

26



ÉTICA EN LA INTELIGENCIA ARTIFICIAL (IA): Análisis global de la Ética en la IA: Hacia una IA responsable, sostenible y confiable.

Fco. Javier Torres Gella



cuadernos de
ética
en clave cotidiana

ÉTICA EN LA INTELIGENCIA ARTIFICIAL (IA):

Análisis global de la Ética en la IA:
Hacia una IA responsable,
sostenible y confiable

FCO. JAVIER TORRES GELLA



Fundación Europea para el
Estudio y Reflexión Ética



Editorial Perpetuo Socorro
Covarrubias, 19. 28010 Madrid

COORDINADORA DE
“CUADERNOS DE ÉTICA EN CLAVE COTIDIANA”

M. Begoña Husillos Tamarit
Enfermera y Doctora en Teología

CONSEJO ASESOR

- *Rafael Junquera de Estéfani*
- *Antonio Fuertes Ortiz de Urbina*
- *Luis Mesa Castilla*
- *Marta Iglesias López*
- *Enrique Lluch Frechina*
- *José Luis Pareja Rivas*

DIRECCIÓN – REDACCIÓN – ADMINISTRACIÓN

Fundación Europea para el Estudio y Reflexión Ética
FUNDERÉTICA

C/ Félix Boix, 13
28036 Madrid (España)
www.funderetica.org | fundraising@funderetica.org

ISBN: 978-84-284-0876-9

DEPÓSITO LEGAL: M-2632-2026

ISSN: 2341-0388

*A mi familia, especialmente a Marina
y a mis hijos Santiago y Marina,
a quienes debo mi plenitud.*

ÍNDICE

INTRODUCCIÓN 7

1. MARCOS ESTATALES/REGIONALES EN VIGOR O EN PROYECTO 11

1.1 Unión Europea: Hacia una ética integral	11
1.2 Estados Unidos: Diversidad interna e individualismo .	16
1.3 China: Valores colectivos y control social	20
1.4 Canadá: Prevención de daños graves	23
1.5 Japón: El ser humano y la armonía social.....	25
1.6 Singapur: Gobernanza ética voluntaria	28
1.7 Corea del Sur: Modelo de los Tres pilares	30
1.8 Brasil: Siguiendo el modelo europeo	32
1.9 Perú: Promoción de IA con Responsabilidad Social....	34
1.10 Kenia: Estrategia nacional y valores africanos	36
1.11 Australia: Pragmatismo regulatorio	37
1.12 Emiratos Árabes Unidos (EAU): Estrategia híbrida .	40
1.13 Israel: Enfoque anti regulatorio horizontal	41

2. TRATADOS Y MARCOS INTERNACIONALES: HACIA UNA CONVERGENCIA ÉTICA GLOBAL DESDE 2019 HASTA 2025 45

2.1 Principios de IA, de la OCDE.....	45
2.2 Recomendación sobre Ética de IA, de la UNESCO	46

2.3 Convenio Marco sobre IA, del Consejo de Europa	49
2.4 "Governing AI for Humanity", de la ONU	51
2.5 "Antiqua et Nova", de la Iglesia Católica	55

3. ANÁLISIS COMPARATIVO DE DIMENSIONES ÉTICAS TRANSVERSALES 59

3.1 Dignidad Humana: Fundamental, pero con interpretaciones diversas	59
3.2 Supervisión Humana: De control técnico a capacidad de juicio moral	60
3.3 Transparencia y Explicabilidad: La lógica de conocer la motivación	62
3.4 Verdad y epistemología: El riesgo de la manipulación social	63
3.5 Justicia Global y Equidad: Del Principio ético al pragmatismo nacional	65
3.6 Sostenibilidad Ambiental: De preocupación marginal a valor central	67
3.7 El Problema de la Antropomorfización: Nuevas dimensiones éticas	68

4. CONCLUSIÓN HACIA UNA ÉTICA GLOBAL45

4.1 Paisaje ético actual: voces corales similares con tonos diferentes	71
4.2 Tensiones éticas fundamentales en la gobernanza de IA.....	75
4.3 Cuestiones éticas emergentes: agenda futura.....	78

4.4 De los principios a la puesta en práctica	81
4.5 Evidencias que emergen del examen realizado.....	88
4.6 La pregunta que permanece en medio de esta nube, aún difusa	92
4.7 Necesidad de la acción de todos: la llamada de la ONU en su informe.....	94

INTRODUCCIÓN

El presente artículo proporciona un análisis de cómo diferentes países y Organismos internacionales abordan la ética en la IA, destacando las convergencias y las divergencias culturales, sociales y filosóficas. Adicionalmente, este análisis ofrece perspectivas sobre las tensiones éticas subyacentes y los desafíos futuros a todos los niveles.

La regulación de la inteligencia artificial representa uno de los desafíos más significativos de nuestra época. A medida que esta nueva *era de la IA* se integra cada vez más en todos los aspectos de la vida humana (desde el otorgamiento de créditos hasta diagnósticos médicos, desde la educación hasta la contratación laboral) surge una pregunta fundamental: ¿cómo podemos garantizar que esta tecnología respete los valores éticos que

La regulación de la inteligencia artificial representa uno de los desafíos más significativos de nuestra época.

sustentan nuestras sociedades y sea responsable, sostenible y confiable?

Este artículo examina en primer lugar los marcos normativos promulgados o en proyecto en los 5 continentes, revelando tanto convergencias como divergencias significativas en su enfoque. Este análisis examina las normas desarrolladas por grandes potencias mundiales: la UE, los EEUU, China, Canadá y Japón. También examina otros países con relevancia tecnológica asiática, como Corea del Sur o Singapur; así como marcos regulatorios de países de Sudamérica (Brasil y Perú), de África (Kenia) y Australia.

En este sentido, veremos cómo, en los desarrollos normativos, la irrupción de la Administración Trump ha tenido un reflejo internacional, potenciando la competitividad económica y la innovación tecnológica frente a una gobernanza que permita la consecución de una IA responsable y confiable.

Posteriormente, se expondrá la visión global de la ética en la IA desde el punto de vista de organizaciones internacionales (ONU, UNESCO, OCDE, Consejo de Europa) de forma cronológica. Un examen de la ética quedaría incompleto sin analizar la ética promovida por la Iglesia Católica, cuyos enunciados aportan criterios humanis-

tas, comprehensivos y globalizados, teniendo en cuenta las necesidades existentes en países con menor capacidad de desarrollar dicha tecnología.

Expuestos todos los marcos normativos, Tratados y recomendaciones será posible realizar una comparativa de las dimensiones éticas que tienen un alcance general, como la Dignidad humana, la Supervisión, la Transparencia, la Epistemología, o la Justicia global.

Por último, en las conclusiones será posible concretar el paisaje actual, las tensiones éticas existentes, cuestiones éticas emergentes, la concreta puesta en práctica de principios y las evidencias que emergen de todo el examen realizado.

Las leyes sobre IA que han surgido en los últimos años no son meros ejercicios teóricos de regulación. Representan declaraciones sobre los valores éticos que cada sociedad considera prioritarios, incluso frente a un avance tecnológico imparable. La ética en la IA no es un añadido cosmético a dicha regulación; es su fundamento conceptual para desarrollar y desplegar la IA de manera responsable, sostenible y confiable.

Las leyes sobre IA que han surgido en los últimos años no son meros ejercicios teóricos de regulación. Representan declaraciones sobre los valores éticos que cada sociedad considera prioritarios.

1. MARCOS ESTATALES/REGIONALES EN VIGOR O EN PROYECTO

1.1 Unión Europea: Hacia una ética integral

a) Fundamentos Éticos del Enfoque Europeo

El Reglamento de Inteligencia Artificial de la Unión Europea (AI Act) entró en vigor en agosto de 2024. Representa el esfuerzo regulatorio más comprehensivo y éticamente fundamentado en materia de IA a nivel global. Un Grupo de Expertos de la Comisión propuso en 2019 unas "Directrices Éticas para una IA Confiable". Estos principios éticos se plasmaron en el AI Act como obligaciones legales.

b) El Enfoque basado en los efectos y en los riesgos

El AI Act no regula los sistemas de IA en sí, pues se asume que cambiarán radicalmente en el futuro. Adopta un enfoque basado en los *efectos* de los sistemas de IA y sus *riesgos*. El Reglamento clasifica los sistemas de IA en cuatro niveles de *riesgo* (inaceptable, alto, limitado y mínimo). De éstos, destacan dos categorías:

El Reglamento de Inteligencia Artificial de la Unión Europea (AI Act) entró en vigor en agosto de 2024. Representa el esfuerzo regulatorio más comprehensivo y éticamente fundamentado en materia de IA a nivel global. Un Grupo de Expertos de la Comisión propuso en 2019 unas "Directrices Éticas para una IA Confiable". Estos principios éticos se plasmaron en el AI Act como obligaciones legales.

- ✓ **Riesgo Inaceptable:** Prohíbe sistemas que son una amenaza clara para los derechos fundamentales (p. ej., sistemas de manipulación subliminal dañina). Esta prohibición refleja el compromiso ético europeo con la autonomía humana y la dignidad, enumerando usos de IA incompatibles con una sociedad democrática, independientemente de sus posibles beneficios.
- ✓ **Alto Riesgo:** Exige el cumplimiento de requisitos técnicos, como gobernanza de datos, transparencia, trazabilidad, supervisión humana, solidez y ciberseguridad. En áreas críticas como el empleo, la educación, o la migración, entre otros del Anexo III, se facilita su cumplimiento mediante estándares y mediante una autoevaluación de los proveedores.

c) **Principios éticos clave:**

- ✓ **Derechos Fundamentales:** El AI Act garantiza que los sistemas de IA respeten la dignidad humana y los Derechos de la Carta de Derechos Fundamentales de la UE.
- ✓ **Transparencia:** Los usuarios deben ser informados, como regla general, cuando inte-

ractúan con un sistema de IA para que sean conscientes de ello.

- ✓ **Explicabilidad:** La lógica de los sistemas de alto riesgo debe ser auditada y explicada a todo afectado por una decisión basada en un sistema de IA.
- ✓ **No Discriminación y Equidad:** Exige medidas para identificar y mitigar los riesgos de sesgo que puedan conducir a la discriminación.

d) Protección de derechos fundamentales y evaluaciones de impacto

El AI Act crea **Autoridades de Supervisión de la IA** para garantizar la protección de los Derechos Fundamentales, de oficio o a través de denuncias. Además, el AI Act permite publicar Directrices para aclarar a los operadores los casos de uso dudosos. Igualmente, fomenta el desarrollo de sistemas de IA en entornos de prueba controlados. Adicionalmente, la UE ha puesto en marcha el “*EU AI Act Compliance Checker*” para facilitar el cumplimiento del reglamento.

Por otro lado, el AI Act exige establecer un sistema de gestión de riesgos durante el entrenamiento de sistemas de alto riesgo, que incluye

Evaluaciones de Impacto en Derechos Fundamentales, de forma que respeten estos derechos durante todo su ciclo de vida. Estas evaluaciones se efectúan durante el entrenamiento en el Sistema de Gestión de Riesgos.

e) Supervisión humana

El Reglamento exige la supervisión humana de los sistemas de IA cuando las decisiones pueden impactar significativamente las vidas de las personas (de alto riesgo) por quienes los pongan en uso, quedando sujetas al juicio humano.

Este requisito reconoce una verdad ética: la automatización completa de decisiones es problemática porque elimina la moral humana de la deliberación. La supervisión humana no es un requisito de control de calidad; es un reconocimiento de que ciertas decisiones requieren capacidades morales y una contextualización que solo los humanos poseen.

f) Gobernanza de datos y su calidad

Dado que los datos de alta calidad son un requisito previo para modelos de IA efectivos y éticos, el AI Act establece requerimientos para los datos y su gobernanza en los sistemas de alto

riesgo. Deben contener información precisa, y mitigar los sesgos potenciales.

Ello aborda otra dimensión ética crítica: los datos defectuosos o sesgados producen sistemas de IA que perpetúan o amplifican desinformación y discriminaciones. El Reglamento establece principios como minimización de datos, limitación de propósito y calidad de datos para garantizar que se recopilan y procesan correctamente.

g) Sostenibilidad Ambiental como consideración ética emergente

El AI Act plantea que la IA debe beneficiar a todos los seres humanos, incluidas generaciones futuras. Por tanto, debe garantizarse que sea sostenible y respetuosa con el medio ambiente. Sin embargo, no aborda medidas concretas en esta materia.

h) Nuevo horizonte tras la victoria de la administración Trump en 2024

No obstante toda esta regulación, desde la victoria de la Administración Trump en 2024 se ha iniciado un movimiento dentro de la UE, por diferentes Estados Miembro, para simplificar o posponer la aplicación del Reglamento de IA y de varios más sobre datos, como el Reglamento de Datos o el de Protección de Datos.

El Colorado AI Act (2024) Es un marco arraigado en los principios de derechos civiles y de protección al consumidor.

Fruto de estas presiones, la Comisión ha publicado la propuesta de Digital ómnibus el 19 de noviembre de 2025, para unificar los Reglamentos de Datos, simplificar exigencias a las PYMES y posponer hasta diciembre de 2027 las obligaciones a sistemas de alto riesgo.

1.2 Estados Unidos: Diversidad interna e individualismo

a) La Ley de IA de Colorado centrada en la Discriminación Algorítmica

El **Colorado AI Act (2024)** representa un enfoque hacia la ética de la IA centrada en el daño potencial al consumidor y la prevención de la *discriminación algorítmica* en decisiones tomadas por sistemas de IA de *alto riesgo*. Es un marco arraigado en los principios de derechos civiles y de protección al consumidor.

