

---

DOCUMENTO DE ANÁLISIS

# Ética en la inteligencia artificial

## Principios, marcos regulatorios y la distancia entre el deber ser y la práctica

*Un análisis de los marcos éticos de la UNESCO, la Unión Europea y la OCDE, y sus implicancias para instituciones educativas y organizaciones*

---

### Nick Perot

Abogado | Consultor en transformación digital | Docente  
www.nickperot.com    Marzo 2026

## 1. Una tecnología que avanza más rápido que las preguntas sobre ella

---

La inteligencia artificial ya no es un tema de laboratorio. Está en las aulas, en los procesos de investigación, en la gestión de organizaciones y en las decisiones cotidianas de millones de personas. Esa velocidad de adopción genera una tensión que vale la pena nombrar: las herramientas se implementan antes de que existan marcos claros para evaluar sus implicancias éticas.

Este documento analiza los principales marcos éticos internacionales aplicables a la IA, identifica los principios sobre los que existe consenso entre organismos multilaterales y examina la distancia que separa esos principios de su implementación efectiva en instituciones educativas y organizaciones. No aborda en profundidad el uso responsable de la IA, que será objeto de un documento específico posterior. El foco aquí está en la ética: en los valores, principios y marcos normativos que deberían orientar el desarrollo y la adopción de estas tecnologías.

## 2. Ética y uso responsable: una distinción necesaria

---

Los términos ética de la IA y uso responsable de la IA aparecen con frecuencia como sinónimos en la literatura y en las políticas institucionales. No lo son. La distinción importa porque orienta decisiones diferentes.

La ética de la IA refiere al marco de valores, principios y reflexiones normativas que definen qué está bien y qué está mal en el desarrollo, la implementación y el uso de sistemas de inteligencia artificial. La UNESCO la define como una reflexión normativa sistemática, basada en un marco integral, holístico, multicultural y evolutivo de valores, principios y acciones interdependientes. Es, en esencia, el territorio de las preguntas fundamentales: ¿qué derechos deben protegerse? ¿Qué daños son inaceptables? ¿Quién es responsable cuando un sistema automatizado toma una decisión que afecta a personas?

*UNESCO, Recomendación sobre la ética de la inteligencia artificial, París, 2021, párrafo 1.*

El uso responsable, en cambio, refiere a las prácticas concretas, los protocolos, las políticas institucionales y los hábitos operativos que traducen esos principios éticos en acciones cotidianas. Incluye cómo se documenta el uso de herramientas de IA, cómo se capacita a los equipos, cómo se transparenta el contenido generado por sistemas automatizados y cómo se miden los resultados.

La ética define el horizonte. El uso responsable construye el camino. Sin el primero, el segundo no tiene dirección. Sin el segundo, el primero no tiene efecto.

Este documento se concentra en el horizonte: los marcos éticos. El camino operativo del uso responsable se desarrollará en un documento posterior.

### 3. Los principios éticos: dónde hay consenso internacional

---

#### La Recomendación de la UNESCO (2021)

En noviembre de 2021, los Estados Miembros de la UNESCO adoptaron la Recomendación sobre la ética de la inteligencia artificial, el primer estándar normativo global en la materia. El documento fue adoptado por 193 países en la Conferencia General de ese año y es hoy aplicable a los 194 Estados Miembros de la organización. Establece cuatro valores fundamentales sobre los que se construye todo el marco: derechos humanos y dignidad humana; sociedades pacíficas, justas e interconectadas; diversidad e inclusión; y florecimiento del medio ambiente y los ecosistemas.

*UNESCO, Recomendación sobre la ética de la inteligencia artificial, adoptada el 23 de noviembre de 2021.  
Disponible en: [unesdoc.unesco.org/ark:/48223/pf0000381137\\_spa](https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa)*

Sobre esos cuatro valores, la Recomendación define diez principios operativos: proporcionalidad y no causar daño; seguridad y protección; equidad y no discriminación; sostenibilidad; privacidad y protección de datos; supervisión y determinación humanas; transparencia y explicabilidad; responsabilidad y rendición de cuentas; sensibilización y alfabetización; y gobernanza y colaboración adaptativas y multipartitas.

