



# Recomendación sobre la inteligencia artificial

Traducción no oficial

**Instrumentos  
jurídicos de la OCDE**



El presente trabajo se publica bajo la responsabilidad del Secretario General de la OCDE. Las opiniones expresadas y los argumentos utilizados en el mismo no reflejan necesariamente el punto de vista oficial de los países miembros de la OCDE.

Tanto este documento, como cualquier dato y cualquier mapa que se incluya en él, se entenderán sin perjuicio alguno respecto al estatus o la soberanía de cualquier territorio, a la delimitación de fronteras y límites internacionales, ni al nombre de cualquier territorio, ciudad o área.

*Este documento está disponible de manera gratuita. Puede ser reproducido y distribuido gratuitamente sin necesidad de ningún otro permiso, siempre que no se modifique de ninguna manera. No puede venderse.*

*Esta es una traducción no oficial. Si bien se ha puesto el mayor empeño para cerciorarse que correspondan con texto original, las únicas versiones oficiales son los textos en inglés y en francés disponibles en el sitio web de la OCDE, <https://legalinstruments.oecd.org>.*

## **EL CONSEJO,**

**TENIENDO EN CUENTA** el artículo 5(b) de la Convención de la Organización para la Cooperación y el Desarrollo Económicos de 14 de diciembre de 1960;

**TENIENDO EN CUENTA** las normas elaboradas por la OCDE en las áreas de privacidad, seguridad digital, protección del consumidor y conducta empresarial responsable;

**TENIENDO EN CUENTA** los Objetivos de Desarrollo Sostenible recogidos en la Agenda 2030 para el Desarrollo Sostenible, adoptada por la Asamblea General de las Naciones Unidas (A/RES/70/1), así como la Declaración Universal de Derechos Humanos de 1948;

**TENIENDO EN CUENTA** el importante trabajo que está efectuándose en materia de inteligencia artificial (en lo sucesivo, «IA») en otros foros internacionales de carácter gubernamental y no gubernamental;

**RECONOCIENDO** que la IA tiene repercusiones generalizadas, de amplio calado y de ámbito mundial que están transformando sociedades, sectores económicos y el mundo laboral, y probablemente lo hagan cada vez más en el futuro;

**RECONOCIENDO** que la IA tiene potencial para mejorar el bienestar general de las personas, para contribuir a una actividad económica mundial positiva y sostenible, para aumentar la innovación y la productividad y para ayudar a responder ante retos mundiales clave;

**RECONOCIENDO** que, al mismo tiempo, estas transformaciones pueden tener efectos dispares en y entre las sociedades y las economías y sobre todo en relación con los cambios económicos, la competencia, las transiciones en el mercado laboral, las desigualdades y las implicaciones para la democracia y los derechos humanos, la privacidad y la protección de datos y la seguridad digital;

**RECONOCIENDO** que la confianza es un vector clave de la transformación digital; que, si bien la naturaleza de las futuras aplicaciones de la IA y sus implicaciones pueden resultar difíciles de predecir, la fiabilidad de los sistemas de IA es un factor determinante para la difusión y adopción de esta tecnología; y que se necesita un debate público bien informado y de toda la sociedad con vistas a aprovechar el potencial beneficioso de la tecnología, limitando al mismo tiempo los riesgos que lleva aparejados;

**SUBRAYANDO** que determinados marcos jurídicos, reglamentarios y normativos nacionales e internacionales ya hacen alusión a la IA, incluidos los relativos a los derechos humanos, la protección de los consumidores y los datos de carácter personal, los derechos de propiedad intelectual, la conducta empresarial responsable y la competencia, señalando, no obstante, que podría ser necesario evaluar la idoneidad de algunos marcos y desarrollar nuevos enfoques;

**RECONOCIENDO** que, a la vista del rápido desarrollo e implantación de la IA, se necesita un marco de políticas estable que promueva un enfoque de la IA fiable centrado en las personas, aliente la investigación, preserve los incentivos económicos para innovar y se aplique a todas las partes interesadas en función de su papel y el contexto;

**CONSIDERANDO** que, con el fin de promover la adopción de una IA fiable en la sociedad y convertir la fiabilidad de la IA en un parámetro competitivo en el mercado internacional, resulta esencial aprovechar las oportunidades que ofrecen las aplicaciones de IA, abordando los retos que plantean, y capacitar a las partes interesadas para interactuar con ella.