Se dirige a los desarrolladores y propietarios de sistemas que tienen un impacto significativo en los consumidores (contratos de préstamos, seguros, el empleo, la vivienda). Los principios éticos clave de esta ley son:

- ✓ **Equidad y no discriminación:** Exige que los desarrolladores realicen evaluaciones de impacto anuales para identificar sesgos

algorítmicos que resulten en resultados discriminatorios basado en la edad, color, discapacidad, raza, religión, sexo, etc.

- ✓ **Transparencia con el Consumidor:** Las empresas deben informar a los consumidores sobre cómo utilizan esta IA en la toma de decisiones que les afectan, otorgándoles el derecho de corregir datos o apelar la decisión, excluyendo la IA generativa salvo que sea usada al tomar decisiones frente a consumidores.
- ✓ **Rendición de Cuentas:** Responsabiliza a las empresas por no mitigar los riesgos conocidos, sin requerir intención discriminatoria.
- ✓ **Deber de reporte:** En caso de producirse una discriminación algorítmica, se impone al desarrollador del sistema el deber de reportarlo al fiscal General. Esta obligación refleja un compromiso con la transparencia y la rendición de cuentas.

Este enfoque reconoce una evidencia sobre la discriminación algorítmica: los sistemas de IA pueden discriminar incluso aunque sus creadores no tengan intención de hacerlo, debido a sesgos en los datos de entrenamiento o en el diseño del algoritmo.

La Ley de Privacidad del Consumidor de California (CCPA), modificada por numerosas leyes, aborda la ética de la IA a través de la privacidad y los derechos del consumidor.

b) Ley de California centrada en la privacidad del individuo

La Ley de Privacidad del Consumidor, de California (CCPA), modificada por numerosas leyes sectoriales y por la Ley de Derechos de Privacidad de California (CPRA), aborda la ética de la IA a través de la privacidad y los derechos del consumidor:

- ✓ **Derecho a la transparencia en decisiones automatizadas:** La Ley otorga a los consumidores el derecho a saber si una entidad utiliza sistemas de toma de decisiones automatizadas en su contra, especialmente si tienen un impacto significativo (por ejemplo, en el empleo, el crédito o la vivienda).
- ✓ **Derecho de oposición y exclusión voluntaria (Opt-out):** Los consumidores tienen derecho a limitar o excluir el uso de réplicas suyas o de su información personal sensible, lo cual es relevante para sistemas de IA que consumen grandes volúmenes de datos.
- ✓ **Deber de reporte** por los desarrolladores de grandes sistemas de IA de incidentes críticos cuando causen la muerte o lesiones graves a personas o daños catastróficos, su incumplimiento puede implicar multas.

- ✓ **Evaluaciones de riesgo:** La norma requiere que las empresas realicen evaluaciones periódicas si su procesamiento de datos personales presenta un riesgo significativo para los derechos de privacidad y seguridad de los consumidores.
- ✓ **Minimización de datos:** exige que las empresas solo recopilen la información estrictamente necesaria y la conserven por el tiempo estrictamente necesario.

En resumen, el enfoque ético de California es reactivo y centrado en el individuo, utilizando la privacidad como palanca para imponer responsabilidad y control sobre el uso ético de la IA, con excepciones en materia de reporte de incidentes graves.

c) America's AI Action Plan (Plan de Acción de IA de EE. UU.)

La publicación del Plan de Acción de IA de EE. UU. en julio de 2025 establece un marco federal que choca frontalmente con el enfoque regulatorio de Colorado y California.

Sus principios se centran en eliminar las "*barreras burocráticas*" estatales (i.e. las leyes de Colorado y California) que considera "*onerosas*". El

Sus principios se centran en eliminar las "barreras burocráticas" estatales que considera "onerosas". El Plan promueve la innovación sin restricciones y libre de sesgos ideológicos.

Plan promueve la innovación sin restricciones, que el uso de IA sea imparcial y libre de sesgos ideológicos (especialmente contra la denominada cultura “Woke”, de ideología progresista). Establece una infraestructura nacional para mantener la seguridad y el liderazgo global.

Este Plan no anula las leyes estatales ni preva-
lece sobre ellas, ya que es una Orden ejecutiva y no una Ley del Congreso. Sin embargo, utiliza la presión económica: Amenaza con retirar fondos federales a Estados con regulaciones "onerosas", obligándoles a alinear sus marcos regulatorios con el Plan para evitar la pérdida de financiación federal. En este sentido, el Estado de Colorado ha votado retrasar la aplicación de su Ley de IA a junio de 2026, no así el de California.

1.3 China: Valores colectivos y control social

a) Marco regulatorio chino: Una perspectiva ética distinta

Su enfoque regulatorio a través de las “**Medidas para la administración de servicios de IA Generativa**” de agosto de 2023 y otras regulaciones relacionadas, es muy diferente al de Occidente. El enfoque de China se centra en el control estatal, la

seguridad nacional y la estabilidad social, al tiempo que fomenta la innovación.

b) Principios éticos clave de su marco normativo

- ✓ **Valores Socialistas Fundamentales:** El contenido generado por la IA, desde las *Normas de contenido de la red* de 2019, debe adherirse a los valores socialistas de su constitución, promoviendo el control ideológico sobre la información.
- ✓ **Seguridad y Fiabilidad:** Los proveedores deben asegurar la legalidad de las fuentes de datos y evitar generar información falsa, dañina o que incite a la subversión.
- ✓ **Doble Vía:** Hay un equilibrio delicado entre la estabilidad social y la innovación, buscando una regulación que no estrangule el desarrollo tecnológico.

La regulación tiene una fuerte impronta de "*lo que debe y no debe hacer*" un sistema de IA desde la perspectiva del Estado, especialmente la IA generativa; y se enfoca en el resultado de la IA, especialmente el contenido generado y su impacto en el orden social.

El enfoque de China se centra en el control estatal, la seguridad nacional y la estabilidad social, al tiempo que fomenta la innovación.

c) Prevención de discriminación con especificidad China

Las normas requieren tomar medidas efectivas durante el diseño, entrenamiento, y puesta en servicio para prevenir la discriminación basada en nacionalidad, creencias, país de origen, región, género, edad, ocupación y salud. La lista de categorías protegidas en China difiere de las occidentales y estas normas deben entenderse en el contexto del compromiso del Estado con la estabilidad social y el control ideológico.

d) Derechos de Propiedad Intelectual y Derechos individuales

Las normas requieren respetar la propiedad intelectual, la ética comercial y el secreto comercial, así como evitar prácticas monopolísticas o de competencia desleal. También exige respetar los derechos e intereses legítimos de terceros, abstenerse de dañar la salud y evitar violaciones de los derechos de imagen, honor, privacidad e información personal. Reflejan preocupación por la IA, aunque supeditada al interés superior del Estado.

e) Obligaciones reforzadas de transparencia y etiquetado.

Desde el 1 de septiembre de 2025 han entrado en vigor nuevas Medidas para la identificación de contenido generado por IA. La norma requiere una doble capa de transparencia y etiquetado. Debe añadirse una etiqueta visible en el contenido o en la interfaz donde se presente. Además, debe incorporar un identificador implícito en el contenido, como marcas de agua digitales incrustadas en los metadatos del archivo para rastrear su origen. Ello facilita la trazabilidad y la investigación por el Estado de los denominados *deepfakes*.

1.4 Canadá: Prevención de daños graves

En proceso de tramitación legislativa desde 2022, la Artificial Intelligence and Data Act ofrecía un enfoque centrado en "sistemas de IA de alto impacto". Esta Ley consideraba de alto impacto los sistemas que podrían causar "*daño físico o psicológico grave a individuos, daño sustancial a la salud o propiedad de individuos, o un impacto económico adverso sustancial*". Esta definición refleja un enfoque basado en los *efectos* de la IA, que prioriza la prevención de daños graves.

Esta Ley consideraba de alto impacto los sistemas que podrían causar daño sustancial a la salud o propiedad de individuos, o un impacto económico adverso sustancial". Refleja un enfoque basado en los efectos de la IA, que prioriza la prevención de daños graves.

Tras la victoria de la Administración Trump en los Estados Unidos en 2024, Canadá ha dejado en suspenso la tramitación de la Ley.

La Ley establecía obligaciones para los responsables de sistemas de alto impacto:

- **Identificar resultados perjudiciales**, evaluando daños que podría causar el sistema.
- **Implementar medidas** para identificar, evaluar y mitigar riesgos y sesgos.
- **Supervisar** para detectar daños emergentes durante el ciclo de vida del sistema.
- **Mantener registros** para documentar las evaluaciones, las medidas y el monitoreo.
- **Proporcionar transparencia**, publicando resúmenes de cómo funciona el sistema y las medidas para mitigar riesgos, con diferentes niveles (v.gr para proteger la seguridad o la propiedad industrial).
- **Establecer sistemas de gobernanza interna**, incluyendo la designación de responsables, políticas y procedimientos documentados.

No obstante, tras la victoria de la Administración Trump en los Estados Unidos en 2024, Canadá ha dejado en suspenso la tramitación de la Ley,

por lo que no se conoce a día de hoy si culminará este proceso legislativo o no.

1.5 Japón: El ser humano y la armonía social

Japón ha publicado normas que reflejan una filosofía influida por sus valores culturales. El enfoque de Japón promueve un modelo de **gobernanza empresarial basado en principios**, junto al **fomento de la innovación tecnológica** por los poderes públicos. Este marco legal refleja valores éticos, como la armonía social (*wa* 和), el equilibrio, y la consideración de impactos comunitarios colectivos junto con derechos individuales para mantener la cohesión y el bienestar social. Se articula a través de distintas leyes:

*Este marco legal refleja valores éticos, como la armonía social (*wa* 和), el equilibrio, y la consideración de impactos comunitarios colectivos junto con derechos individuales.*

Los “Principios sociales de la IA centrada en el ser humano” (2019) promueve una ética inclusiva y proactiva de la IA basada en los siguientes principios:

- ✓ **Dignidad Humana y Diversidad:** Se exige que la IA respete los Derechos Humanos y que los resultados no marginen sectores de población.
- ✓ **Seguridad:** Los sistemas de IA deben evitar daños a las personas y propiedades.

- ✓ **Transparencia y Trazabilidad:** Para generar confianza, los sistemas deberían ser trazables y explicables en lo posible.
- ✓ **Sostenibilidad y Bienestar:** La IA debe contribuir a la prosperidad económica, al desarrollo sostenible y a solucionar problemas sociales (salud, medio ambiente, etc.) beneficiando a generaciones presentes y futuras. Incluye normas sobre el envejecimiento de la población, debiendo ayudar a preservar su autonomía y dignidad.

Las “Directrices de IA para Negocios” (2025) proporcionan orientación práctica sobre cómo las empresas deben aplicar los principios éticos. Las directrices enfatizan:

- ✓ Los derechos de propiedad intelectual deben protegerse, aunque permite el uso de datos en el entrenamiento de sistemas siempre que no perjudique al titular.
- ✓ Propugna un **proceso continuo** durante todo el ciclo de vida del sistema de IA.
- ✓ El desarrollo de IA debe involucrar la **colaboración** entre proveedores, propietarios de la IA, usuarios y otros afectados.

- ✓ Las obligaciones éticas específicas dependerán del **contexto** de uso y el riesgo potencial. El Gobierno podrá iniciar **investigaciones ante incidentes graves** relacionados con la IA. según los resultados, proporcionará orientación y asesoramiento a empresas e informará al público sin establecer sanciones. Si hay daños a terceros, se recurrirá al Código Penal o al civil según corresponda.

La “Ley de promoción de tecnologías relacionadas con IA” (2025) busca traducir los principios éticos en directrices operativas que el sector público debe adoptar:

- ✓ El gobierno debe **garantizar la seguridad, la confiabilidad y la robustez** de los sistemas de IA utilizados en el sector público, cumpliendo estándares técnicos rigurosos para ser considerada ética y apta.
- ✓ El gobierno debe **promover la investigación y el desarrollo**.

En conjunto, el enfoque japonés establece que la **ética y la innovación no son mutuamente excluyentes**. Las medidas japonesas se centran en el largo plazo y en la responsabilidad compartida. El Estado debe invertir en educación, infraestructura

y directrices, lo cual permitirá que el sector privado innove en un marco de confianza social.

1.6 Singapur: Gobernanza ética voluntaria

El concepto ético de la libre competencia con la promoción de la confianza en el mercado es recogido en regulaciones de países como Singapur.

Continuando en Asia, Singapur ha adoptado un enfoque parecido al japonés, basado en marcos voluntarios en lugar de una regulación vinculante. "Propugna un concepto ético de la libre competencia con la promoción de la confianza en el mercado. El **Marco de Gobernanza Modelo para la IA** (actualizado en 2020) y la Ley para IA Generativa de mayo de 2024 representan este enfoque de "soft law". No son leyes vinculantes en sí mismas, sino un conjunto de directrices éticas y técnicas para ayudar a las empresas a implementar la IA de manera responsable y transparente, enfocado en la confianza.