Lo que distingue a la Recomendación de la UNESCO de otros marcos es su alcance universal y su operatividad: incluye once áreas de acción política concretas que van desde la gobernanza de datos hasta la igualdad de género, pasando por educación, salud, cultura y medio ambiente.

#### La Ley de Inteligencia Artificial de la Unión Europea (2024)

La Unión Europea adoptó en 2024 la primera legislación integral que regula el uso de sistemas de IA en el mundo. El Reglamento (UE) 2024/1689, conocido como AI Act, entró en vigor el 1 de agosto de 2024 y será plenamente aplicable el 2 de agosto de 2026, con aplicación escalonada según el tipo de obligación.

*Unión Europea, Reglamento (UE) 2024/1689 (AI Act), Diario Oficial de la UE, 13 de junio de 2024.*

El AI Act adopta un enfoque basado en riesgo. En materiales introductorios, la propia Comisión Europea lo presenta en cuatro niveles: riesgo inaceptable (sistemas prohibidos), riesgo alto (regulado con obligaciones estrictas), riesgo limitado (obligaciones de transparencia) y riesgo mínimo (sin

regulación específica). A estos cuatro niveles se suman obligaciones diferenciadas para los modelos de propósito general (GPAI), que el Reglamento trata en un bloque específico.

### Categorías de riesgo del AI Act

Categoría	Sistemas alcanzados	Obligaciones
<b>Inaceptable</b>	Manipulación subliminal, puntuación social, categorización biométrica por características sensibles, reconocimiento de emociones en el trabajo y la educación.	Prohibidos. Aplicable desde el 2 de febrero de 2025.
<b>Alto</b>	Sistemas que impactan seguridad o derechos fundamentales: educación, empleo, infraestructura crítica, servicios esenciales, aplicación de la ley.	Gestión de riesgos, gobernanza de datos, documentación técnica, supervisión humana, registro en base de datos de la UE.
<b>Limitado</b>	Sistemas que interactúan directamente con usuarios: chatbots, generadores de contenido sintético (deepfakes), sistemas de reconocimiento de emociones permitidos.	Informar al usuario que interactúa con IA o que el contenido fue generado artificialmente.
<b>Mínimo</b>	Sistemas de uso normalizado: filtros de spam, videojuegos, herramientas de productividad sin impacto en derechos fundamentales.	Sin obligaciones específicas. Se sugieren códigos voluntarios.

*Fuente: Reglamento (UE) 2024/1689; Parlamento Europeo, "EU AI Act: first regulation on artificial intelligence," ficha informativa, 2024.*

Las obligaciones para modelos de propósito general (GPAI) se regulan aparte y no encajan directamente en la escala de cuatro niveles. Los proveedores de GPAI deben cumplir obligaciones de transparencia y documentación técnica. Los modelos que superan umbrales de riesgo sistémico enfrentan requisitos adicionales: evaluación del modelo, pruebas adversariales, reporte de incidentes graves y protecciones de ciberseguridad. Estas obligaciones son aplicables desde el 2 de agosto de 2025.

Un dato relevante: desde febrero de 2025, las prácticas de IA prohibidas y las obligaciones de alfabetización en IA ya son exigibles. La aplicación plena del Reglamento será en agosto de 2026, y los sistemas de alto riesgo integrados en productos regulados tienen plazo hasta agosto de 2027. La regulación no es futura. Ya está en vigencia parcial.

### La perspectiva de la OCDE y la definición sistémica

La OCDE adoptó en 2024 una definición actualizada de sistema de IA que enfatiza su naturaleza sistémica: un sistema basado en máquinas que, para objetivos explícitos o implícitos, infiere a partir de la información que recibe y genera resultados como predicciones, contenidos, recomendaciones

o decisiones que pueden influir en entornos físicos o virtuales. La definición incluye una referencia explícita a los niveles variables de autonomía y adaptabilidad tras el despliegue.

