE.

### Clasificación de los sistemas de IA

- Con arreglo al reglamento propuesto, las actividades de calificación crediticia que utilicen sistemas de inteligencia artificial estarán sujetas a los requisitos mínimos aplicables a los sistemas de inteligencia artificial de alto riesgo. El **BCE sugiere que métodos como los árboles de decisión o las regresiones logísticas no se consideren de alto riesgo** siempre que su impacto en la evaluación de la solvencia o la puntuación crediticia de las personas físicas sea menor.
- Para la calificación crediticia, el **BCE sugiere retrasar la entrada en vigor hasta que haya especificaciones sobre las condiciones para verificar la conformidad con los requisitos aplicables**, y definir cuándo debe considerarse que los sistemas de IA "son puestos en servicio por proveedores a pequeña escala para su propio uso".
- El BCE sugiere que se considere la posibilidad de **actualizar la lista de sistemas de IA de alto riesgo para tener en cuenta otros sistemas de IA implantados por las entidades financieras, como la modelización de datos de IA que vincula datos de ventas, transacciones y rendimiento para garantizar una visión clara y general del riesgo de conducta en un área determinada.** Del mismo modo, los sistemas de IA podrían utilizarse en el seguimiento en tiempo real de los pagos, o en la elaboración de perfiles de clientes o transacciones, con fines de lucha contra el blanqueo de capitales y la financiación del terrorismo...



# A Anexo 3

## Supervisión e iniciativas - Posición del BCE (2/2)



En el sector bancario, el BCE, como supervisor prudencial de las entidades de crédito, ha publicado su posición sobre los requisitos del AI Act

Aclaración sobre el papel del BCE en virtud del reglamento propuesto

- **Aclaración de las competencias de supervisión prudencial del BCE en relación con la vigilancia del mercado**
  - El BCE entiende que, con arreglo al reglamento propuesto, el BCE no es en modo alguno una autoridad de vigilancia del mercado.
  - El BCE considera que la vigilancia del mercado no tiene por objeto garantizar la seguridad y la solidez de las entidades de crédito, sino que se centra en la protección de los intereses de los particulares y estima que el texto debería modificarse para aclarar este extremo.
  - Los Estados miembros podrían considerar la designación de las autoridades nacionales competentes que participan en la supervisión de las entidades de crédito como responsables de la vigilancia del mercado en el contexto del reglamento propuesto, en la medida en que lo permita su mandato.
  - El BCE observa que las disposiciones de vigilancia del mercado del reglamento propuesto no abordan adecuadamente las situaciones en las que un sistema de IA se pone en servicio para uso propio.
- **Aclaración de las competencias de supervisión prudencial del BCE en materia de evaluación de la conformidad**
  - Se invita al legislador de la Unión a considerar en qué medida varios elementos de la evaluación de la conformidad podrían no ser de naturaleza prudencial en la medida en que se refieren en gran medida a la evaluación técnica de los sistemas de IA para salvaguardar la salud y la seguridad de las personas y garantizar que los derechos fundamentales.
  - Destaca la necesidad de designar a las autoridades competentes como responsables de la supervisión de la evaluación de la conformidad de los requisitos en materia de salud, seguridad y derechos fundamentales.
  - Ciertos requisitos para los sistemas de IA de alto riesgo no son del todo claros o lo suficientemente específicos como para proporcionar una comprensión suficiente para informar las expectativas de supervisión (por ejemplo, que los datos de entrenamiento, validación y prueba sean pertinentes y representativos).
  - El BCE considera que el reglamento propuesto debería modificarse para reflejar la naturaleza ex post de la evaluación específica como parte del PRES.
- **Aclaración de las competencias de supervisión prudencial del BCE:** El BCE sólo podrá ser considerado autoridad competente en la medida en que sea necesario para el desempeño de las funciones que le atribuye el Reglamento del MUS.
- **Aclaración de la independencia del BCE en el desempeño de las funciones que le atribuye el Tratado:** El BCE entiende que cuando actúe como proveedor que comercializa o pone en servicio sistemas de inteligencia artificial, o como usuario, él (o los BCN) pueden estar sujetos al reglamento propuesto, manteniendo al mismo tiempo su independencia para desempeñar las funciones que le atribuye el Tratado.



# A | Anexo 3

## Supervisión e iniciativas - AI Sandbox en España



El Gobierno español ha puesto en marcha una iniciativa, en colaboración con la Comisión Europea, para implantar un sandbox regulatorio de Inteligencia Artificial en la UE

### Objetivo



- El objetivo de esta colaboración es **poner en contacto a las autoridades competentes con las empresas desarrolladoras de Inteligencia Artificial para definir conjuntamente las mejores prácticas a la hora de aplicar el futuro reglamento europeo** sobre Inteligencia Artificial, promovido por la Comisión Europea.
- El resultado de estas pruebas se **recopilará en una guía de buenas prácticas**, que se presentará durante la Presidencia española del Consejo de la UE en el segundo semestre de 2023.
- La guía será **accesible a todos los Estados miembros y a la Comisión Europea**.

