

Bonn, Bucarest, Dublín, Lisboa, Madrid, Milán, París, La Haya, Viena, Varsovia

Cómo definir un sistema de IA en virtud del Reglamento europeo de IA

Grupo de trabajo sobre IA y datos de CEDPO

Serie Micro-Insights

Octubre de 2025

Autores:

Marc Bellon

Lionel Capel

Maria Moloney

Laura Morató Pascual

Alessandro Vasta

Información de contacto:

<https://cedpo.eu>

info@cedpo.eu

Acerca de la serie Micro-Insights

La serie Micro-Insights es una iniciativa editorial del Grupo de Trabajo sobre IA y Datos de CEDPO. Ofrecerá artículos breves, digeribles y definitivos sobre áreas clave de interés en la intersección entre los datos y la inteligencia artificial. Con un enfoque práctico, centrado por un lado en la explicación de temas complejos y por otro en su aplicación, se destacará la importancia de las áreas clave y se asesorará a los profesionales sobre el impacto y los próximos pasos. Con la entrada en vigor del Reglamento europeo de IA ("RIA") en 2024, el escenario está ahora preparado para todos los profesionales, y es posible debatir la regulación de los datos y la IA con mucha más claridad.

La serie Micro-Insights seguirá la evolución de la IA y los datos en los próximos años y, a medida que se acerque el final del periodo crucial de aplicación del RIA y que las tecnologías de IA evolucionen de forma cada vez más novedosa e inesperada, la serie responderá con orientaciones actualizadas y autorizadas sobre los principales ámbitos de interés.

Entre otros temas, analizaremos:

- Regulación de los Sistemas de Inteligencia Artificial de Propósito General según el Reglamento europeo de IA
- Comentarios al Pacto de IA.
- Formación para profesionales de protección de datos sobre la realización de evaluaciones de impacto sobre derechos fundamentales según RIA.
- El papel que tendrán los reguladores de protección de datos en la regulación de la IA.
- Examinar si el Delegado de Protección de Datos es o no es la persona adecuada para ser el "IA Officer"
- La base legal para el uso de datos de entrenamiento en el aprendizaje automático.
- El conjunto de herramientas de preparación para el RIA.

Índice

1. Introducción.....	4
2. Criterios de los sistemas de IA	5
Un sistema basado en máquinas.....	5
Niveles variables de autonomía	5
Adaptabilidad tras la implementación.....	6
Objetivos.....	6
Inferencia	6
Resultados	6
Influencia en los entornos	7
3. Conclusión	7

1. Introducción

Uno de los principales objetivos del Reglamento europeo de IA (“RIA”) es fomentar la innovación en materia de IA, garantizando al mismo tiempo un alto nivel de protección de la salud, la seguridad y los derechos fundamentales dentro de la UE. Es fundamental comprender el alcance de la normativa del RIA, ya que nos permite reconocer exactamente qué entra dentro del ámbito de aplicación de la normativa y qué no. Solo los sistemas que cumplen la definición de sistema de IA de la normativa entran dentro de su ámbito de aplicación. Dada la complejidad de la normativa, la Comisión Europea debe proporcionar directrices para ayudar a las partes interesadas a comprender y aplicar la definición de sistema de IA que se recoge en el Reglamento.

Por este motivo, el 6 de febrero de 2025, la Comisión Europea publicó unas directrices para ayudar a todas las partes interesadas a comprender mejor la definición de sistemas de IA que figura en el RIA y para apoyar la aplicación inicial de las normas de la ley, que entraron en vigor ese mismo mes.

Las directrices son una ayuda, pero no son jurídicamente vinculantes. El Tribunal de Justicia de la Unión Europea (TJUE) es la autoridad competente para la interpretación definitiva del RIA. Dada la rápida evolución de la tecnología de IA, las directrices hacen hincapié en la necesidad de una aplicación flexible y no mecánica de la definición de sistema de IA.

El artículo 3, apartado 1, del Reglamento se define un sistema de IA como un sistema **basado en máquinas** con **distintos niveles de autonomía**, que puede mostrar **capacidad de adaptación** tras su implementación y que **deduce resultados** (predicciones, contenidos, recomendaciones, decisiones) a partir de **entradas para influir en entornos físicos o virtuales**.

Si desglosamos esta definición, podemos ver que hay siete criterios principales para los sistemas de IA:

1. Debe ser un sistema basado en máquinas.
2. Existen distintos niveles de autonomía.
3. Puede mostrar capacidad de adaptación tras su implementación.
4. Debe haber (al menos uno o más) objetivos.
5. Debe existir inferencia.
6. Debe haber (uno o más) resultados.
7. Debe tener influencia en su entorno.