El enfoque de Singapur es **pragmático y centrado en la empresa**, buscando fomentar la innovación a través de un marco de **gobernanza voluntario y adaptable**. La ética se implementa a través de la autorregulación corporativa, guiada por los principios Éticos de:

- ✓ **Explicabilidad, Transparencia y Justicia:**
Las organizaciones que usan IA en la toma

de decisiones deben garantizar que la toma de decisiones sea: **explicable** –que las razones detrás de la decisión puedan explicarse en términos no técnicos; **transparente** – informando a las personas que se está usando IA en relación con ellas y cómo las afecta; y **justa** – garantizando que las decisiones no creen impactos discriminatorios o injustos en diferentes líneas demográficas (por ejemplo, raza o sexo).

- ✓ **Diseño Centrado en el Humano:** Las soluciones de IA deben ser "humano-céntricas" – lo que significa que la protección de los intereses humanos (incluido el bienestar y la seguridad) deben ser las consideraciones primarias al diseñar, desarrollar y desplegar IA. La IA debe mejorar la calidad de vida y tomar decisiones que sean ética y socialmente aceptables.
- ✓ **Responsabilidad Organizativa:** Las empresas pueden establecer procesos internos para la toma de decisiones sobre IA, la gestión de riesgos y la supervisión. Estos procesos se refieren a la gestión de los datos, al entrenamiento, reporte de incidentes, ciberseguridad, transparencia, propiedad intelectual y sostenibilidad.

Singapur ha desarrollado *AI Verify*, un marco de prueba de gobernanza de IA y un kit de herramientas de software, proporcionando a las organizaciones medios concretos para evaluar si sus sistemas de IA cumplen con estándares éticos.

1.7 Corea del Sur: Modelo de los Tres pilares

En Corea del Sur se busca el equilibrio entre la libre competencia, la promoción estatal y la gestión de los riesgos de la IA.

En Corea del Sur se busca el equilibrio entre la libre competencia, la promoción estatal y la gestión de los riesgos de la IA. Este país asiático, como Japón y Singapur, pone el énfasis en promover un marco de confianza, sin imposiciones. La "Ley sobre el desarrollo de IA y establecimiento de confianza" (2024) integra 3 estrategias: Coordinar la *dirección gubernamental*, una *Política industrial* para promover la IA, y Regulación para gestionar los riesgos de la IA.

La competitividad económica y la innovación tecnológica son consideradas inseparables de la gobernanza ética. La ética no es vista como una restricción sobre la innovación sino como su guía interna. La ley puede requerir que los desarrolladores se alineen con principios éticos, estableciendo pocas obligaciones legales explícitas. La Ley expone los Principios de ética de la IA:

- ✓ **seguridad y confiabilidad** para garantizar que la creación y uso de IA no ponga en peligro la vida humana, salud física o bienestar mental;
- ✓ **accesibilidad** para garantizar que todos puedan utilizar productos y servicios impulsados por IA sin restricciones o inconvenientes;
- ✓ **bienestar humano:** creación y aplicación de IA para mejorar el bienestar humano.

Las compañías, universidades e institutos de investigación pueden formar Comités de Ética autónomos y privados, que pueden verificar cumplimientos y proporcionar educación ética interna. Ello refleja una aproximación que enfatiza la autorregulación y la responsabilidad organizacional sobre la fiscalización gubernamental directa.

No obstante, impone obligaciones a los sistemas de IA de "alto riesgo". Son los que tienen o pueden causar un impacto serio en la vida humana, seguridad física y derechos básicos (contratación, préstamos, dispositivos médicos, etc.). Les impone deberes de *transparencia* y de *gestión de riesgos*, e implementar medidas de *mitigación*. A los sistemas de IA generativa también les impone el deber de *etiquetar* el contenido generado por IA

y dar *explicaciones* de cómo tomaron decisiones que afectan a humanos.

1.8 Brasil: Siguiendo el modelo europeo

Propuso en 2023 un modelo seguidor del europeo, aunque no ha llegado a aprobar el proyecto de Ley.

Brasil propuso en 2023 un modelo seguidor del europeo, aunque no ha llegado a aprobar el proyecto de Ley que regula la IA. El **Proyecto** de Ley en trámite en la Cámara de Diputados, intenta crear un marco regulatorio inspirado en el AI Act, pero adaptado al contexto brasileño. Establece principios éticos fundamentales que deben guiar el desarrollo y uso de sistemas de IA en Brasil:

- **Centralidad del ser humano:** Los sistemas de IA deben respetar la dignidad, los derechos y las libertades fundamentales de las personas.
- **Desarrollo sostenible:** Promoción del desarrollo económico inclusivo y la protección del medio ambiente como principio de justicia intergeneracional.
- **No discriminación e igualdad:** Deben promover la igualdad y prevenir sesgos injustos basados en raza, género, origen étnico, orientación sexual, religión, etc.

- **Transparencia y explicabilidad:** Reconoce un derecho a información clara sobre el uso de sistemas y a explicaciones sobre las decisiones automatizadas.
- **Seguridad y privacidad:** Los sistemas de IA deben operarse de manera segura, protegiendo datos personales y respetando la privacidad de los individuos.

Adicionalmente, el proyecto brasileño adopta un enfoque basado en riesgo similar al AI Act, estableciendo obligaciones diferenciadas según el nivel de riesgo:

- **Sistemas de IA de “Riesgo Excesivo”:** Ciertos usos están prohibidos, incluyendo sistemas que manipulan comportamiento humano de manera subliminal o explotan vulnerabilidades de grupos específicos para causar daño físico o psicológico.
- **Sistemas de “Alto Riesgo”:** Requieren evaluaciones de impacto, documentación extensiva, supervisión humana, y cumplimiento con estándares técnicos específicos.

Perú promulga una Ley para el desarrollo responsable y la gestión de los riesgos de la IA, pero sin establecer autoridades ni sanciones.

1.9 Perú: Promoción de IA con Responsabilidad Social

Perú promulga una Ley para el desarrollo responsable y la gestión de los riesgos de la IA, pero sin establecer autoridades ni sanciones especiales en caso de incumplimiento. La **Ley N° 31814**, (2023) y su Reglamento de desarrollo (2025) establecen un marco para promover el uso de la inteligencia artificial con el objetivo del desarrollo económico y social a través de la IA. La ley peruana establece principios éticos fundamentales que deben guiar el desarrollo y uso de IA:

- ✓ **Derechos Humanos:** La IA debe usarse respetando los derechos humanos y libertades fundamentales de su Constitución y de Tratados internacionales.
- ✓ **Transparencia y Explicabilidad:** Se promueve la comprensibilidad de los sistemas de IA de riesgo alto, permitiendo a los individuos entender decisiones que les afectan.
- ✓ **Responsabilidad:** Debe existir claridad sobre quién es responsable por las decisiones y acciones de los sistemas de IA, manteniendo la capacidad de rendición de cuentas.

- ✓ **Enfoque Antropológico:** La IA debe estar siempre al servicio de la humanidad, priorizando el bienestar social y el desarrollo sostenible del país.
- ✓ **No discriminación:** Los sistemas de IA no deben generar discriminación o sesgos.
- ✓ **Inclusión digital y reducción de brechas de acceso:** Un aspecto distintivo del enfoque peruano es su aplicación a **contextos de desarrollo**. Aborda cómo se usa la tecnología y además quién tiene acceso a ella. Por ello, promueve programas de alfabetización y capacitación en IA para vulnerables; que la IA considere la diversidad lingüística y cultural; y que sus beneficios alcancen a todas las regiones del país.

Clasifica los sistemas por su riesgo: (i) “*Uso indebido*” o prohibido, cuando impacta de manera irreversible, significativa y negativa en derechos fundamentales o en el bienestar de las personas. (ii) “*Uso de riesgo alto*”, cuando suponen un riesgo para la vida, dignidad y demás derechos fundamentales. Y el resto (iii) “*Uso de riesgo aceptable*”. Ante incumplimientos de la norma, la Autoridad de IA puede derivar el caso a la entidad competente en materia de datos, propiedad intelectual, etc.

Kenia busca alinear el desarrollo de la IA con valores africanos y con las necesidades de desarrollo específicas del continente africano.

1.10 Kenia: Estrategia nacional y valores africanos

La **Estrategia Nacional de IA de Kenia 2025-2030**, publicada en marzo de 2025, busca alinear el desarrollo de la IA con valores africanos, con las necesidades de desarrollo específicas del continente africano:

- ✓ **Ubuntu** —"Soy porque nosotros somos"— Este concepto proporciona un fundamento filosófico que difiere tanto del individualismo occidental como del colectivismo autoritario. Enfatiza interdependencia, comunidad y responsabilidad mutua, fortaleciendo lazos sociales, no solo optimizando la eficiencia individual.
- ✓ **Soberanía de datos:** refleja preocupaciones postcoloniales sobre extracción de recursos (en este caso, datos) uso de las infraestructuras tecnológicas por poderes extranjeros. La ética de la IA en contexto africano debe abordar legados de colonialismo y garantizar que la IA no replique patrones de dependencia y explotación.
- ✓ **Desarrollo inclusivo:** La IA debe ser una herramienta para reducir desigualdades, no ampliarlas. Esto incluye atención particular

a comunidades rurales, poblaciones pastoriales, y grupos tradicionalmente marginalizados.

- ✓ **Alfabetización digital diferenciada:** Reconocimiento de que poblaciones con diferentes niveles de alfabetización tecnológica requieren diferentes enfoques para consentimiento informado y transparencia.
- ✓ **Gobernanza Participativa:** Las comunidades locales tienen voz al implementar sistemas de IA en sus contextos, pudiendo supervisarla, modificarla o rechazarla. Este enfoque va más allá del modelo experto-regulador típico de marcos occidentales.

No obstante estos principios, un informe de la UNESCO de 2025 afirma que Kenia carece de los marcos de gobernanza necesarios para proteger y promover los derechos humanos dentro del ecosistema de la IA, no pudiendo garantizar los principios de una IA responsable en todas las fases del ciclo de vida y la cadena de valor de la IA.

1.11 Australia: Pragmatismo regulatorio

La legislación australiana sobre inteligencia artificial (IA) se basa actualmente en un **enfoque**

voluntario, más que en un marco vinculante integral. En 2019 Australia adoptó los *Australia's AI Ethics Principles*, que establecen ocho principios básicos:

- ✓ **Bienestar humano, social y ambiental:** Los sistemas de IA deben beneficiar a individuos, sociedad y medio ambiente.
- ✓ **Valores humanos;** deben informar el diseño, desarrollo, despliegue y uso de IA.
- ✓ **Equidad:** Los sistemas de IA deben ser inclusivos, accesibles y no discriminatorios.
- ✓ **Protección de la privacidad y seguridad de los datos:** Deben respetarla y mejorarla.
- ✓ **Fiabilidad y seguridad operativa:** Deben operar confiablemente según su propósito.
- ✓ **Transparencia y explicabilidad:** Las personas deben conocer cuándo interactúan con sistemas de IA y puedan entender resultados.
- ✓ **Posibilidad de impugnar resultados** cuando un sistema de IA impacta significativamente a una persona, comunidad, grupo o ambiente; y

- ✓ **Rendición de cuentas:** Quienes diseñan, desarrollan, despliegan y operan sistemas de IA deben ser responsables de su operación continua.

Sin embargo, estos principios tienen carácter “aspiracional” y voluntario: no crean obligaciones legales directas. Buscan guiar a organizaciones públicas y privadas a reflexionar sobre los impactos de sus sistemas de IA.

Adicionalmente, Australia ha publicado *estándares* para el desarrollo y uso de IA por Entidades públicas, incluyendo normas sobre gobernanza, ciclo de vida del sistema, documentación, pruebas y supervisión humana, así como un marco de “implementación” que exige evaluaciones de impacto de derechos humanos, auditorías de sesgo, control de datos, pruebas piloto, monitorización continua, transparencia y vías de recurso.

En resumen, el enfoque australiano prioriza valores éticos amplios, proporciona mecanismos de gobernanza y responsabilidad, pero deja en la persona afectada el ejercicio de acciones para corregir decisiones erróneas o remediar impactos de la IA.

El enfoque australiano prioriza valores éticos amplios, proporciona mecanismos de gobernanza y responsabilidad, pero deja en la persona afectada el ejercicio de acciones para corregir decisiones erróneas o remediar impactos de la IA.

El marco normativo de los EAU está compuesto por una combinación de Leyes vinculantes y guías voluntarias dirigidas a promover la ética de los sistemas de IA.

1.12 Emiratos Árabes Unidos (EAU): Estrategia híbrida

El marco normativo de los EAU en IA está compuesto por una combinación de Leyes vinculantes (v.gr. datos personales), y guías voluntarias dirigidas a promover la ética de los sistemas de IA.

En cuanto a los mecanismos vinculantes, la Ley de Protección de Datos de EAU **regula** el procesamiento automatizado o la elaboración de perfiles. En todo caso, esta ley **prohíbe** desde 2025 la difusión de información errónea o vejatoria mediante IA Generativa, y considera contraria a la ética la toma decisiones sobre la vida humana sin intervención humana.

Por otro lado, las AI Ethics Principles & Guidelines (2022) y la Charter for the development & use of AI (2024) promueven numerosos valores, priorizando:

- ✓ **Human Well-being y Derechos Fundamentales:** Enfoca la IA al beneficio humano y proteger sus derechos.
- ✓ **Seguridad y Confiabilidad:** Garantizar que los sistemas operen de forma robusta y mitiguen los riesgos.

- ✓ **Transparencia y Explicabilidad:** Fomentar la comprensión sobre el funcionamiento y las decisiones de la IA.
- ✓ **Responsabilidad y Rendición de Cuentas:** Establece quién es responsable de los resultados de la IA.
- ✓ **Equidad e Inclusividad:** pretende evitar el sesgo algorítmico.