### **En lo concerniente a la propuesta del Comité de Política Digital:**

**I. ACUERDA** que, a los efectos de la presente Recomendación, se utilicen las siguientes definiciones:

- *Sistema de IA*: Un sistema de IA es un sistema basado en máquinas que, para objetivos explícitos o implícitos, infiere, a partir de los datos de entrada que recibe, cómo generar información de salida como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos reales o virtuales. Una vez implementados, los distintos sistemas de IA presentan

diversos niveles de autonomía y varían en su capacidad de adaptación.

- *Ciclo de vida del sistema de IA*: El ciclo de vida de un sistema de IA comprende generalmente varias fases: planificación y diseño; recopilación y tratamiento de datos; creación de modelo(s) y/o adaptación de modelo(s) existente(s) a tareas específicas; prueba, evaluación, verificación y validación; entrada en servicio/despliegue; explotación y supervisión; y retirada/desmantelamiento. Estas fases a menudo se desarrollan de forma iterativa y no son necesariamente secuenciales. La decisión de dejar de explotar un sistema de IA puede producirse en cualquier momento durante la fase de explotación y supervisión.
- *Actores de la IA*: Los actores de IA son aquellos que desempeñan un papel activo en el ciclo de vida del sistema de IA, como las entidades y personas que despliegan y explotan la IA.
- *Conocimientos sobre IA*: Los conocimientos sobre IA se refieren a las competencias y recursos, como los datos, el código informático, los algoritmos, los modelos, la investigación, el *know-how*, los programas de formación, la gobernanza, los procesos y las buenas prácticas que se requieren para comprender y participar en el ciclo de vida de la IA, incluida la gestión de los riesgos.
- *Partes interesadas*: Por partes interesadas se entiende todas las entidades y personas que participan o se ven afectadas directa o indirectamente por los sistemas de IA. Los actores de la IA son un subconjunto de las partes interesadas.

### **Sección 1: Principios de administración responsable en aras de una IA fiable**

**II. RECOMIENDA** que los Miembros y no Miembros que se adhieran a esta Recomendación (en lo sucesivo, los «Adherentes») promuevan y apliquen los siguientes principios de administración responsable en aras de una IA fiable, que son importantes para todas las partes interesadas.

**III. INSTA** a todos los actores de la IA a promover e implantar, de acuerdo con sus respectivas funciones, los siguientes principios de administración responsable en aras de una IA fiable.

**IV. SUBRAYA** que los siguientes principios son complementarios y deberían considerarse como un conjunto.

#### **1.1. Crecimiento inclusivo, desarrollo sostenible y bienestar**

Las partes interesadas deberían poner en marcha activamente la administración responsable en aras de una IA fiable en busca de resultados beneficiosos para las personas y el planeta, como aumentar las capacidades del ser humano y estimular la creatividad, avanzar en la inclusión de poblaciones infrarrepresentadas, reducir las desigualdades económicas, sociales, de género y de otra índole, y proteger los entornos naturales, reforzando de este modo el crecimiento inclusivo, el bienestar, el desarrollo sostenible y la sostenibilidad medioambiental.

#### **1.2. Respetar el Estado de derecho, los derechos humanos y los valores democráticos, incluidas la equidad y la privacidad**

a) Los actores de la IA deben respetar el Estado de derecho, los derechos humanos, los valores democráticos y los centrados en el ser humano a lo largo del ciclo de vida de los sistemas de IA. Por ejemplo, la no discriminación y la igualdad, la libertad, la dignidad, la autonomía de las personas, la privacidad y la protección de datos, la diversidad, la equidad, la justicia social y los derechos laborales reconocidos internacionalmente. Esto también comprende abordar la información engañosa y la desinformación amplificadas por la IA, respetando al mismo tiempo la libertad de expresión y otros derechos y libertades protegidos por el Derecho internacional aplicable.

b) Para ello, los actores de la IA deben poner en marcha mecanismos y garantías, como la capacidad de intervención y supervisión humana, lo que incluye abordar riesgos derivados de usos distintos del fin perseguido, usos indebidos intencionados o usos indebidos no intencionados, de forma apropiada

al contexto y coherente con el estado actual de la tecnología.