### Alcance



Dado que la propuesta de **Reglamento sobre Inteligencia Artificial se centra en particular en las obligaciones que deben cumplir los denominados sistemas de inteligencia artificial de alto riesgo**, la participación en el entorno controlado regulado se centra en ellos:

- Sistemas de inteligencia artificial clasificados como de alto riesgo;
- Sistemas de inteligencia artificial de propósito general;
- Modelos fundacionales;
- Sistemas de inteligencia artificial generativa.

### Beneficios



- Aportar **claridad sobre los nuevos requisitos para los sistemas de IA establecidos** en el Reglamento sobre IA;
- Transmitir **a las entidades** que desarrollan soluciones de IA **conocimientos especializados sobre la aplicación de la próxima legislación sobre IA**;
- **Fomentar la innovación** y permitir el desarrollo de sistemas de IA innovadores y fiables;
- **Crear capacidad** e iniciar consultas en España que conduzcan finalmente a la creación de una Autoridad Nacional de Supervisión;
- **Poner a prueba futuras obligaciones y requisitos en un entorno controlado** y proporcionar experiencia práctica de aprendizaje para apoyar el desarrollo de normas, orientaciones y herramientas a escala nacional y europea.

# A Anexo 3

## Supervisión e iniciativas - Autoridad de Supervisión de la IA en España



En España, el Real Decreto 729/2023 por el que se aprueba la constitución de la Agencia Española de Supervisión de la IA, como agencia estatal adscrita al Ministerio de Economía y Transformación Digital

### Agencia Española para la Supervisión de la Inteligencia Artificial (AESIA)

#### Constitución de AESIA

- La Agencia tendrá **personalidad jurídica propia, autonomía de gestión y potestades administrativas** para el cumplimiento de sus fines. Su sede estará en A Coruña.
- La **AESIA supervisará el uso de sistemas de inteligencia artificial para proteger los derechos fundamentales y minimizar los riesgos**. Colaborará con las autoridades nacionales y europeas.
- Sus competencias incluyen: promover entornos de prueba, fomentar el uso ético y sostenible de la IA, reforzar la confianza en la tecnología, coordinarse con otros agentes, formar y concienciar sobre el uso responsable de la IA.
- La **AESIA contará con órganos de gobierno** (Presidencia y Consejo de Administración), órganos ejecutivos (Dirección, Secretaría General, dos Subdirecciones) y órganos de control (Comisión de Control).
- Se regula el régimen de personal, económico-financiero, presupuestario, patrimonial y de contratación de AESIA.
- La asistencia jurídica correrá a cargo de la Fiscalía General del Estado.

#### Principales competencias

**Las principales competencias de la Agencia Española para la Supervisión de la Inteligencia Artificial (AESIA), detalladas en el Real Decreto, son:**

- Promover entornos de prueba regulados para que los sistemas de IA puedan probarse de forma segura y conforme a la ley.
- Promover un uso ético, sostenible y respetuoso con el medio ambiente de la IA.
- Crear un sistema voluntario de certificación para garantizar las normas técnicas y el diseño responsable de las soluciones de IA.
- Identificar tendencias y evaluar el impacto social de la IA mediante estudios e informes.
- Coordinarse con otras iniciativas públicas y privadas relacionadas con la IA.
- Generar conocimiento, formación y concienciación sobre la IA ética y humanista.
- Dinamizar el mercado para fomentar prácticas innovadoras de IA.
- Colaborar con el sector privado para fomentar un desarrollo humanista de la IA.
- Supervisar y, en su caso, sancionar el uso de sistemas de IA para garantizar el cumplimiento de la normativa europea y nacional.
- Proporcionar asistencia técnica a jueces y tribunales en casos legales relacionados con la IA.
- Otras funciones relacionadas con la supervisión de la IA que puedan atribuírsele debido a cambios normativos o tecnológicos.

# A | Anexo 4

## Abreviaturas

Abreviatura	Significado
AI / IA	Inteligencia artificial
EBA	Autoridad Bancaria Europea
CRR	Regulación de los requisitos de capital
CAC	Administración del Ciberespacio de China
CE	Comisión Europea
UE	Unión Europea
GDPR	Reglamento General de Protección de Datos
Gato Nat	Organización Internacional de Comisiones de Valores
OICV	Autoridad nacional competente
ML	Aprendizaje automático
MINECO	Ministerio de Economía y Transformación Digital
METI	Ministerio de Economía, Comercio e Industria
OCDE	Organización para la Cooperación y el Desarrollo Económico
IRB	Basado en calificaciones internas
PDPC	Comisión de Protección de Datos Personales

**MS AI**  
**Management Solutions**  
*Making things happen*



Una firma internacional



Equipo multidisciplinar



Know-how especialista



Experiencia demostrada



Máximo Compromiso

**Javier Calvo Martín**

Socio de Management Solutions

Javier.calvo.martin@managementsolutions.com

**Manuel Ángel Guzmán**

Socio de Management Solutions

Manuel.guzman@managementsolutions.com

**Marta Hierro**

Socia de Management Solutions

Marta.Hierro@msspain.com