Para promover estos principios éticos, las Guías establecen una herramienta de autoevaluación ética disponible para desarrolladores / operadores de sistemas de IA que permite evaluar riesgos relacionados con datos, sesgos, etc. Además, se contempla una certificación de sistemas de IA éticamente aceptables para incentivar su cumplimiento.

En resumen, los EAU siguen una estrategia híbrida: promover la innovación en IA estableciendo guías y ciertas obligaciones legales; si bien, la reclamación civil o penal en caso de perjuicios por sistemas de IA se deja a los perjudicados.

1.13 Israel: Enfoque anti regulatorio horizontal

El Modelo Israelí se alinea con el enfoque seguido por la Administración Trump en materia

La prosperidad nacional y el liderazgo tecnológico son considerados por Israel como valores esenciales en sí mismos, no meramente objetivos instrumentales.

de IA. Rechaza legislar sobre IA, favoreciendo intervenciones sectoriales específicas y la promoción voluntaria de valores mediante estándares industriales voluntarios. No obstante, dada su situación geopolítica, la prosperidad nacional y el liderazgo tecnológico son considerados por Israel como valores esenciales en sí mismos, no meramente objetivos instrumentales.

Este modelo refleja las prioridades nacionales específicas: Al depender fuertemente de la innovación tecnológica, Israel es especialmente sensible a riesgos de regulación que podría obstaculizar la competitividad.

La *Política sobre Regulación y Ética* identifica siete desafíos principales que surgen del uso de IA en el sector privado: discriminación, supervisión humana, explicabilidad, divulgación de interacciones con IA, seguridad, responsabilidad y privacidad. Para cada desafío, la norma proporciona recomendaciones de mejores prácticas, pero no impone obligaciones legales vinculantes ni establece responsabilidades más allá de las previstas con carácter general en sus normas de responsabilidad criminal o civil.

Sin embargo, en 2024 Israel firmó la *ConvenCIÓN Marco sobre IA* del Consejo de Europa, que

ahora examinaremos, junto con otros países, comprometiéndose con el primer Tratado internacional vinculante sobre IA y derechos humanos.

2. TRATADOS Y MARCOS INTERNACIONALES: HACIA UNA CONVERGENCIA ÉTICA GLOBAL DESDE 2019 HASTA 2025

2.1 Principios de IA, de la OCDE

Los Principios de IA de la OCDE, adoptados en 2019 y actualizados en 2024, representan el primer estándar intergubernamental sobre IA. Los cinco principios son:

- ✓ **Crecimiento inclusivo, desarrollo sostenible y bienestar:** La IA debe beneficiar a personas y al planeta aumentando las capacidades humanas y su creatividad.
- ✓ **Valores centrados en el ser humano y equidad:** Los sistemas de IA deben respetar Estado de derecho, derechos humanos, valores democráticos y diversidad, e incluir salvaguardas apropiadas.
- ✓ **Transparencia y explicabilidad:** Debe haber transparencia y garantizar que las personas entiendan resultados basados en IA
- ✓ **Robustez y seguridad:** Los sistemas de IA deben funcionar robustamente de manera segura a lo largo de sus ciclos de vida.

Los Principios de IA de la OCDE, adoptados en 2019 y actualizados en 2024, representan el primer estándar intergubernamental sobre IA.

- ✓ **Responsabilidad:** Las organizaciones e individuos que desarrollan, despliegan u operan sistemas de IA deben ser responsables de su operación apropiada.

Estos principios han sido influyentes, informando marcos nacionales alrededor del mundo. Su adopción por 51 países representa una convergencia ética significativa sobre cómo la IA debe ser gobernada.

Los Principios de IA de la OCDE son un instrumento de "soft law" ("derecho suave"), lo que significa que no son un tratado internacional vinculante ni crean obligaciones legales. Sin embargo, son muy influyentes al ser una recomendación política adoptada por consenso y un estándar voluntario desarrollado con una amplia consulta experta. Por ello, los países firmantes se comprometen a alinear estos principios con sus políticas nacionales. V.gr. Japón, Singapur y Corea del Sur los han integrado en sus directrices.

En noviembre de 2021, los 193 Estados que son parte de la UNESCO adoptaron el primer instrumento normativo global sobre ética de IA.

2.2 Recomendación sobre Ética de IA, de la UNESCO

En noviembre de 2021, los 193 Estados que son parte de la UNESCO adoptaron la "Recomendación sobre Ética de Inteligencia Artificial", el primer instrumento normativo global sobre ética de

IA. La Recomendación articula valores y principios compartidos que pueden guiar la construcción de un marco legal necesario para garantizar el desarrollo de la IA. Los diez principios centrales incluyen:

1. **Proporcionalidad e inocuidad:** El uso de sistemas de IA no debe ir más allá de lo necesario para alcanzar un objetivo legítimo.
2. **Seguridad y protección:** Los daños no deseados (riesgos de seguridad) y las vulnerabilidades a los ataques (riesgos de protección) deberían ser evitados y tomados en consideración.
3. **Derecho a la intimidad y protección de datos:** La privacidad debe protegerse y promoverse a lo largo de todo el ciclo de vida de la IA.
4. **Gobernanza y colaboración adaptativas** entre todos los interesados afectados.
5. **Responsabilidad y rendición de cuentas:** Los sistemas de IA deben ser auditables y trazables. *Deben existir mecanismos de supervisión, evaluación de impacto, auditoría y diligencia debida.*

6. **Transparencia y explicabilidad**, adecuada al contexto.
7. **Supervisión y decisión humanas**: Debería ser posible atribuir la responsabilidad ética y jurídica a personas físicas o a entidades jurídicas existentes.
8. **Sostenibilidad**: Las tecnologías de IA deben evaluarse en función de su impacto en la "sostenibilidad", entendida como un conjunto de objetivos en constante evolución, incluidos los ODS.
9. **Sensibilización y educación**: La sensibilización y la comprensión de la IA y el valor de los datos deberían promoverse.
- 10. Equidad y no discriminación**: Los actores de la IA deberían promover la justicia social, salvaguardar la equidad y luchar contra todo tipo de discriminación.

Lo distintivo del marco de la UNESCO es su énfasis en consideraciones de desarrollo global y justicia intergeneracional. Reconoce explícitamente que la ética de IA debe abordar desigualdades entre y dentro de países, y considerar impactos en generaciones futuras.

Esta Recomendación tampoco es un tratado internacional vinculante ni crea obligaciones legales. Es un compromiso moral y político de los Estados Miembro, un marco de referencia ético para el desarrollo de las políticas nacionales utilizable como Guía de principios y valores compartidos.

2.3 Convenio Marco sobre IA, del Consejo de Europa

En septiembre de 2024, el Consejo de Europa abrió para firma el primer **Convenio Marco internacional sobre IA, Derechos Humanos, Democracia y Estado de Derecho**. Este Tratado vinculante representa un hito histórico: Es el primer instrumento jurídico internacional que crea obligaciones legales vinculantes y que requiere que las Partes garanticen que la actividad durante el ciclo de vida de los sistemas de IA será consistente con la protección de los derechos humanos, la democracia y el Estado de derecho.

El Convenio establece que los Estados adoptarán o mantendrán medidas para que los sistemas de IA:

1. Respeten los *Derechos Humanos, la dignidad humana y la autonomía individual*.
2. Respeten la integridad de los *procesos democráticos y el estado de derecho*.

*El Convenio
Marco internacional sobre IA,
Derechos Humanos, Democracia
y Estado de Derecho.
Representa un
hitos histórico:
Es el primer
instrumento jurídico internacional
que crea obligaciones legales
vinculantes.*

3. Sean *transparentes* y estén sujetos a *supervisión* según el contexto.
4. Rindan *cuentas* y *respondan* por impactos en derechos humanos, la democracia y el estado de derecho.
5. Respeten la *igualdad* y la *no discriminación*.
6. Protejan la *privacidad* y los datos personales.
7. Sean *seguros* y estén sujetos a *evaluaciones* de riesgo.
8. Puedan ser desarrollados en *entornos controlados* bajo supervisión de las autoridades.
9. Sean *explicables* en las decisiones que puedan impactar significativamente derechos humanos y que puedan ser *impugnadas* por los afectados.

El Convenio Marco del Consejo de Europa es vinculante en el sentido de que crea obligaciones legales internacionales, pero utiliza mecanismos de "soft law enforcement" (aplicación suave) en lugar de sanciones coercitivas. Estos mecanismos son *Conferencias* para posible cooperación y el deber de informar periódicamente los avances efectuados por cada parte. Constituye más un ins-

trumento de armonización de principios y estándares mínimos que un régimen regulatorio con capacidad sancionadora.

No obstante, su fortaleza radica en que establece un consenso internacional sobre **principios éticos fundamentales**, creando la obligación legal de respetar los derechos humanos en el desarrollo de la IA y sentando las bases para desarrollos regulatorios nacionales más específicos, como el garantizar una transparencia y supervisión apropiada o realizar evaluaciones de impacto en derechos humanos para sistemas de alto riesgo.

En octubre de 2025 suman 37 signatarios (con los 27 Estados de la UE), e incluye países no miembros del Consejo de Europa: como Reino Unido, Israel, Estados Unidos, Canadá o Japón. Sin embargo, aún no se han alcanzado las 5 ratificaciones necesarias (con al menos 3 de Estados miembros del Consejo de Europa) para que el Convenio entre en vigor.

2.4 "Governing AI for Humanity", de la ONU

En septiembre de 2024, la ONU publicó "*Gobernar la IA para la Humanidad*", el esfuerzo más ambicioso para articular una visión global de gobernanza ética de IA. Establece 5 principios

En septiembre de 2024, la ONU publicó "Gobernar la IA para la Humanidad", el esfuerzo más ambicioso para articular una visión global de gobernanza ética de IA.

rectores de la gobernanza de la IA y 7 recomendaciones para reducir brechas de Gobernanza global.

Parte de una premisa ética: Para mitigar los riesgos de la IA, abordar casos de uso novedosos y garantizar que la IA pueda beneficiar a toda la humanidad, los **derechos humanos** y el ser humano deben estar en el centro de toda gobernanza de la IA (**antropocéntrica**). Ello incluye (i) garantizar la igualdad y la no discriminación, (ii) prevenir o mitigar los sesgos algorítmicos, y (iii) proteger la privacidad y los datos personales.

Establece cinco **Principios rectores de la Gobernanza de la IA**:

1. **Inclusiva:** Debe tener en cuenta todas las regiones, culturas y sectores, pues la toma de decisiones está concentrada en unos pocos países y compañías.
2. **En interés público:** Los sistemas de IA deben servir al bien público y a necesidades sociales, no solo intereses comerciales o estatales
3. Alineada con la **gobernanza de datos** y la promoción de **datos comunes**.
4. Debe ser **Universal**, en red, y basarse en la colaboración de las partes interesadas.

5. Anclada en la **Carta de la ONU**, los Derechos Humanos y los ODS. Evita la trampa de tratar la IA como un dominio completamente nuevo que requiere una reinvención ética. Aplica los principios éticos y legales ya establecidos a esta nueva tecnología.

Ofrece recomendaciones a los Estados para reducir brechas de Gobernanza global:

1. Establecer un **Panel científico internacional** e independiente sobre la IA
2. Establecer un **Foro de Política global sobre IA**, inaugurado en octubre de 2025.
3. **Intercambiar Estándares sobre IA** nacionales e internacionales.
4. **Desarrollar las capacidades nacionales** de gobernanza de IA, proporcionando apoyo técnico y financiero, especialmente en países en desarrollo.
5. Crear un **Fondo global** para el desarrollo de la IA
6. Establecer un **Marco global de datos** de entrenamiento, que equilibre innovación con la propiedad intelectual y la sostenibilidad.
7. Crear una **Oficina de IA** en la ONU.

Contribuciones éticas clave del Informe:

- ✓ **Enfoque del riesgo basado en vulnerabilidad:** La categorización de los riesgos debe considerar no solo su gravedad, sino también *quién* es más vulnerable a ellos. Prioriza la protección de comunidades marginadas, poblaciones en países en desarrollo y otros grupos vulnerables.
- ✓ **Justicia global:** Pone el énfasis en la necesidad de abordar las asimetrías de poder y desigualdades estructurales entre Norte y Sur Global en la gobernanza de la IA.
- ✓ **Coordinación contra la fragmentación:** La elaboración de muchos documentos sobre IA por organismos de la ONU crean confusión en los Estados y permite que actores poderosos busquen los regímenes regulatorios más favorables.
- ✓ **IA militar:** Advierte en particular sobre las carreras armamentísticas, la reducción del umbral de conflicto y la erosión del derecho internacional humanitario, llamando a mantener un control *humano* significativo sobre uso de armas autónomas letales.

La Resolución de la ONU es un toque de atención global: la IA debe ser una fuerza positiva para el desarrollo sostenible, pero su potencial solo se realizará si se mantiene dentro de un marco de principios éticos universales. Exige un equilibrio entre la innovación tecnológica y la protección de los derechos humanos y la estabilidad global.

2.5 "Antiqua et Nova", de la Iglesia Católica

"**Antiqua et nova**" (Lo antiguo y lo nuevo), aprobada por el Papa Francisco en enero de 2025, es una Nota de reflexión teológica, filosófica y antropológica sobre la IA. Su objetivo es proporcionar una base ética para guiar el desarrollo de la IA, asegurando que esta tecnología sirva a la dignidad humana y al bien común, a la luz de la religión cristiana.