### **1.3. Transparencia y explicabilidad**

Los actores de la IA deberían comprometerse con la transparencia y una divulgación responsable en relación con los sistemas de IA. Para ello, deben proporcionar información significativa que sea adecuada para la situación específica y se ajuste al estado actual de la tecnología:

- i. para fomentar una comprensión general de los sistemas de IA, incluidas sus capacidades y limitaciones,
- ii. para poner en conocimiento de las partes interesadas sus interacciones con los sistemas de IA, también en el lugar de trabajo,
- iii. cuando resulte viable y útil, para suministrar información transparente y comprensible sobre las fuentes de datos/entradas, los factores, los procesos y/o el razonamiento que subyace a las predicciones, contenidos, recomendaciones y decisiones, para permitir a los afectados por un sistema de IA comprender los resultados, y
- iv. para suministrar información que permita a los afectados negativamente por un sistema de IA cuestionar sus resultados.

### **1.4. Solidez, seguridad y protección**

- a) A lo largo de su vida útil, los sistemas de IA deben ser robustos, seguros y estar protegidos de tal modo que, en condiciones de uso normal, uso previsible o uso indebido, u otras condiciones adversas, funcionen adecuadamente y no planteen riesgos excesivos en materia de seguridad y/o protección.
- b) Deben existir mecanismos para, cuando proceda, garantizar que, si los sistemas de IA amenazan con provocar daños indebidos o muestran un comportamiento indeseado, estos puedan ser invalidados, corregidos y/o desmantelados de forma segura, según sea necesario.
- c) Cuando sea técnicamente viable, deben existir mecanismos para reforzar la integridad de la información y, al mismo tiempo, garantizar el respeto de la libertad de expresión.

### **1.5. Responsabilidad**

- a) Los actores de la IA deben ser responsables del correcto funcionamiento de los sistemas de IA y del respeto de los principios anteriores, de acuerdo con sus funciones, el contexto y el estado de la tecnología.
- b) Para ello, los actores de la IA deben garantizar la trazabilidad, entre otros elementos, de los conjuntos de datos, los procesos y las decisiones tomadas durante el ciclo de vida del sistema de IA con el fin de permitir el análisis de los resultados del sistema de IA y garantizar que las respuestas sean adecuadas al contexto específico y coherentes con el estado de la tecnología.
- c) Los actores de la IA, de acuerdo con sus funciones, el contexto y su capacidad de actuación, deben aplicar de forma permanente un enfoque sistemático de la gestión de riesgos a cada fase del ciclo de vida del sistema de IA y adoptar una conducta empresarial responsable a la hora de abordar los riesgos relacionados con los sistemas de IA, lo que comprende, cuando proceda, la cooperación entre diferentes actores de la IA, proveedores de conocimientos y recursos de IA, usuarios de sistemas de IA y otras partes interesadas. Los riesgos comprenden los relacionados con un sesgo dañino, con los derechos humanos, como la seguridad, la protección y la privacidad, así como con los derechos laborales y de propiedad intelectual.

## **Sección 2: Políticas nacionales y cooperación internacional en aras de una IA fiable**

**V. RECOMIENDA** que los Adherentes pongan en marcha las siguientes recomendaciones, que son coherentes con los principios de la sección 1, en sus políticas nacionales y su cooperación internacional,

prestando especial atención a las pequeñas y medianas empresas (pymes).

### **2.1. Invertir en investigación y desarrollo de la IA**

- a) Los gobiernos deben considerar la inversión pública a largo plazo y alentar también la inversión privada en investigación y desarrollo y ciencia abierta, incorporando los esfuerzos interdisciplinarios, para estimular la innovación en IA fiable que se centre en cuestiones técnicas complejas y en las implicaciones sociales, jurídicas y éticas y las cuestiones normativas relacionadas con la IA.
- b) Los gobiernos también deberían considerar la inversión pública y alentar la inversión privada en herramientas de código fuente abierto y conjuntos de datos abiertos que sean representativos y respeten la privacidad y la protección de datos con vistas a sostener un entorno de investigación y desarrollo de la IA que esté libre de sesgos perjudiciales y mejore la interoperabilidad y uso del estándares.

### **2.2. Fomentar un ecosistema inclusivo que propicie la IA**

Los gobiernos también deberían fomentar el desarrollo y el acceso a un ecosistema digital inclusivo, dinámico, sostenible e interoperable en aras de una IA fiable. Un ecosistema de estas características comporta, *inter alia*, datos, tecnologías de IA, infraestructuras informáticas y de conexión y mecanismos para compartir los conocimientos sobre la IA, cuando proceda. A este respecto, los gobiernos deberían considerar la promoción de mecanismos para garantizar una difusión segura, justa, legal y ética de los datos, como los denominados *data trusts* o administradores de datos.