El resto de este artículo ofrece explicaciones muy breves de cada uno de estos siete criterios para ayudar al lector a determinar de forma práctica la naturaleza de un sistema de IA tal y como se define en el RIA.

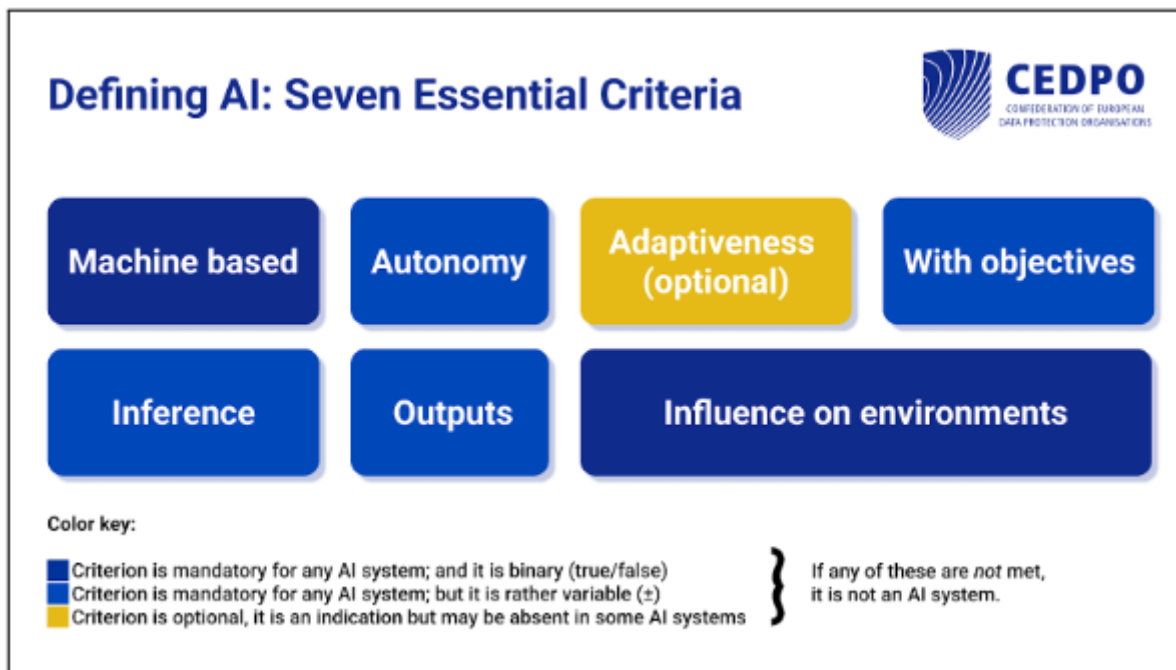


Figura 1. Los siete criterios para los sistemas de IA

2. Criterios de los sistemas de IA

Un sistema basado en máquinas

Todos los sistemas de IA se desarrollan y funcionan con máquinas: hardware (elementos físicos) y software (código informático, instrucciones, programas, sistemas operativos y aplicaciones). Los sistemas de IA pueden depender de muchos de estos componentes.

El término «basado en máquinas» se entiende en sentido amplio e incluye los sistemas emergentes que se alejan de la informática tradicional, entre los que se incluyen:

- sistemas de computación cuántica; y
- sistemas computacionales biológicos u orgánicos (computación biológica)

Niveles variables de autonomía

Para determinar si un sistema puede considerarse un sistema de IA, el nivel de autonomía es un factor crucial. Los sistemas diseñados para funcionar con un grado razonable de independencia en sus acciones cumplen el criterio de autonomía de la definición de sistema de IA.

Un aspecto clave de la autonomía es el nivel o grado de participación e intervención humana. Por ejemplo, un sistema como ChatGPT, que requiere entradas proporcionadas manualmente para generar una salida de forma independiente, demuestra «cierto grado de independencia de

acción». Esto se debe a que el sistema es capaz de producir una salida sin que sea especificada con precisión por un humano.

Adaptabilidad tras la implementación

Un sistema de IA «puede mostrar adaptabilidad tras su implementación». Esto se refiere a las capacidades de autoaprendizaje, lo que significa que el sistema puede aprender automáticamente, descubrir nuevos patrones o identificar relaciones en los datos más allá de lo que se le enseñó inicialmente. Esto permite al sistema cambiar su comportamiento mientras se utiliza, lo que puede producir resultados diferentes para las mismas entradas.