Considera que la inteligencia del ser humano es un *don de Dios* pues fuimos "creados a imagen de Dios". A diferencia de la IA, la inteligencia humana es espiritual y orientada al amor.

La IA puede imitar funciones cognitivas, pero carece de dimensiones morales o espirituales. La IA es un "producto técnico extraordinario", una

Su objetivo es proporcionar una base ética para guiar el desarrollo de la IA, asegurando que esta tecnología sirva a la dignidad humana y al bien común, a la luz de la religión cristiana.

herramienta sofisticada para el procesamiento de datos y la resolución de tareas específicas. Sin embargo, la IA carece de la riqueza de la comunicación, la emoción, el juicio moral y la apertura a la verdad y al bien (lo trascendente). Representar a la IA como persona (*Antropomorfización*) es una "violación ética grave" que erosiona la confianza social y explota vulnerabilidades.

A partir de esta visión **antropológica**, el documento establece principios éticos pragmáticos y universales:

- ✓ **Responsabilidad humana:** La IA debe permanecer bajo supervisión humana; rechazo del determinismo tecnológico
- ✓ **Transparencia y explicabilidad** del uso de la IA en todas las fases del proceso de toma de decisiones.
- ✓ **Verdad:** Contenido generado por IA debe identificarse claramente; expresa su preocupación por los *deepfakes* para engañar o manipular en contextos como la educación, las relaciones personales o la política.
- ✓ **Dignidad y justicia:** Las aplicaciones deben promover equidad y evitar discriminación

- ✓ **Supervisión Humana Significativa:** Se exige que el ser humano mantenga la capacidad de **juicio moral** y de **intervención** en las decisiones críticas, especialmente donde está en juego la vida o los derechos fundamentales.
- ✓ **Vigilancia y Privacidad:** El texto condena el uso de la IA para el control social invasivo, y aboga por un mayor respeto a los **datos personales**, que son vistos como una extensión de la naturaleza relacional de la persona.

Alarma ante las Compañías Tecnológicas: El Vaticano señala con preocupación que el desarrollo de la IA está concentrado en un pequeño grupo de empresas poderosas, lo que constituye un riesgo inherente al limitar la influencia y el acceso a esta tecnología, aumentando desigualdades.

No obstante, expresa **optimismo** pues si se dirige la IA hacia el bien común, puede ayudar a reducir pobreza, dar acceso a educación, mejorar diagnósticos de salud, promover sostenibilidad ambiental e incluso fomentar comprensión más profunda de nuestras tradiciones culturales y espirituales.

La Contribución de “Antiqua et Nova” aporta una perspectiva teológica única: La verdadera medida del progreso tecnológico no reside en la

Aporta una perspectiva teológica única: La verdadera medida del progreso tecnológico no reside en la eficiencia de la máquina, sino en su capacidad para promover la solidaridad, el cuidado de los vulnerables y la sabiduría, que son cualidades que la tecnología no puede replicar.

eficiencia de la máquina, sino en su *capacidad* para promover la solidaridad, el cuidado de los vulnerables y la sabiduría, que son cualidades que la tecnología no puede replicar.

Su relevancia viene dada por influir moralmente en más de mil trescientos millones de católicos en todo el mundo.

3. ANÁLISIS COMPARATIVO DE DIMENSIONES ÉTICAS TRANSVERSALES

3.1 Dignidad Humana: Fundamental, pero con interpretaciones diversas

La Dignidad humana emerge como el principio ético más universalmente invocado, pero con interpretaciones significativamente diferentes:

La Dignidad humana emerge como el principio ético más invocado, pero con interpretaciones significativamente diferentes.

Interpretación de Derechos Humanos (ONU, Consejo de Europa, UE): Dignidad humana se materializa a través de *derechos humanos*. Violaciones de dignidad son violaciones de derechos que pueden ser remediadas a través de mecanismos legales.

Interpretación Teológica (Vaticano): Dignidad humana deriva de ser creado a imagen de Dios y es inalienable e incondicional. Esta fundamentación teológica proporciona una base absoluta para ciertos límites éticos: algunos usos de IA son simplemente incompatibles con la dignidad humana independientemente de consecuencias.

Interpretación Cultural (Kenia, marcos africanos): Dignidad se entiende a través de concepto de Ubuntu — "Soy porque nosotros somos".

Estas diferentes concepciones de dignidad humana conducen a diferentes prioridades regulatorias.

La dignidad es fundamentalmente relacional y comunitaria, no solo individual.

Interpretación de Armonía Social: Dignidad está vinculada con membresía en sociedad armónica bajo dirección del Estado (China) O también, la dignidad individual es importante pero no puede separarse del bienestar colectivo (Japón).

Estas diferentes concepciones de dignidad humana conducen a diferentes prioridades regulatorias. Por ejemplo, la prohibición del AI Act de sistemas de puntuación social (“Social Scoring”) refleja concepto de dignidad centrado en autonomía individual, mientras que el uso de China de sistemas similares (“Gran Hermano”) refleja concepto de dignidad vinculado con armonía social.

No se trata simplemente de mantener control técnico sobre sistemas de IA, sino de mantener la capacidad de juicio moral y de intervención en las decisiones críticas.

3.2 Supervisión Humana: De control técnico a capacidad de juicio moral

La publicación de "Governing AI for Humanity" y "Antiqua et Nova" en 2025 revela una dimensión profunda del principio de "supervisión humana". No se trata simplemente de mantener control técnico sobre sistemas de IA, sino de mantener la *capacidad de juicio moral* y de *intervención* en las decisiones críticas, especialmente donde está en juego la vida de seres humanos o los derechos fundamentales.

Por un lado, el informe de la ONU enfatiza la supervisión humana como mecanismo de *rendición de cuentas* basada en *derechos*. La supervisión humana es necesaria para garantizar que cuando los sistemas de IA causen daño, haya humanos identificables que puedan ser *responsabilizados*.

Por su parte, Antiqua et Nova va más allá, argumentando que ciertas decisiones (particularmente aquellas que involucran vida y muerte) no deben jamás ser delegadas a máquinas, independientemente de cuán sofisticada sea la supervisión técnica. Esto refleja una posición ética deontológica: algunas formas de delegación son intrínsecamente incorrectas porque violan la dignidad humana.

Si comparamos la regulación de los marcos nacionales, el concepto es muy diferente:

- El **AI Act** requiere supervisión humana para sistemas de alto riesgo, pero permite una considerable automatización con supervisión apropiada.
- **China** enfatiza la supervisión, pero dentro del contexto de control estatal último.
- **Colorado** requiere una revisión humana de decisiones adversas cuando sea "técnicamente factible", reconociendo limitaciones prácticas.

- **Corea del Sur** menciona la supervisión como principio, pero deja implementación flexible.
- **Singapur y Japón** articulan la supervisión humana como valor pero confían en autorregulación.

El espectro va desde posiciones deontológicas absolutas (ciertas decisiones nunca pueden automatizarse) hasta enfoques pragmáticos (la automatización es aceptable si los resultados son buenos y hay posibilidad de recurrir ante errores).

3.3 Transparencia y Explicabilidad: La lógica de conocer la motivación

Todos los marcos nacionales valoran la transparencia como un principio ético relevante en la IA, pero difieren significativamente en qué significa y cuánta se requiere:

Todos los marcos nacionales valoran la transparencia como un principio ético relevante en la IA, pero difieren significativamente en qué significa y cuánta se requiere.

- **Transparencia estratificada (UE):** Establece niveles de transparencia según categoría de riesgo, con obligaciones de documentación y un derecho del afectado a la explicación por la decisión para sistemas de alto riesgo, pero con requisitos sencillos para sistemas de menor riesgo (solo ha de informar que está originado por IA si interactúa con un ser humano).

- **Transparencia con control estatal (China):** Divulgación obligatoria pero dentro de un marco donde el gobierno mantiene acceso privilegiado a información.
- **Máxima divulgación (Australia y Colorado):** Requiere explicaciones claras de cómo se tomaron decisiones adversas, enfatizando la explicabilidad individual.
- **Transparencia condicionada (Canadá):** Su proyecto de Ley enfatiza el dar "información en lenguaje sencillo" pero acepta que la transparencia varíe según el contexto.
- **Principios programáticos (Singapur, Japón, EUA):** Son marcos voluntarios que enfatizan la transparencia como valor, pero dejan los detalles de su implementación a organizaciones.

Estas diferencias reflejan tensiones éticas: Una transparencia absoluta puede entrar en conflicto con la propiedad intelectual, la seguridad y la privacidad. Por ello, los marcos nacionales resuelven estas tensiones de forma diferente.

3.4 Verdad y epistemología: El riesgo de la manipulación social

La dimensión epistemológica de la ética de IA (cómo la IA afecta nuestras capacidades para co-

La dimensión epistemológica de ética en la IA emerge como preocupación central en documentos internacionales.

nocer la verdad) emerge como preocupación central en documentos internacionales:

“Governing AI for Humanity” identifica la desinformación generada por IA como una amenaza para la democracia y el discurso público. La preocupación no es tanto epistémica (¿cómo sabemos qué es verdad?) como *política* (¿cómo puede funcionar la democracia cuando la verdad compartida es erosionada o manipulada?).

“Antiqua et Nova” enfatiza que la IA puede erosionar la confianza de los ciudadanos a través de *deepfakes*, desinformación y contenido sintético mal etiquetado. En la tradición católica, la verdad no es instrumental sino fundamentalmente valiosa. El engaño, por lo tanto, no es solo dañino en sus consecuencias, sino intrínsecamente incorrecto.

En cuanto a los marcos nacionales, se regula de forma igualmente diferente:

- La UE requiere etiquetar contenido generado por IA en las ultrasuplantaciones (contenido que se asemeja mucho a personas, objetos, lugares, o sucesos reales y que puede inducir a personas a pensar que son verídicos).
- **China** requiere el etiquetado, pero también permite un control estatal extensivo sobre la información.

- **Colorado** no aborda directamente desinformación, centrándose en discriminación.
- **Singapur** incluye transparencia de contenido como una de sus nueve dimensiones de confianza

La tensión ética aquí es entre *libertad de expresión* y *evitar la desinformación social*. Las naciones tratan estos valores de forma muy distinta, con China priorizando el control de la información, Estados Unidos priorizando la libertad de expresión, y Europa buscando un equilibrio a través de regulación únicamente de las ultrasuplantaciones.

La tensión ética aquí es entre libertad de expresión y evitar la desinformación social.

3.5 Justicia Global y Equidad: Del Principio ético al pragmatismo nacional

"Governing AI for Humanity" coloca la justicia global en el centro. En primer lugar, denuncia la Brecha Norte-Sur: El informe de la ONU cuestiona que los beneficios de la IA fluyan abrumadoramente hacia países ricos, mientras que los riesgos se distribuyen más uniformemente. Esta asimetría plantea profundas cuestiones de justicia. Además, pone el énfasis en la necesidad de financiar y desarrollar capacidades nacionales de gobernanza en los países en desarrollo, pues las políticas sin capacidad de implementación son vacías.

El informe de la ONU cuestiona que los beneficios de la IA fluyan abrumadoramente hacia países ricos, mientras que los riesgos se distribuyen más uniformemente.

“Antiqua et Nova” defiende como principio ético pragmático y universal la justicia, la cual se alcanzará cuando la IA sirva a la dignidad humana y al bien común, a la luz de la tradición cristiana.

En cuanto a la regulación de los marcos nacionales, el concepto es nuevamente diferente:

- **UE** tiene mecanismos de cooperación al desarrollo, pero no integrados en el AI Act.
- **Estados Unidos, Israel y Corea** se centran en la competitividad nacional, frente a una redistribución global.
- **Australia y EUA**: promueven los valores y derechos humanos para el beneficio de los individuos y de la sociedad.
- **Perú y Brasil** integran objetivos de desarrollo inclusivo en sus marcos de IA.
- **Kenia** enfatiza la soberanía de datos y el control africano sobre recursos de IA.

Se plantea, pues, una tensión ética fundamental entre la IA como *herramienta de competencia económica nacional* versus *recurso global* que pueda beneficiar a toda la humanidad. El

informe de la ONU articula fuertemente la segunda visión.

3.6 Sostenibilidad Ambiental: De preocupación marginal a valor central

Tanto "Governing AI for Humanity" como "Antiqua et Nova" elevan sostenibilidad ambiental de preocupación secundaria a valor ético central:

"Antiqua et Nova" vincula la ética de la IA con la enseñanza católica sobre ecología integral, argumentando que el impacto ambiental de IA debe gestionarse cuidadosamente para prevenir daños al planeta. El documento señala específicamente preocupaciones sobre consumo de agua y uso de energía por el hardware de IA.

Tanto "Governing AI for Humanity" como "Antiqua et Nova" elevan sostenibilidad ambiental de preocupación secundaria a valor ético.

"Governing AI for Humanity" conecta sostenibilidad de IA con los Objetivos de Desarrollo Sostenible de la ONU, particularmente con la acción climática. El informe reconoce que el entrenamiento de modelos grandes de IA consume recursos energéticos masivos con considerable huella de carbono.

A la luz, pues, de estos documentos cabe preguntarse ¿Los modelos cada vez más grandes son éticamente justificables dado su impacto climático? Aunque todos los países coinciden en promo-

A la luz, pues, de estos documentos cabe preguntarse ¿Los modelos cada vez más grandes son éticamente justificables dado su impacto climático?