*OCDE, Explanatory memorandum on the updated OECD definition of an AI System, OECD Artificial Intelligence Papers, N.º 8, marzo de 2024.*

Esta definición importa porque desplaza el foco desde la herramienta individual hacia el sistema completo: los datos con los que se entrena, los algoritmos que lo componen, las decisiones que produce y el contexto en el que opera. La ética de la IA, en esta perspectiva, no puede evaluarse en abstracto. Requiere considerar el sistema en su totalidad, incluyendo las condiciones de acceso, recursos e infraestructura de quien lo adopta.

## 4. El marco ético en la educación superior

La integración de la IA en instituciones educativas plantea una dualidad que el diagnóstico de la DGTIC de la UNAM, publicado en octubre de 2025, describe con precisión: enriquecimiento de la docencia y la investigación por un lado, y desafíos éticos que exigen un manejo cuidadoso y reflexivo por el otro.

*Castañeda de León, L.M. et al., Uso y desarrollo ético de la Inteligencia Artificial en la Universidad: docencia e investigación, DGTIC-UNAM, octubre de 2025, pp. 34–35.*

El documento identifica cinco principios éticos sustantivos derivados de la UNESCO y aplicables a la docencia e investigación universitarias: beneficencia, no maleficencia, autonomía, justicia y transparencia. Estos cinco principios conforman la base sobre la que deberían construirse las políticas institucionales de IA en cualquier universidad.

### Los cinco principios éticos para la IA en educación superior

Principio	Aplicación en docencia e investigación
<b>Beneficencia</b>	Los sistemas de IA deben buscar el beneficio de la comunidad universitaria y la sociedad: aprendizaje significativo, desarrollo de competencias y solución de problemas relevantes.
<b>No maleficencia</b>	Evitar daños: sesgos de discriminación, manipulación, deshumanización, pérdida de control. Garantizar la seguridad y calidad de datos y algoritmos.
<b>Autonomía</b>	Respetar la libertad de decisión de docentes, investigadores y estudiantes. Informar sobre el uso y funcionamiento de la IA. Ofrecer alternativas.
<b>Justicia</b>	Promover equidad e inclusión educativa. Evitar generar o agravar brechas, desigualdades o discriminación por género, edad, origen, cultura o discapacidad.
<b>Transparencia</b>	Facilitar el entendimiento y la confianza. Honestidad sobre el origen, propósito y límites de la IA utilizada. Rendición de cuentas sobre decisiones.

*Fuente: UNESCO (2021); Castañeda de León et al. (2025), pp. 36–37.*

## El consenso multilateral: más amplio de lo esperado

Un hallazgo relevante del diagnóstico de la UNAM es que el consenso ético internacional sobre IA es más extenso de lo que suele reconocerse. El trabajo de Nguyen et al. (2023) mapea los principios éticos presentes en seis marcos multilaterales (UNESCO 2020, UNESCO 2021, Consenso de Beijing 2019, OCDE 2021, Comisión Europea 2019 y Parlamento Europeo 2021) e identifica convergencias en siete grandes áreas: gobierno y administración, transparencia y rendición de cuentas, sostenibilidad, privacidad, seguridad y protección, inclusividad, e IA centrada en el ser humano.

*Nguyen, A., Ngo, H.N., Hong, Y., Dang, B. y Nguyen, B.P.T., "Ethical principles for artificial intelligence in education," Education and Information Technologies, 28(4), 4221–4241, 2023.*

El principio de transparencia, por ejemplo, aparece en los seis marcos analizados. Lo mismo ocurre con la privacidad de datos, la inclusividad, la justicia y la centralidad del ser humano en los sistemas de IA. Esto no significa que la implementación sea uniforme, pero sí que el debate ético tiene una base común más sólida de lo que la fragmentación regulatoria podría sugerir.

En la misma dirección, Seldon, Lakhani y Luckin (2020) elaboraron un marco ético para la IA en educación con nueve objetivos y 33 criterios verificables, diseñado para ser aplicable a cualquier entidad educativa independientemente de su tamaño o ubicación geográfica. Ese enfoque de universalidad intenta responder al problema que Jin et al. (2024) plantean sobre las asimetrías entre el Norte y el Sur Globales en materia de IA educativa.