A diferencia de otros criterios, todos ellos esenciales para definir un sistema de IA, la adaptabilidad tras la implementación no es un criterio decisivo, sino más bien un indicador. Por lo tanto, un sistema que ha sido entrenado inicialmente y que no muestra adaptabilidad una vez implementado no queda excluido de la definición de sistema de IA.

Objetivos

Los «objetivos» de un sistema de IA se refieren a sus metas internas, que determinan cómo funciona el sistema. Estos objetivos pueden definirse explícitamente mediante funciones de optimización o derivarse implícitamente de patrones aprendidos de interacción con su entorno. Estos objetivos pueden diferir del «propósito previsto» asignado por el proveedor en un contexto específico, ya que se centran en los resultados que debe alcanzar el sistema, sin implicar necesariamente cómo se integra o se utiliza. En consecuencia, la IA puede desarrollar objetivos no formulados, como es el caso de los modelos de aprendizaje por refuerzo o los grandes modelos lingüísticos. Si bien esta característica es esencial para describir un sistema de IA, no es suficiente por sí sola para distinguirlo de otros algoritmos.

Inferencia

La «inferencia» en IA se refiere a la capacidad de un sistema para generar resultados, como predicciones, contenidos, recomendaciones o decisiones, basándose en la información que recibe. Este proceso implica extraer conclusiones a partir de los datos, a menudo utilizando técnicas de IA como el aprendizaje automático o enfoques basados en el conocimiento. En el contexto de la IA, la inferencia abarca tanto la generación de resultados a partir de entradas (durante la fase de uso del sistema) como la derivación de modelos o algoritmos a partir de datos durante la fase de construcción. Esta capacidad de inferir distingue a los sistemas de IA del software tradicional, que se basa en reglas predefinidas y no genera resultados basados en patrones aprendidos o derivados.

Resultados

La capacidad de un sistema para generar resultados, como predicciones, contenidos y recomendaciones que pueden influir en entornos físicos o virtuales, basándose en las entradas que recibe y utilizando el aprendizaje automático y enfoques basados en la lógica y el conocimiento, es fundamental para lo que hacen los sistemas de IA y lo que distingue a esos

sistemas de otras formas de software. Estos resultados se dividen en cuatro categorías principales: **predicciones, contenidos, recomendaciones y decisiones**. Las **predicciones** implican estimar valores desconocidos con una intervención humana mínima. El **contenido** se refiere a la creación de material nuevo, como texto e imágenes. Las **recomendaciones** sugieren acciones o productos basados en los datos del usuario. Las **decisiones** son conclusiones tomadas por el sistema que, en efecto, pueden automatizar el juicio humano. Los sistemas de IA, especialmente los que utilizan el aprendizaje automático, pueden destacar en el manejo de datos complejos y en la generación de resultados más matizados en comparación con los sistemas que no utilizan IA.

Influencia en los entornos

Un elemento importante a la hora de definir un sistema de IA es su capacidad para producir resultados que puedan influir tanto en entornos físicos como virtuales. Esto pone de relieve que los sistemas de IA son agentes activos que influyen en los entornos en los que operan. El término «entornos físicos o virtuales» abarca tanto objetos físicos tangibles (como un brazo robótico) como entornos virtuales, incluidos los espacios digitales, los flujos de datos y los ecosistemas de software. Un agente de IA que reserva billetes de avión es un ejemplo de un sistema de IA que influye en un entorno físico.

3. Conclusión

El Reglamento europeo de IA ofrece una definición exhaustiva de los sistemas de IA. Describe dichos sistemas como aquellos que cumplen siete criterios clave, tal y como se describe en este artículo. La conclusión principal que se extrae del ejercicio de definir los sistemas de IA según el RIA es que estos sistemas son esencialmente códigos de software desarrollados utilizando técnicas como el aprendizaje automático, métodos estadísticos y/u otros enfoques basados en la lógica, que dan lugar a decisiones, predicciones o recomendaciones que tienen el potencial de influir en nuestros entornos físicos y virtuales.

Esta compleja definición pretende captar la naturaleza dinámica y en constante evolución de la IA, al tiempo que trata de establecer bases jurídicas claras para supervisarla. Al integrar los sistemas de IA en este marco regulador, la UE busca establecer un equilibrio entre el fomento de la innovación y la protección de la salud, la seguridad y los derechos fundamentales. Esto, a su vez, promoverá la transparencia, la rendición de cuentas y la confianza en el desarrollo y el despliegue de las tecnologías de IA en toda la Unión y en el futuro.