### **2.3. Dar forma a un entorno político y de gobernanza interoperable propicio para la IA**

- a) Los gobiernos deben promover un entorno de políticas ágil que sostenga la transición desde la fase de investigación y desarrollo hasta la fase de despliegue y explotación de sistemas de IA fiables. Para ello, deberían considerar usar la experimentación con el fin de crear un entorno controlado en el que los sistemas de IA puedan probarse y, en su caso, ampliarse. También deberían adoptar enfoques basados en los resultados que aporten flexibilidad a la hora de alcanzar los objetivos de gobernanza y cooperar tanto dentro de las jurisdicciones como entre ellas para promover entornos políticos y de gobernanza interoperables, cuando proceda.
- b) Los gobiernos deberían revisar y adaptar, cuando proceda, sus marcos políticos y normativos, así como sus mecanismos de evaluación a la hora de aplicarlos a los sistemas de IA para fomentar la innovación y la competencia en aras de una IA fiable.

### **2.4. Reforzar las capacidades del ser humano y prepararse para la transformación del mercado laboral**

- a) Los gobiernos deberían trabajar en estrecha colaboración con las partes interesadas con el fin de prepararse para la transformación del mundo laboral y de la sociedad. Deberían capacitar a las personas para usar e interactuar eficazmente con los sistemas de IA en todo el espectro de aplicaciones, dotándolas también de las competencias necesarias.
- b) Los gobiernos deberían tomar medidas, en particular a través del diálogo social, a fin de garantizar una transición justa para los trabajadores a medida que se despliegue la IA, por ejemplo mediante programas de formación a lo largo de la vida laboral, prestando apoyo a los afectados por el cambio, entre otras cosas, mediante la protección social, y brindando acceso a nuevas oportunidades laborales.
- c) Los gobiernos también deberían fijarse como meta que los efectos beneficiosos de la IA se repartan de forma amplia y justa y, a tal fin, deberían colaborar estrechamente con las partes interesadas para fomentar el uso responsable de la IA en el trabajo, para mejorar la seguridad de los trabajadores y la calidad del empleo y de los servicios públicos, y para alentar el emprendimiento y la productividad.

### **2.5. Cooperación internacional en aras de una IA fiable**

- a) Los gobiernos, incluidos los países en desarrollo y en colaboración con las partes interesadas, deberían cooperar activamente para promover estos principios y progresar en la administración

responsable en aras de una IA fiable.

b) Los gobiernos deberían trabajar juntos en el seno de la OCDE y otros foros internacionales y regionales para alentar la difusión de los conocimientos sobre la IA, cuando proceda. Deberían fomentar iniciativas internacionales, intersectoriales y abiertas entre las diferentes partes interesadas para recabar conocimientos especializados a largo plazo sobre la IA.

c) Los gobiernos deberían promover el desarrollo de normas técnicas internacionales consensuadas por diferentes partes interesadas en aras de una IA interoperable y fiable.

d) Los gobiernos también deberían impulsar el desarrollo de —y utilizar ellos mismos— indicadores comparables a escala internacional para medir la investigación, el desarrollo y el despliegue de la IA y reunir pruebas para evaluar los progresos en la implantación de estos principios.

**VI. INVITA** al secretario general y a los Adherentes a difundir esta Recomendación.

**VII. INVITA** a los no Adherentes a que tengan debidamente en cuenta la presente Recomendación y se adhieran a ella.

**VIII. ENCARGA** al Comité de Política Digital, a través de su Grupo de Trabajo sobre Gobernanza de la IA:

a) continúe su importante labor sobre la inteligencia artificial tomando como base la presente Recomendación y teniendo en cuenta el trabajo realizado en otros foros internacionales, siga desarrollando el marco de medición para las políticas de IA basadas en datos empíricos;

b) desarrolle y perfeccione nuevas orientaciones prácticas sobre la implantación de esta Recomendación para responder a la evolución de los acontecimientos y a las nuevas prioridades políticas;

c) proporcione un foro para el intercambio de información sobre políticas y actividades de IA, incluida la experiencia acumulada con la implantación de esta Recomendación, y fomente el diálogo interdisciplinar y entre las múltiples partes interesadas para alentar la confianza en la adopción de la IA; y

d) informe al Consejo, consultando con los correspondientes Comités, sobre la aplicación, difusión y pertinencia a lo largo del tiempo de la presente Recomendación, a más tardar cinco años después de su revisión y, a partir de entonces, al menos cada diez años.