*Seldon, A., Lakhani, P. y Luckin, R., The Ethical Framework for AI in Education, The Institute for Ethical AI in Education, 2020.*

## 5. La distancia entre los principios y la práctica

---

Michael Sandel, profesor de la Universidad de Harvard, identifica tres áreas principales de preocupación ética que la IA plantea a la sociedad: privacidad y vigilancia, sesgo y discriminación, y el papel del juicio humano en las decisiones que más importan. Sobre esta última, Sandel plantea la pregunta más difícil: si las máquinas inteligentes pueden superar nuestro pensamiento, o si hay elementos del juicio humano que son indispensables para decidir sobre las cosas más importantes de la vida.

*Harvard Gazette, "Ethical concerns mount as AI takes bigger decision-making role," Harvard University, 2020. Declaraciones de Michael Sandel.*

Esa tensión entre regulación e innovación también aparece en el ámbito académico. En el Winter Workshop 2025 de la Harvard Data Science Initiative, dedicado a IA y modelos de lenguaje, uno de

los temas recurrentes fue la dificultad de equilibrar controles regulatorios con el ritmo de la innovación, y la sugerencia de que un punto de partida viable puede ser construir gobernanza sólida sobre los datos antes de regular los sistemas que los procesan.

*Harvard Data Science Initiative, "Human vs. AI — or is it?," resumen del HDSI Winter Workshop 2025, datascience.harvard.edu, marzo de 2025.*

## El caso de las universidades latinoamericanas

El diagnóstico de la UNAM aporta un dato que ilustra la magnitud de la brecha entre principios y práctica. Al revisar publicaciones de académicos e investigadores de la UNAM sobre la institucionalización de la ética en los sistemas de IA durante el período 2020–2025, los autores no identificaron propuestas institucionales sobre uso y desarrollo ético de la IA publicadas en revistas indexadas de alto impacto. La integración de la ética en los sistemas de IA de la UNAM, concluyen, apenas comienza a ser tratada.

*Castañeda de León et al. (2025), pp. 42–43.*

Este hallazgo no es exclusivo de la UNAM. Jin et al. (2024), citados en el mismo diagnóstico, señalan que la investigación sobre sistemas de IA en educación superior se concentra mayoritariamente en universidades del Norte Global, es decir, en economías desarrolladas que no representan la realidad de las instituciones ubicadas en países emergentes. La perspectiva de análisis asumida como global en numerosas investigaciones, argumentan, no lo es realmente.

Los marcos éticos existen. Lo que falta, en la mayoría de las instituciones, es la decisión de adoptarlos y la capacidad de implementarlos.

## El estado regulatorio en América Latina

En México, durante los últimos años, se han tramitado múltiples iniciativas de ley y propuestas en materia de regulación de sistemas de IA. Alcalá y Hernández (2025) documentan un conjunto significativo de propuestas presentadas entre 2021 y 2024, ninguna de las cuales prosperó. Los autores señalan que esto evidencia la complejidad de construir un marco regulatorio que equilibre regulación e innovación sin caer en procesos de sobrerregulación.

*Alcalá, P. y Hernández, O., "México: Las diversas iniciativas de regulación de la IA," centrocompetencia.com, enero de 2025.*

La ausencia de regulación no implica ausencia de necesidad. Implica que el vacío normativo deja las decisiones éticas en manos de cada institución, sin estándares mínimos comunes. En ese contexto, los marcos internacionales como la Recomendación de la UNESCO y los principios multilaterales

mapeados por Nguyen et al. funcionan como referencia de mínima para cualquier organización que quiera operar con criterio.

## 6. Qué puede hacer una institución hoy

---

Los marcos éticos internacionales no son instrumentos decorativos. Ofrecen tres aportes concretos para instituciones que quieran tomar decisiones informadas sobre IA.