Antiqua et Nova introduce el peligro de antropomorfizar la IA o tratarla como si fuera persona.

ver como valor o principio que los sistemas de IA beneficien el medio ambiente o sean sostenibles, ninguno concreta estándares ni exige evaluaciones ni impone límites por impacto climático... dejando esta pregunta sin respuesta todavía.

3.7 El Problema de la *Antropomorfización*: Nuevas dimensiones éticas

Antiqua et Nova introduce una dimensión ética que está subdesarrollada en la mayoría de marcos nacionales: el peligro de antropomorfizar la IA o tratarla como si fuera una persona. El documento argumenta que representar IA como persona es una "violación ética grave" que puede erosionar la confianza social y explotar vulnerabilidades humanas.

Esta preocupación nos lleva hacia cuestiones filosóficas profundas sobre categorías ontológicas, la definición de la conciencia y de estatus morales, que se ahondarán a medida que la IA se integre con neurotecnología para conectarse a cerebros de seres humanos para un uso personal, como ahora se utilizan los móviles. Algunos autores predicen la denominada "*singularidad*" (Interfaces cerebro-computadora) como el futuro del ser humano a medio o largo plazo¹.

1 "The Singularity Is Nearer: When We Merge with AI", Ray Kurzweil, junio 2024, Editorial Viking

Esta capacidad mental aumentada, usando a la vez un soporte biológico y tecnológico para pensar, sentir y trabajar ¿implicará un cambio de nuestra comprensión de lo que significa el ser *humano*? ¿Podríamos comenzar a valorar *relaciones* genuinas de humanos con su propio sistema de IA? ¿Cuáles son consecuencias éticas y psicológicas respecto de la *intimidad* personal o incluso respecto de una posible futura *conciencia* artificial?

Esta es un área de la ética donde los enfoques religiosos y filosóficos seguro influirán en los marcos nacionales que, por ahora, ni han previsto ni regulado esta singularidad.

4. CONCLUSIONES PARA UNA ÉTICA GLOBAL

4.1 Paisaje ético actual: voces corales similares con tonos diferentes

La publicación de la Convención Marco sobre IA, del Consejo de Europa en 2024, del Informe "Governing AI for Humanity" por la ONU y de "Antiqua et Nova" en 2025, transforma el paisaje ético global para los sistemas de IA. Ya no vemos cómo diferentes naciones regulan autónomamente la IA, sino cómo múltiples voces (gubernamentales, intergubernamentales, religiosas, e incluso de la sociedad civil) están contribuyendo a una conversación global emergente sobre qué significa gobernar la IA éticamente.

Esta conversación revela tanto convergencias notables como diversidades irreducibles:

Hay convergencias notables en que la IA debe (1) respetar la dignidad humana y los derechos fundamentales; (2) ser transparente y explicable en contextos apropiados; (3) evitar discriminaciones injustas; (4) incluir supervisión humana; (5) ser segura y robusta; (6) proteger la privacidad; y (7) estar sujeta a reclamaciones cuando cause daño.

La publicación de la Convención Marco sobre IA, del Consejo de Europa en 2024, del Informe "Governing AI for Humanity" por la ONU y de "Antiqua et Nova" en 2025, transforma el paisaje ético global para los sistemas de IA.

Hay convergencias notables en que la IA debe respetar la dignidad humana y los derechos fundamentales; ser transparente y explicable en contextos apropiados; evitar discriminaciones injustas; incluir supervisión humana; ser segura y robusta; proteger la privacidad; estar sujeta a reclamaciones cuando cause daño.

Sin embargo, la materialización de estos principios varía dramáticamente por las diferencias profundas en cómo se interpretan e implementan estos principios, reflejando diferentes tradiciones filosóficas, sistemas políticos, contextos culturales y prioridades de desarrollo. Diferentes sociedades tienen diferentes concepciones de equilibrios apropiados entre valores que ocasionalmente están en tensión: autonomía individual vs. armonía colectiva, normativa vinculante vs. programática, derechos negativos vs. positivos, desarrollo económico vs. protección social.

Esperar una convergencia completa sobre cómo equilibrar estos valores sería ingenuo e incluso éticamente cuestionable, por imponer una visión hegemónica sobre la diversidad legítima de tradiciones éticas o culturales.

Destacan por su relevancia dos divergencias conceptuales en las regulaciones estatales:

a) Individualismo vs. colectivismo vs. inclusividad social:

➤ Los marcos occidentales tienden a enfatizar los *derechos individuales* y la *autonomía*. Dentro de éstos, **Europa** parte de un protagonismo del Estado, que garantiza la protección de los derechos individuales. En cambio, Es-

tados Unidos da el protagonismo al individuo para defender sus derechos.

- El enfoque **asiático y africano** prioriza la *armonía social y bienestar colectivo*. En el extremo, el marco chino explícitamente prioriza estabilidad social y seguridad nacional con mecanismos que difieren significativamente del enfoque occidental. Esto no significa que el modelo chino carezca de consideraciones éticas, sino que opera dentro de un marco filosófico diferente sobre la relación entre individuo, sociedad y Estado.
 - En **Iberoamérica**, Perú y Brasil ejemplifican un enfoque que integra gobernanza ética de IA con objetivos más amplios de desarrollo económico inclusivo, reducción de desigualdad y justicia social. Estos marcos reflejan realidades de países en desarrollo donde los principales desafíos éticos no son solo cómo prevenir que la IA cause daño, sino cómo garantizar que la IA contribuya a reducir desigualdades estructurales existentes.
- b) Normas vinculantes vs. programáticas o voluntarias (autorregulación):**

- La Unión Europea favorece la regulación legal coercitiva para proteger los derechos, pues ve la dignidad humana y los derechos humanos

como límites absolutos al poder tecnológico. Este enfoque refleja una filosofía ética deontológica donde ciertos usos de IA son simplemente prohibidos independientemente de sus supuestos beneficios, porque violan derechos fundamentales. No obstante, hemos visto como la propia Comisión Europea ha propuesto en noviembre de 2025 posponer la aplicación de los requisitos a los sistemas de “alto riesgo” a través del Digital Omnibus para evitar un posible impacto en el ecosistema europeo de la IA. Igualmente, ha propuesto suprimir obligaciones alfabetización en materia de IA a las empresas, imponiendo a los poderes públicos el fomento de dicha alfabetización.

- Los países de Oriente medio y Asia proporcionan principios y estándares voluntarios, confiando en que la industria los implemente voluntariamente. Este modelo refleja una cultura política subyacente: confianza en que las corporaciones se autorregulan cuando se les proporciona orientación clara y en que marcos flexibles permiten adaptarse mejor a la rápida innovación tecnológica.
- Canadá y Emiratos Árabes Unidos: Han promovido una regulación vinculante para imponer estándares mínimos o para evitar

riesgos de incidentes graves, junto con una orientación voluntaria y mejoras prácticas a proveedores en áreas donde la innovación rápida requiere flexibilidad. Esta vía parece que va a ser potenciada por la UE para favorecer la competitividad mediante la denominada “Digital Omnibus”.

4.2 Tensiones éticas fundamentales en la gobernanza de IA

La incorporación de perspectivas internacionales revela ciertas tensiones éticas, que no son meramente desafíos técnicos, sino dilemas que requieren una gobernanza continua:

a) Derechos universales vs. Derechos contextualizados

¿Son principios éticos universalmente aplicables o dependen del contexto?

La incorporación de perspectivas internacionales revela ciertas tensiones éticas que requieren una gobernanza continua.

- **El principio del respeto de los derechos humanos** afirma la universalidad de ciertos derechos fundamentales.
- Perspectivas **culturales y postcoloniales** argumentan que la ética debe ser sensible al contexto local.

➤ **Antiqua et Nova** articula verdades universales, pero admite su aplicación prudente en contextos diversos.

La tensión no puede resolverse eligiendo uno u otro. Requiere compromiso continuo con mantener tanto compromisos universales como sensibilidad contextual lo que algunos filósofos morales² llaman "universalismo contextual".

b) Innovación versus Precaución

¿Cuántos riesgos en las vidas de los ciudadanos o vulneraciones de derechos son aceptables en búsqueda de beneficios de la IA?

- **Europa** adopta un enfoque precautorio, con entornos de prueba controlados.
- **Estados Unidos, Israel, Corea y Singapur** favorecen la experimentación industrial y la competitividad.
- **Governing AI for Humanity** enfatiza la necesidad de un equilibrio innovación-protección.
- **Antiqua et Nova** llama a la prudencia y sabiduría en la adopción de IA.

2 Rainer Forst, "Justificación y crítica. Perspectivas de una teoría crítica de la política." (Madrid: Katz, 2014). y Mark Hunyadi, "La aproximación contextualista de la moral", Université catholique de Louvain-La-Neuve, Revista Stultifera de Humanidades y Ciencias Sociales, Vol. 5, N°. 1, 2022.

Esta tensión no puede resolverse mediante una fórmula simple. Requiere una ponderación profunda entre los potenciales beneficios y la gravedad de riesgos potenciales, así como quién se beneficia y quién soporta tales riesgos, o si los potenciales daños son reversibles.

c) Equidad versus "Neutralidad Ideológica"

Una nueva tensión ética emerge en el contexto estadounidense, aunque resuena globalmente:

- **Colorado y California** tratan la equidad, no discriminación y la justicia como imperativos éticos fundamentales que deben integrarse en el diseño y evaluación de los sistemas de IA.
- El **Plan de Acción de IA de Trump** considera estas preocupaciones como "sesgo ideológico" y ordena que los sistemas federales sean "ideológicamente neutrales", específicamente rechazando marcos de "*diversidad, equidad e inclusión*".
- **Marcos europeos**, la ONU, UNESCO, el Consejo de Europa y Vaticano integran equidad y no discriminación como derechos fundamentales a proteger en la IA.

- **Marcos asiáticos** tienden a enfatizar la armonía social y el equilibrio sobre la equidad individualizada.

Esta tensión refleja un desacuerdo filosófico profundo: ¿Puede IA ser verdaderamente "*neutral*" cuando se despliega sin fronteras, en sociedades diferentes y en países diversos, con brechas y desigualdades estructurales?

La propia caracterización que hace el Plan de Trump de considerar la *diversidad, equidad e inclusión* como "ideológicas", mientras se afirma que su ausencia es "neutral", es en sí misma una posición ideológica. Esta tensión se intensificará a medida que diferentes países tomen posiciones sobre qué constituye desarrollo "*ético*" vs. "*sesgado*" de la IA.

4.3 Cuestiones éticas emergentes: agenda futura

Este análisis permite identificar 4 dimensiones éticas que requerirán mayor atención en futuros desarrollos de gobernanza.

Este análisis permite identificar 4 dimensiones éticas que requerirán mayor atención en futuros desarrollos de gobernanza:

a) La epistemología y la verdad

Cómo la IA afecta y afectará aún más en el futuro nuestra capacidad de conocer la verdad. Ello

es una preocupación plasmada por la mayoría de los Estados, que exigen etiquetar ultrasuplantaciones o niveles diferentes de transparencia:

- Debe alcanzarse la explicabilidad de las decisiones en sistemas basados en redes neuronales, en las que a día de hoy no es posible conocer los motivos que llevan a la IA a tomar determinadas decisiones.
- Las alucinaciones (errores) de la IA en sus resultados vulneran la confiabilidad.
- Los *Deepfakes* y el contenido sintético erosionan también la confiabilidad.
- La IA generativa puede crear "océanos de desinformación".

b) La Neurotecnología y los interfaces Cerebro-IA

A medida que la IA se integre con la neurotecnología, y dada la tendencia humana a antropomorfizar la IA (que se agudizará en soportes robóticos), se plantearán cuestiones éticas novedosas:

- Riesgos de manipulación de colectivos vulnerables (menores, ancianos).

- Cómo preservar la privacidad y la libertad mental individual.
- Riesgo de desigualdades en el acceso a tecnologías médicas.
- Discernir la frontera entre ser *humano vs. máquina*.
- Dilema sobre la posible existencia de “*conciencias*” en soporte robótico.

c) La IA y el cambio de paradigma del trabajo para colectivos enteros.

El impacto de IA en grandes sectores laborales requiere una reflexión ética profunda:

- La recapacitación y el sustento económico de colectivos de trabajadores (v.gr. transportistas, vendedores, restauración, operarios, etc.)
- La redefinición de la contribución productiva y la contribución impositiva.
- La transformación del significado de *trabajo*: ¿también en régimen de subsidios?

d) La IA y su beneficio a la humanidad

Más allá de prevenir daños, ¿cómo puede la IA contribuir al beneficio de la humanidad? ¿puede la IA apoyar el desarrollo de capacidades humanas en lugar de suplantarlas?

Debe promoverse siempre una visión positiva del futuro con la IA, no solo defensiva, pues se impone una evolución tecnológica imparable que será ciertamente positiva en muchos ámbitos, como la ciencia, la medicina, el arte, la educación, etc.

4.4 De los principios a la puesta en práctica

Para hacer que la ética de IA sea práctica, se pueden hacer propuestas y llamadas específicas a la acción para diferentes actores en la gobernanza de la IA, como las Organizaciones internacionales, los Gobiernos nacionales, los investigadores académicos y las propias compañías tecnológicas.

Para hacer que la ética de IA sea práctica, se pueden hacer propuestas y llamadas específicas a la acción para diferentes actores en la gobernanza de la IA.

