



## *Inteligencia artificial en la educación: revisión de aplicaciones y desafíos éticos*

### *Artificial Intelligence in Education: A Review of Applications and Ethical Challenges*

### *Inteligência Artificial na Educação: Uma Revisão das Aplicações e Desafios Éticos*

Irma Judith Guzmán Brito <sup>I</sup>

[irmgbrito@gmail.com](mailto:irmgbrito@gmail.com)

<https://orcid.org/0009-0001-0153-8940>

Sofia Natalia Pastuña Salcedo <sup>II</sup>

[sofypastuna@gmail.com](mailto:sofypastuna@gmail.com)

<https://orcid.org/0000-0003-4179-0689>

María Teresa Garay Baño <sup>III</sup>

[mariategaray@gmail.com](mailto:mariategaray@gmail.com)

<https://orcid.org/0009-0004-8976-4193>

Marco Antonio Pazan Reyes <sup>IV</sup>

[marcopazzan@gmail.com](mailto:marcopazzan@gmail.com)

<https://orcid.org/0009-0006-9925-0959>

**Correspondencia:** [irmgbrito@gmail.com](mailto:irmgbrito@gmail.com)

Ciencias de la Educación

Artículo de Investigación

\* **Recibido:** 26 de julio de 2025 \* **Aceptado:** 22 de agosto de 2025 \* **Publicado:** 24 de septiembre de 2025

- I. Magíster en Gestión Educativa, Unidad Educativa "24 de Mayo", Ecuador.
- II. Magíster en Gestión Educativa, Unidad Educativa FAE N. 5, Ecuador.
- III. Magister en Educación Superior, Unidad Educativa Fiscomisional Pacífico Cembranos, Ecuador.
- IV. Diploma Superior en Diseño Curricular por Competencias, Unidad Educativa Fiscal "General Eloy Alfaro Delgado", Ecuador.



## Resumen

La inteligencia artificial (IA) está redefiniendo la experiencia educativa: desde tutores inteligentes y analítica del aprendizaje hasta asistentes de escritura y accesibilidad inclusiva. Este artículo ofrece una revisión narrativa y crítica de las aplicaciones actuales de la IA en educación y examina los desafíos éticos que emergen: sesgos algorítmicos, privacidad y gobernanza de datos, transparencia de modelos, integridad académica, desplazamiento de funciones docentes y sostenibilidad. Sostenemos que el valor pedagógico de la IA depende de diseños responsables, participación docente y marcos de uso centrados en el estudiante, alineados con normas internacionales como la Recomendación de la UNESCO sobre la Ética de la IA (2021) y las Orientaciones de la UNESCO sobre IA generativa en educación e investigación (2023), así como los Principios de IA de la OCDE (2019) y estándares IEEE para sistemas responsables.

### Objetivo

Revisar las principales aplicaciones educativas de la IA (tutoría, evaluación, analítica, accesibilidad, gestión) y analizar los desafíos éticos asociados, proponiendo criterios prácticos para una adopción pedagógica, segura y equitativa.

**Palabras Clave:** Inteligencia artificial; educación; tutoría inteligente; analítica del aprendizaje; ética de la IA; privacidad; sesgo algorítmico; transparencia.

## Abstract

Artificial intelligence (AI) is redefining the educational experience: from intelligent tutors and learning analytics to writing assistants and inclusive accessibility. This article offers a narrative and critical review of current applications of AI in education and examines emerging ethical challenges: algorithmic bias, data privacy and governance, model transparency, academic integrity, teacher role displacement, and sustainability. We argue that the pedagogical value of AI depends on responsible designs, teacher engagement, and student-centered frameworks for use, aligned with international standards such as the UNESCO Recommendation on the Ethics of AI (2021) and the UNESCO Guidance on Generative AI in Education and Research (2023), as well as the OECD AI Principles (2019) and IEEE standards for responsible systems. Objective

To review the main educational applications of AI (tutoring, assessment, analytics, accessibility, management) and analyze the associated ethical challenges, proposing practical criteria for safe, equitable, and pedagogical adoption.