### Un vocabulario común

Los principios de la UNESCO y las categorías de riesgo del AI Act proveen un lenguaje estandarizado para discutir las implicancias éticas de la IA. Ese vocabulario común es el primer paso para que una organización pueda formular políticas propias sin reinventar definiciones.

### Un marco de evaluación

La Recomendación de la UNESCO propone la Evaluación de Impacto Ético como instrumento para que Estados y organizaciones evalúen los beneficios y riesgos de los sistemas de IA, implementen medidas de prevención y mitigación, y establezcan mecanismos de reparación. No es un formulario. Es un proceso.

### Un punto de partida que no requiere legislación local

En países donde la regulación de IA aún no existe, los marcos internacionales ofrecen una base legítima para tomar decisiones éticas sin esperar que el poder legislativo actúe. El Código de Ética de la UNAM, por ejemplo, contiene al menos tres principios (honestidad, integridad y autoría intelectual) que ya son directamente aplicables al uso de herramientas de IA en docencia e investigación, aun sin mención explícita de la tecnología.

*UNAM, Código de Ética de la Universidad Nacional Autónoma de México, 2015.*

## 7. Consideraciones finales

---

El consenso ético internacional sobre IA es más amplio y más sólido de lo que la fragmentación regulatoria podría sugerir. Los principios de transparencia, privacidad, justicia, supervisión humana e inclusión aparecen de manera consistente en todos los marcos analizados, desde la Recomendación de la UNESCO hasta el AI Act europeo.

Sin embargo, la distancia entre esos principios y su implementación efectiva sigue siendo considerable. Las universidades latinoamericanas, como muestra el diagnóstico de la UNAM, están

en etapas tempranas de integración ética. La ausencia de legislación local en países como México agrava el problema pero no lo explica completamente: las herramientas para actuar ya existen.

La pregunta no es si una institución necesita un marco ético para su uso de IA. La pregunta es cuánto tiempo puede operar sin uno antes de que las consecuencias sean irreversibles.

La ética de la IA no es un problema que se resuelve con un documento. Pero empieza por tener uno que haga las preguntas correctas.

## Referencias

---

- Alcalá, P. y Hernández, O. (2025). *México: Las diversas iniciativas de regulación de la IA*. centrocompetencia.com, 22 de enero de 2025.
- Castañeda de León, L.M., Ramírez Molina, A.Y., Castillejos Reyes, J.M. y Ventura Miranda, M.T. (2025). *Uso y desarrollo ético de la Inteligencia Artificial en la Universidad: docencia e investigación*. México: DGTIC-UNAM. ISBN 978-607-587-954-3.
- Harvard Data Science Initiative (2025). *Human vs. AI — or is it?* Resumen del HDSI Winter Workshop 2025. [datascience.harvard.edu](https://datascience.harvard.edu).
- Harvard Gazette (2020). *Ethical concerns mount as AI takes bigger decision-making role*. Harvard University. [news.harvard.edu](https://news.harvard.edu).
- Nguyen, A., Ngo, H.N., Hong, Y., Dang, B. y Nguyen, B.P.T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, 28(4), 4221–4241.
- OCDE (2024). *Explanatory memorandum on the updated OECD definition of an AI System*. OECD Artificial Intelligence Papers, N.º 8.
- Parlamento Europeo (2024). *EU AI Act: first regulation on artificial intelligence*. Ficha informativa. [europarl.europa.eu](https://europarl.europa.eu).
- Seldon, A., Lakhani, P. y Luckin, R. (2020). *The Ethical Framework for AI in Education*. The Institute for Ethical AI in Education.
- UNESCO (2021). *Recomendación sobre la ética de la inteligencia artificial*. París: UNESCO. [unesdoc.unesco.org/ark:/48223/pf0000381137\\_spa](https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa).
- UNAM (2015). *Código de Ética de la Universidad Nacional Autónoma de México*. México: UNAM.
- Unión Europea (2024). *Reglamento (UE) 2024/1689 — Ley de Inteligencia Artificial (AI Act)*. Diario Oficial de la UE, 13 de junio de 2024.