Sin embargo, dado el enfoque del presente artículo, procede aterrizar los conceptos éticos en concretas propuestas para Instituciones religiosas, para la Sociedad civil y para los ciudadanos:

4.4.1 Para Instituciones religiosas:

Se pueden desglosar las siguientes acciones:

a) Continuar contribuyendo al discurso ético global:

- Articular propuestas sobre la IA desde enfoques espirituales y sociales.
- Ofrecer una crítica al reduccionismo tecnológico y al materialismo.
- Enfatizar el valor de la vida humana más allá de eficiencia y optimización.
- Proporcionar la visión del beneficio humano que puede guiar el desarrollo de IA.

b) Educar comunidades de fe:

- Integrar la reflexión sobre la IA en la educación religiosa.
- Ayudar a comunidades a navegar los desafíos éticos de IA en sus vidas.
- Alfabetizar y capacitar a religiosos para poder abordar las cuestiones de IA.

- Crear espacios para diálogo permanente entre tecnología y espiritualidad.

c) Defensa de la justicia y dignidad:

- Alzar la voz contra usos de la IA que violen la dignidad humana y los derechos fundamentales.
- Proteger a los vulnerables en la gobernanza de la IA.
- Presionar a las grandes tecnológicas por la distribución equitativa de los beneficios de IA.
- Mantener la justicia global y la solidaridad como horizontes irrenunciables.

d) Promover la integración reflexiva de la tecnología:

- Resistirse a una adopción acrítica de las tecnologías.
- Practicar discernimiento sobre cuándo y cómo usar la IA, sin socavar los valores espirituales.
- Mantener espacios para contemplación, silencio y reflexión en la era digital.

4.4.2 Para la Sociedad civil y ONGs

a) Supervisar la implementación y aplicación:

- Rastrear si las regulaciones se están implementando efectivamente respetando los Derechos fundamentales.
- Publicitar las brechas y debilidades en marcos regulatorios que vulneren derechos fundamentales.
- Documentar casos donde sistemas de IA discriminan o causan daño a colectivos o sectores sociales.
- Responsabilizar a los gobiernos y/o proveedores responsables de tales sistemas.

b) Amplificar voces marginalizadas:

- Garantizar que las comunidades afectadas tengan voz en la gobernanza de la IA.
- Documentar los impactos de la IA en grupos vulnerables.
- Traducir el lenguaje técnico a términos accesibles al público o a colectivos.

- Crear canales para la participación de colectivos normalmente excluidos.

c) Abogar por cambios de política:

- Presionar por una regulación más garantista en los países donde sea necesaria.
- Oponerse a usos dañinos de la IA, incluso promoviendo el ejercicio de acciones legales.
- Promover una mayor distribución equitativa de los beneficios de la IA.
- Mantener un enfoque de justicia y derechos humanos en procesos participativos.

d) Proporcionar educación pública:

- Aumentar la alfabetización sobre la IA entre el público general.
- Desmitificar la IA y hacerla comprensible a todos los públicos y colectivos.
- Empoderar a personas para ejercer sus derechos ante desarrollos dañinos de IA.
- Crear conciencia social sobre las dimensiones éticas de la IA.

4.4.3 Para los ciudadanos individuales

a) Informarse sobre IA:

- Buscar comprender los elementos básicos de cómo funciona la IA.
- Aprender sobre los diferentes derechos que regulan la IA.
- Ser consciente de cuándo se está usando IA en interacciones contigo mismo.
- Desarrollar una alfabetización crítica sobre la IA desde la perspectiva de los derechos humanos.

b) Ejercer los derechos existentes:

- Usar cuando existan los derechos de indemnización de los daños causados.
- Denunciar a los sistemas de IA que causen daño.
- Participar en consultas públicas sobre políticas de IA.
- Responsabilizar a las organizaciones que usan IA que cause daños.

c) Practicar un discernimiento personal sobre la tecnología:

- Considerar los impactos en uno mismo en tiempo y formación intelectual.
- Reflexionar sobre cuándo y cómo usar tecnologías de IA.
- Mantener espacios libres de tecnología para reflexionar y participar de la naturaleza.

d) Participar en el diálogo público:

- Contribuir a las conversaciones sobre futuro de la IA.
- Compartir las experiencias positivas y negativas con sistemas de IA.
- Apoyar políticas que promuevan una IA responsable, sostenible y confiable.
- Publicitar en medios sociales las preocupaciones sobre desarrollos dañinos.

4.5 Evidencias que emergen del examen realizado

Al concluir este análisis, emergen varias evidencias esenciales para construir una ética global en la IA:

Primera evidencia:

La Ética de la IA es fundamentalmente sobre valores humanos

En su esencia, los debates sobre gobernanza de IA no son *técnicos* sino normativos y sobre *valores*: qué tipo de sociedad queremos, qué sociedad futura queremos crear, qué aspectos de la humanidad queremos preservar y cuáles transformar. Ya sea en lenguaje de derechos humanos, dignidad, justicia o bien común, todos los marcos examinados hacen fundamentalmente afirmaciones sobre lo que importa y por qué importa.

Segunda evidencia:

No hay una “Solución única”.

La diversidad de enfoques no es un fracaso de coordinación sino reflejo de diferencias legítimas en valores, contextos, culturas, filosofías y prioridades. El enfoque de derechos fundamentales de Europa, el enfoque individualista de Colorado, el énfasis de China en valores socialistas, los marcos voluntarios de Singapur, la perspectiva teológica

del Vaticano, la llamada de la ONU a justicia global y los Derechos humanos... Cada uno ofrece aspectos valiosos y ninguno tiene un monopolio sobre la verdad ética.

Tercera evidencia: Convergencia es posible sin uniformidad

A pesar de las diferencias, hay notables convergencias sobre principios fundamentales. Prácticamente todos afirman que la IA debe respetar la dignidad humana, evitar la discriminación, y mencionan la transparencia, la supervisión humana, la privacidad, la seguridad y servir al bien común. Esta convergencia es una base sólida para la cooperación internacional mientras se respeta diversidad legítima al implementar la IA.

Cuarta evidencia: El Tiempo importa en la toma de decisiones sobre IA

La IA está evolucionando extraordinariamente rápido. Decisiones que tomamos ahora sobre cómo gobernarla moldearán trayectorias tecnológicas por décadas que influirán en la humanidad. No podemos esperar a la certeza perfecta o al consenso completo. Debemos actuar con la información actual, reconociendo que necesitaremos adaptación continua.

Quinta evidencia: La Inclusión no es opcional

Como "Governing AI for Humanity" afirma, 118 países no han tenido voz en iniciativas significativas de gobernanza de IA. Esta exclusión es éticamente insostenible. Si la IA va a afectar a toda la humanidad de forma radical, entonces toda la humanidad debería tener voz en su gobernanza. La inclusión no es gesto simbólico, sino un imperativo moral y un requisito práctico para una gobernanza de la IA legítima y efectiva.

Sexta evidencia: La Responsabilidad es compartida

Ningún actor único (ni gobiernos, ni empresas, ni organismos internacionales, ni la sociedad civil) puede abordar los desafíos éticos de la IA solo. Una gobernanza efectiva requiere colaboración *multi-stakeholder*, de todas las partes interesadas, donde los diferentes actores asuman diferentes roles proactivamente y responsabilidades, trabajando hacia objetivos compartidos.

Séptima evidencia: La Vigilancia debe ser continua

La ética de la IA no es un problema que puede resolverse de una vez. Es un proyecto continuo

que requiere vigilancia constante, reflexión continua, adaptación y compromiso sostenido. Las capacidades de la IA continuarán evolucionando, surgirán nuevos desafíos éticos, y nuestro entendimiento de sus implicaciones se profundizará con la experiencia.

Octava evidencia: La Esperanza está Justificada

A pesar de desafíos y riesgos reales, hay muchas razones para la esperanza. El hecho de que tantas naciones, organizaciones internacionales, empresas e investigadores ya estén comprometidos con la ética de la IA es alentador. Desde una AI Act de la UE holística hasta el Diálogo Global de la ONU, desde Antiqua et Nova del Vaticano hasta las estrategias nacionales de países en desarrollo, hay un compromiso generalizado de gobernar la IA de manera que respeten dignidad humana y promuevan bien común.

Como Antiqua et Nova expresó: "*Si dirigimos la IA hacia el bien común, puede ayudar a reducir la pobreza, expandir el acceso a la educación, mejorar diagnósticos de salud, promover la sostenibilidad ambiental e incluso fomentar una comprensión más profunda de nuestras tradiciones culturales y espirituales.*"

La IA no es un destino inevitable sino una herramienta cuyo valor depende de cómo la usamos. Las regulaciones y marcos examinados representan nuestros mejores esfuerzos colectivos actuales para garantizar que usamos esta herramienta poderosa sabiamente, éticamente y al servicio del beneficio humano.

4.6 La pregunta que permanece en medio de esta nube, aún difusa

Al final, todos los marcos éticos examinados están tratando de responder una pregunta básica: *¿Cómo podemos garantizar que la inteligencia artificial amplifique lo mejor de la humanidad en lugar de lo peor y la beneficie?*

Esta pregunta no tiene respuesta final porque está atada a preguntas futuras sobre qué significará ser “humano” o “conciencia”, o qué constituirá una sociedad armónica, qué obligaciones tenemos ya con las generaciones futuras, y cómo equilibraremos el valor de la competitividad en la economía, en un mundo de gran complejidad social.

No obstante, no teniendo respuestas finales, tenemos una hoja de ruta. Los marcos examinados (desde tratados internacionales hasta marcos nacionales, desde las perspectivas religiosas hasta

los estándares técnicos) nos ofrecen un conocimiento colectivo que puede guiarnos al indicarnos que debemos:

- Centrarnos en dignidad humana
- Respetar los derechos fundamentales
- Buscar la justicia e inclusión
- Mantener una transparencia y rendición de cuentas
- Preservar la supervisión humana
- Proteger a colectivos vulnerables
- Distribuir los beneficios de la IA equitativamente, sean médicos, científicos, etc.
- Actuar con precaución y evaluaciones que permitan sacar evidencias.
- Establecer estructuras de monitorización y adaptables a las innovaciones.
- Trabajar juntos, conociendo y aceptando nuestras legítimas diferencias éticas, sociales, filosóficas, religiosas y culturales.

Estos principios no resuelven cada dilema ético que enfrentamos, pero proporcionan una brújula que puede orientarnos al navegar dentro de esta nube difusa aún: la *era de la IA*.

4.7 Necesidad de la acción de todos: la llamada de la ONU en su informe

Este análisis ha examinado cómo diferentes naciones y organismos internacionales abordan la ética en la regulación de la IA. Hemos visto la articulación de principios claros, con creatividad en el diseño institucional, y un compromiso para traducir los valores éticos en acciones, e incluso protecciones, concretas.

Pero este análisis solo será útil, como concluye la ONU en su Informe final “*Governing AI for humanity*”, si realmente promueve la acción y la proactividad de todas las partes interesadas, incluyendo gobiernos, industria, la academia y a nosotros como ciudadanos.

El desafío es trasladar el conocimiento comunitario al mundo real: implementando las regulaciones de forma efectiva, desarrollando sistemas de IA de forma responsable, distribuyendo beneficios equitativamente en el planeta, protegiendo los derechos fundamentales y estando abiertos al aprendizaje y la adaptación. Cada

uno en sus capacidades personales, económicas y profesionales.

El futuro de la IA no está escrito ni será determinado por una lógica tecnológica inexorable. Será escrito cada día por las opciones que tomemos, como políticos, tecnólogos, académicos, divulgadores, activistas y ciudadanos individuales.

Los marcos examinados proporcionan fundamentos éticos para ello. Es trabajo de todos el contribuir en la nueva *era de la IA* al desarrollo y uso responsable, sostenible y confiable de la inteligencia artificial.

La tarea es urgente, pues los riesgos son reales. Pero también son oportunidades de crecer más y mucho mejor. Con conocimiento, compromiso y cooperación podemos gobernar la IA de manera que respete la ética, mientras alcanzamos las inmensas posibilidades transformadoras de esta nueva era que se abre en la historia de la humanidad.