**Keywords:** Artificial intelligence; education; intelligent tutoring; learning analytics; AI ethics; privacy; algorithmic bias; transparency.

## Resumo

A inteligência artificial (IA) está a redefinir a experiência educativa: desde tutores inteligentes e análise de aprendizagem a assistentes de escrita e acessibilidade inclusiva. Este artigo oferece uma revisão narrativa e crítica das aplicações atuais da IA na educação e examina os desafios éticos emergentes: enviesamento algorítmico, privacidade e governação de dados, transparência de modelos, integridade académica, deslocamento de papéis do professor e sustentabilidade. Argumentamos que o valor pedagógico da IA depende de designs responsáveis, do envolvimento dos professores e de estruturas de utilização centradas no aluno, alinhadas com normas internacionais como a Recomendação da UNESCO sobre a Ética da IA (2021) e a Orientação da UNESCO sobre IA Generativa na Educação e Investigação (2023), bem como os Princípios de IA da OCDE (2019) e as normas do IEEE para sistemas responsáveis. Objetivo

Rever as principais aplicações educativas da IA (tutoria, avaliação, análise de dados, acessibilidade, gestão) e analisar os desafios éticos associados, propondo critérios práticos para uma adoção segura, equitativa e pedagógica.

**Palavras-chave:** Inteligência artificial; educação; tutoria inteligente; análise da aprendizagem; ética da IA; privacidade; viés algorítmico; transparência.

## Introducción

La IA ha pasado de ser una promesa tecnológica a una infraestructura que atraviesa aulas, plataformas virtuales y procesos de gestión educativa. En el plano pedagógico, los sistemas tutoriales inteligentes (STI) y los asistentes de escritura y programación apoyan la retroalimentación personalizada; la analítica del aprendizaje permite monitorear el progreso y anticipar riesgos de deserción; los modelos de lenguaje generativos ayudan a planificar clases, crear recursos y adaptar materiales a distintos niveles; y las tecnologías de accesibilidad (reconocimiento de voz, subtítulo automático, lectura de pantalla, conversión texto-voz/voz-texto) amplían la participación de estudiantes con discapacidad (Luckin et al., 2016; Baker & Inventado, 2014; UNESCO, 2023).

El entusiasmo, sin embargo, convive con preguntas de fondo:

- ¿Qué tan justas son las decisiones algorítmicas? Conjuntos de datos desbalanceados pueden reproducir sesgos de género, etnia o territorio, afectando recomendaciones, calificaciones o apoyos (OECD, 2019; UNESCO, 2021).
- ¿Quién controla los datos educativos? La privacidad, la minimización de datos y el propósito pedagógico deben regir la recolección y el tratamiento; la transparencia y los derechos de acceso/corrección son innegociables (GDPR; UNESCO, 2023).
- ¿Cómo preservar la integridad académica y el aprendizaje auténtico? La IA generativa exige rediseñar tareas, evaluación y enseñanza de la metacognición y la verificación de fuentes (UNESCO, 2023).
- ¿Cuál es el papel del docente? Lejos de sustituirlo, la IA efectiva opera como copiloto pedagógico: libera carga repetitiva y potencia la mediación didáctica, ética y socioemocional del profesorado (Luckin et al., 2016).
- ¿Qué huella ambiental tiene la IA? El cómputo intensivo plantea un dilema de sostenibilidad que los sistemas educativos no deben soslayar (OECD, 2019).

Ante estas tensiones, organismos como la UNESCO proponen principios de legalidad, seguridad, equidad, explicabilidad, inclusión y supervisión humana significativa (UNESCO, 2021, 2023). En la misma línea, los Principios de IA de la OCDE y las normas IEEE 7000/7001 enfatizan gobernanza del ciclo de vida, gestión de riesgos, trazabilidad y mecanismos de reparación. Este artículo organiza el debate en dos ejes: (1) qué puede hacer hoy la IA en educación y (2) bajo qué condiciones éticas y de política institucional debe hacerse para ser pedagógicamente valiosa y socialmente justa.