DOCUMENTACIÓN

UNION EUROPEA

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689

<https://artificialintelligenceact.eu/the-act/>

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

https://ec.europa.eu/commission/presscorner/detail/es/ip_25_2718

ESTADOS UNIDOS

COLORADO

<https://leg.colorado.gov/bills/sb24-205>

<https://www.naag.org/attorney-general-journal/a-deep-dive-into-colorados-artificial-intelligence-act/>

<https://cdt.org/insights/faq-on-colorados-consumer-artificial-intelligence-act-sb-24-205/>

<https://www.akingump.com/en/insights/ai-law-and-regulation-tracker/colorado-postpones-implementation-of-colorado-ai-act-sb-24-205>

<https://fpf.org/wp-content/uploads/2024/05/FPF-FINAL-CO-SB-205-Two-Pager-.pdf>

<https://www.wiley.law/alert-Colorado-Enacts-Landmark-AI-Legislation>

<https://www.akingump.com/en/insights/ai-law-and-regulation-tracker/colorado-postpones-implementation-of-colorado-ai-act-sb-24-205>

CALIFORNIA

<https://www.pillsburylaw.com/en/news-and-insights/california-ai-laws.html>

<https://www.crowell.com/en/insights/client-alerts/californias-landmark-ai-law-demands-transparency-from-leading-ai-developers#:~:text=California%20is%20the%20first%20state,the%20first%20time%2C%20to%20release>

<https://www.diarioconstitucional.cl/2025/10/03/california-ee-uu-promulga-ley-que-regula-tecnologias-de-inteligencia-artificial/>

<https://www.gov.ca.gov/2025/09/29/governor-newsom-signs-sb-53-advancing-californias-world-leading-artificial-intelligence-industry/>

AMERICA'S AI ACTION PLAN

<https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf>

<https://www.ai.gov/action-plan>

CHINA

<https://www.chinalawvision.com/2025/01/digital-economy-ai/ai-ethics-overview-china/>

<https://www.chinalawtranslate.com/en/generative-ai-interim/>

<https://www.chinalawtranslate.com/en/recommendation-algorithms/>

https://cset.georgetown.edu/wp-content/uploads/t0592_china_ai_law_draft_EN.pdf

https://www.cac.gov.cn/2025-03/14/c_1743654684782215.htm

<https://www.thelegaletters.com/el-deber-de-transparencia-en-ia-europa-vs-china/>

CANADA

<https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act>

<https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document>

<https://montrealethics.ai/the-death-of-canadas-artificial-intelligence-and-data-act-what-happened-and-whats-next-for-ai-regulation-in-canada/>

JAPAN

<https://theofficer.es/japon-legisla-la-ia-por-primera-vez-asi-busca-evitar-abusos-y-delitos-digitales/>

<https://practiceguides.chambers.com/practice-guides/artificial-intelligence-2025/japan>

<https://aisigil.com/es/regulacion-de-la-ia-en-japon-nuevas-leyes-y-desafios/>

<https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-japan>

<https://www.escudodigital.com/tecnologia/japon-convierte-crisis-demografica-en-oportunidad-tecnologia-con-alma-cuidar-mayores.html>

<https://www.swissinfo.ch/spa/jap%C3%A9n-aprueba-su-primera-legislaci%C3%B3n-en-materia-de-desarrollo-y-uso-de-ia/89422685>

<https://www.protecto.ai/blog/ai-regulations-and-governance-april-2025-update>

SINGAPUR

<https://www.diligent.com/resources/blog/Singapore-AI-regulation>

<https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2024/public-consult-model-ai-governance-framework-genai>

<https://www.pdpc.gov.sg/help-and-resources/2020/01/model-ai-governance-framework>

<https://aiverifyfoundation.sg/>

COREA DEL SUR

<https://spanish.korea.net/NewsFocus/Policies/view?articleId=264075>

<https://aibasicact.kr/>

<https://iaparabogados.com/corea-del-sur-y-la-ley-basica-de-ia-un-paso-adelante-en-la-regulacion-de-la-inteligencia-artificial/>

BRASIL

Brazil AI Act | Artificial Intelligence Act artificialintelligenceact.com

<https://retina-der.uniovi.es/index.php/2025/11/07/perspectivas-de-regulacion-de-la-ia-en-brasil-y-su-impacto-en-los-derechos-fundamentales/>

<https://institutoautor.org/el-senado-de-brasil-aprueba-el-proyecto-de-ley-2338-2023-sobre-el-uso-de-la-inteligencia-artificial/>

The Artificial Intelligence Legislation in Brazil: Technical Analysis of the Text to Be Voted on in the Federal Senate Plenary - Data Privacy Brasil Research dataprivacybr.org

Artificial Intelligence 2025 - Brazil | Global Practice Guides | Chambers and Partners chambers.com

Brazil AI Act | Artificial Intelligence Act artificialintelligenceact.com

AI Watch: Global regulatory tracker - Brazil | White & Case LLP whitecase.com

PERU

<https://www.gob.pe/institucion/congreso-de-la-republica/normas-legales/4565760-31814>

<https://www.mofa.com/artificial-intelligence/peru>

<https://digitalpolicyalert.org/event/19625-drafted-regulation-of-law-promoting-the-use-of-artificial-intelligence-for-the-economic-and-social-development-of-peru-no-31814-including-data-protection-regulation>

<https://digital.nemko.com/regulations/ai-regulation-in-peru>

https://www.garrigues.com/es_ES/noticia/peru-aprueba-reglamento-ley-promueve-uso-inteligencia-artificial-favor-desarrollo-economico

<https://www.cuatrecasas.com/es/latam/tecnologia-medios-digitales/art/que-establece-el-reglamento-de-la-ley-de-inteligencia-artificial>

<https://www.dentons.com/es/insights/articles/2025/september/19/peru-aprueba-el-reglamento-que-regula-el-uso-de-la-inteligencia-artificial>

KENIA

<https://www.unesco.org/ethics-ai/en/kenya>

<https://unesdoc.unesco.org/ark:/48223/pf0000392693>

<https://ict.go.ke/sites/default/files/2025-03/Kenya%20AI%20Strategy%202025%20-%202030.pdf>

<https://imbizomag.co.za/2025/11/17/ubuntu-ai-makes-debut-in-africa/>

<https://itlawco.com/kenya-national-ai-strategy-2025-2030/>

<https://www.techpolicy.press/redefining-ai-for-africa-the-role-of-artificial-intelligence-in-kenyas-grassroots-movement/>

AUSTRALIA

<https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles/australias-ai-ethics-principles>

<https://www.finance.gov.au/government/public-data/data-and-digital-ministers-meeting/national-framework-assurance-artificial-intelligence-government/implementing-australias-ai-ethics-principles-government>

<https://www.digital.nsw.gov.au/policy/artificial-intelligence/artificial-intelligence-ethics-policy/mandatory-ethical-principles>

https://www.csiro.au/-/media/D61/NAIC/Gradient-Report/23-00122-DATA61_REPORT_NAIC-ResponsibleAITools_WEB_230620.pdf

EUA

<https://uaelegislation.gov.ae/en/policy/details/uae-s-international-stance-on-artificial-intelligence-policy>

<https://uaelegislation.gov.ae/en/policy/details/uae-s-international-stance-on-artificial-intelligence-policy#:~:text=These%20principles%20are%3A%20fairness%2C%20ensuring, AI%20systems%20being%20-reliable%20and>

<https://alkabban.com/news/ai-regulation-uae-2025-business-guide/>

<https://www.lw.com/en/insights/ai-in-the-uae-understanding-the-regulatory-landscape-and-key-authorities>

<https://www.globallegalinsights.com/practice-areas/ai-machine-learning-and-big-data-laws-and-regulations/uae/>

<https://www.pinsentmasons.com/out-law/news/uae-launches-ai-charter-with-focus-on-human-wellbeing>

<https://sec.gov/https://securiti.ai/uae-charter-for-the-development-and-use-of-ai/uriti.ai/uae-charter-for-the-development-and-use-of-ai/>

ISRAEL

https://www.gov.il/BlobFolder/policy/ai_2023/en/Israels%20AI%20Policy%202023.pdf

<https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-israel>

<https://dig.watch/resource/israels-policy-on-artificial-intelligence-regulation-and-ethics>

<https://blog.ai-laws.org/israeli-ai-regulation-and-policy-white-paper-a-first-glance/>

<https://www.lexology.com/library/detail.aspx?g=23b88410-c417-4261-a04f-877e1ee772cb>

PRINCIPIOS DE IA, OCDE

<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>

<https://legalinstruments.oecd.org/api/download/?uri=/public/db5053b5-93e0-4cf5-a7cf-edce5ee6e893.pdf>

<https://www.cuatrecasas.com/es/spain/propiedad-intelectual/art/ocde-actualiza-principios-ia>

<https://www.onetrust.com/es/blog/approaching-the-oecd-framework-for-the-classification-of-ai-systems/>

RECOMENDACIÓN SOBRE ÉTICA DE IA, UNESCO

<https://www.unesco.org/es/articles/recomendacion-sobre-la-etica-de-la-inteligencia-artificial>

<https://www.unesco.org/es/artificial-intelligence/recommendation-ethics>

CONVENIO MARCO SOBRE IA, CONSEJO DE EUROPA

<https://rm.coe.int/1680afae3c>

<https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-81308>

<https://digital-strategy.ec.europa.eu/en/news/commission-signed-council-europe-framework-convention-artificial-intelligence-and-human-rights>

https://spain.representation.ec.europa.eu/noticias-eventos/noticias-0/la-comision-firma-el-convenio-marco-sobre-inteligencia-artificial-del-consejo-de-europa-2024-09-05_es

"GOVERNING AI FOR HUMANITY", ONU

https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf

<https://www.hsfkramer.com/notes/tmt/2024-posts/UN-releases-its-final-report-on--Governing-AI-for-Humanity->

"ANTIQUA ET NOVA", IGLESIA CATÓLICA

https://www.vatican.va/roman_curia/congregations/cfaith/documents/rc_ddf_doc_20250128_antiqua-et-nova_sp.html

<https://www.fpablovi.org/documentacion-huella-digital/2018-antiqua-et-nova-nuevo-documento-sobre-ia-el-riesgo-de-que-el-ser-humano-se-convierta-en-esclavo-de-su-propia-obra>

<https://www.rtve.es/noticias/20250128/vaticano-ia-oportunidad-advierte-manos-empresas-poderosas/16424917.shtml>

FCO. JAVIER TORRES GELLA

Fco. Javier Torres Gella es Abogado del Estado desde el año 2002 y Académico correspondiente de la Real Academia de Jurisprudencia y Legislación. Ha ocupado destinos en Tribunales de todas las jurisdicciones y en arbitrajes, comerciales y de inversiones. De igual forma, ha ocupado destinos en diferentes Ministerios y Organismos Públicos y ha representado al Reino de España en la negociación de Reglamentos y Directivas comunitarios, en el Consejo Europeo.

Actualmente ostenta el cargo de Secretario General de la *Agencia Española de Supervisión de Inteligencia Artificial*; organismo encargado de supervisar y promover el uso ético, responsable y confiable de la IA en España. Simultáneamente participa en diferentes grupos de trabajo del *AI Office* y en el proyecto *TAIEX Workshop on EU approach to artificial intelligence*, para asistir a países externos a la UE en la implementación de una regulación nacional en materia de IA.

Otros números de Cuadernos de ética en clave cotidiana:

- Nº25: Ética y sanidad: Un proyecto de talleres éticos en Hospitales. Begoña Husillos Tamarit.
- Nº24: Ética y menores: Protección integral del menor ante la sanidad y la justicia. Manuel Ruiz Martínes-Cañavete.
- Nº23: El educador cosmopolita global: pedagogías, valores y competencias profesionales. Luis Aranguren Gonzalo y César García-Rincón de Castro
- Nº22: Bioética y soporte nutricional en el final de la vida. Francisco Rivas García
- Nº21: Conversar, un arte a recuperar. Enrique Lluch Frechina y Luis Illueca Vicente
- Nº20: Sobre el género de la identidad. Aspectos fundamentales para comprender los debates en torno al sexo, el género y la identidad. Silvia Martínez Cano.
- Nº19: Una mirada a la pandemia desde Cáritas. Rafael Allepuz Capdevila y Jesús Pérez Mayo.
- Nº18: Ética en atención primaria. Nazaret Maldonado del Arco.
- Nº17: Cuidados paliativos. Jacinto Bátiz Cantera.
- Nº16: Ética social para niñas y niños. Guía práctica para padres y educadores. César García-Rincón de Castro.
- Nº15: Los confinados. Manual de resistencia (para salir fortalecidos tras el covid-19). F.J. de la Torre Díaz.
- Nº14: Vivir (des)conectado. Aspectos éticos del uso (y abuso) de los dispositivos móviles. Isidro Catela Marcos.
- Nº 13: Acompañar en la fase final de la vida. Natividad Jiménez Sánchez, Nazaret Maldonado del Arco y María Dolores Nieto Martín.

- Nº 12: Ética y Derechos Humanos. Gisela Giner Rommel.
- Nº 11: Paseos para hacer en compañía. José Luis Pareja Rivas.
- Nº 10: Internet y las redes sociales, aspectos éticos. Margarita Martín Martín.
- Nº 9: Ética y vida: La bioética. Rafael Junquera Estefani y Ana Mª Marcos del Cano.
- Nº 8: Los agentes de la Cooperación Internacional al Desarrollo. Una mirada ética. Patricia Rodríguez González.
- Nº 7: Videojuegos, gamificación y reflexiones éticas. Margarita Martín Martín y Luis Fernando Vilchez Martín.
- Nº 6: Función Social de la empresa: Una propuesta de evaluación ética. Elisa Marco Crespo y Enrique Lluch Frechina.
- Nº 5: Ética del Cuidado y Mayores. Los cuidados a las personas mayores desde un horizonte ético y en la búsqueda de la calidad de vida. Rosario Paniagua Fdez.
- Nº 4: Introducción a la Ética Familiar. Victor Chacón, CSsR.
- Nº 3: Interpelación ética de las mujeres que ejercen la prostitución. Mª Luisa del Pozo.
- Nº 2: Ética y Escuela. Juan José Medina Rodríguez y Mª Isabel Rodríguez Peralta.
- Nº 1: Bases éticas para la mejora de nuestra organización económica y política. Enrique Lluch Frechina y Rafael S. Hernández.
- Nº 0: El Don que transforma, una mirada moral desde el carisma redentorista. Carlos Sánchez de la Cruz.

Puedes descargarlos en la página web: <http://funderetica.org/cuadernos/>

éti

cuadernos de
ética
en clave cotidiana



Editorial Perpetuo Socorro
Covarrubias, 19. 28010 Madrid

C/ Félix Boix, 13 28036 Madrid
fundraising@funderetica.org
www.funderetica.org