## **Fundamentación teórica**

### **2.1. IA educativa: de la automatización al copiloto pedagógico**

La IA en educación ha evolucionado desde sistemas de recomendación y chatbots simples hacia modelos adaptativos y tutores inteligentes capaces de personalizar la instrucción, diagnosticar errores y ofrecer retroalimentación inmediata (Baker & Inventado, 2014; Luckin et al., 2016). Con la irrupción de modelos generativos (LLM), la IA opera como copiloto para docentes y estudiantes: asiste en la preparación de clases, la creación de materiales diferenciados y el apoyo a la escritura

y la programación. Este tránsito requiere marcos que preserven agencia humana, finalidades pedagógicas y evaluación auténtica (UNESCO, 2023).

## 2.2. Marcos pedagógicos que habilita la IA

- Aprendizaje adaptativo: ajusta dificultad, ritmo y contenidos a evidencias de desempeño (trazas de interacción, respuestas, tiempo). Basado en la idea de diferenciación instruccional y mastery learning, puede mejorar apoyo a quienes se rezagan sin estigmatizarlos (Pane et al., 2015).
- Sistemas Tutoriales Inteligentes (STI): modelan conocimiento del estudiante, dominio de la tarea y estrategias de tutoría para ofrecer pistas y andamiaje en el momento oportuno (VanLehn, 2011).
- Analítica del aprendizaje (LA): convierte datos en alertas tempranas, perfiles de riesgo y visualizaciones para docentes y equipos de apoyo, con el objetivo de anticipar la deserción y personalizar intervenciones (Siemens & Long, 2011).
- Accesibilidad y diseño universal (DUA): reconocimiento de voz, subtítulo automático, lectores de pantalla y síntesis de voz amplían la inclusión de estudiantes con discapacidad, barreras idiomáticas o conectividad limitada (UNESCO, 2023).

Condición clave: estas aplicaciones generan valor si se insertan en secuencias didácticas bien diseñadas, con metas de aprendizaje explícitas, rúbricas transparentes y retroalimentación formativa (Hattie & Timperley, 2007).

## 2.3. Ética de la IA en educación: principios y estándares

La Recomendación de la UNESCO sobre la Ética de la IA (2021) y las Orientaciones sobre IA generativa en educación e investigación (2023) sientan principios aplicables en escuelas y universidades:

1. Legalidad y derechos: cumplimiento de normativas de protección de datos (consentimiento informado, minimización, limitación de finalidad, derechos de acceso/rectificación).

2. Equidad y no discriminación: identificación y mitigación de sesgos algorítmicos (de representación, medición y selección) que pueden perjudicar a grupos por género, etnia o condición socioeconómica (OECD, 2019).

3. Transparencia y explicabilidad: informes claros de qué datos se usan, para qué y cómo impactan en decisiones educativas (UNESCO, 2021; IEEE, 2022).

4. Supervisión humana significativa (human-in-the-loop) y rendición de cuentas: las decisiones de alto impacto (evaluación sumativa, promoción, disciplina) no deben automatizarse sin control docente/institucional (UNESCO, 2023; OECD, 2019).

5. Seguridad y robustez: pruebas de desempeño, stress tests, planes de contingencia ante fallas y mecanismos de reparación cuando un sistema cause daño o decisión injusta (OECD, 2019; IEEE, 2022).

6. Sostenibilidad: considerar costes energéticos y de infraestructura de los modelos, priorizando soluciones eficientes y abiertas cuando sea posible (OECD, 2019).

#### **2.4. Gobernanza de datos educativos (data governance)**

La analítica y los STI dependen de datos sensibles (trayectorias, calificaciones, interacciones). Su tratamiento requiere políticas institucionales de:

- Minimización y propósito: recolectar solo lo necesario para mejorar el aprendizaje, prohibiendo usos colaterales (perfilado comercial, vigilancia no pedagógica).
- Anonimización/pseudonimización y retención limitada: borrar o desidentificar cuando la finalidad pedagógica concluye.
- Transparencia y participación: informar a estudiantes y familias; consentimiento claro y canales para oponerse a usos no educativos (UNESCO, 2023).
- Evaluaciones de Impacto Algorítmico (AIA): detectar riesgos, sesgos y efectos no intencionales antes de desplegar sistemas a gran escala (UNESCO, 2021; IEEE, 2022).

#### **2.5. Riesgos pedagógicos y académicos**

- Sesgo y estigmatización: modelos entrenados con datos históricos pueden reproducir brechas (p. ej., subestimar a minorías). Mitigación: muestreo estratificado, auditorías de equidad y evaluación continua (OECD, 2019).

- Opacidad (black box): dificulta la explicación de calificaciones o alertas. Mitigación: preferir modelos explicables cuando afecten evaluación; ofrecer justificaciones legibles (UNESCO, 2023).

- Integridad académica en la era generativa: la IA puede facilitar plagio o delegación. Respuesta: rediseño de tareas (portafolios, orales, vivas, proyectos con trazabilidad), enseñanza de metacognición y citación responsable de ayudas de IA (UNESCO, 2023).

- Desplazamiento funcional del docente: riesgo de desprofesionalización si se automatiza la enseñanza. Mitigación: IA como asistente, con foco docente en mediación, evaluación formativa y dimensiones socioemocionales (Luckin et al., 2016).

- Seguridad y privacidad: fuga de prompts/datos, prompt injection, model hallucinations. Mitigación: políticas de uso seguro, guardrails, revisión humana y listas de exclusión de datos sensibles.

## **2.6. Criterios de despliegue responsable (síntesis operativa)**

Antes de adoptar un sistema de IA educativa, una institución debería verificar:

1. Finalidad pedagógica clara y evidencia de efectividad (impacto en aprendizaje, no solo en eficiencia).
2. Datos: minimizados, seguros, con consentimiento y propiedades definidas (quién accede, por qué, cuánto tiempo).
3. Equidad: pruebas A/B y auditorías de desempeño por subgrupos; umbrales de parity acordados.
4. Explicabilidad: reportes comprensibles para estudiantes/docentes; bitácoras de decisiones.
5. Supervisión humana: límites a la automatización en evaluación sumativa/disciplinaria.
6. Gobernanza: comité de ética/tecnología educativa, procesos de queja y reparación.
7. Sostenibilidad: costos energéticos, infraestructura y soporte docente (formación).

### **Aplicaciones educativas de la IA (panorama con buenas prácticas)**

En cada subapartado te doy: qué hace la IA, para qué sirve pedagógicamente, buenas prácticas y alertas éticas. Las referencias clave (UNESCO, OECD, IEEE, y literatura especializada) se listarán completas en el bloque final.

#### **3.1. Tutores inteligentes y aprendizaje adaptativo**

Qué es: Sistemas que modelan el conocimiento del estudiante y ajustan pistas, ejercicios y ritmo según su desempeño (Baker & Inventado, 2014; VanLehn, 2011).

Valor pedagógico: Retroalimentación inmediata, práctica deliberada y andamiaje personalizado; útil en Matemáticas, Lengua, Programación.

Buenas prácticas: Metas de mastery, explicaciones paso a paso, detección de errores frecuentes y visualización del progreso para docentes y estudiantes.

Alertas: Evitar “enseñar al test”. Requiere supervisión docente y revisión de sesgos en datos (UNESCO, 2021; OECD, 2019).

#### **3.2. Analítica del aprendizaje y sistemas de alerta temprana**

Qué es: Paneles que convierten interacciones (tiempo en tareas, entregas, aciertos) en indicadores de riesgo de deserción o rezago (Siemens & Long, 2011).

Valor: Intervención temprana y tutorías focalizadas; gestión de grupos grandes en entornos híbridos.

Buenas prácticas: Explicar qué variables alimentan las alertas; acciones concretas (mensajes empáticos, tutorías, recursos remediales).

Alertas: Privacidad y propósito pedagógico claro; auditorías de desempeño por subgrupos para evitar perfiles estigmatizantes (UNESCO, 2023; OECD, 2019).

#### **3.3. Evaluación asistida por IA (formativa y sumativa)**

Qué es: Corrección automática de ítems cerrados, apoyo a la evaluación de textos/código y generación de rúbricas.

Valor: Ahorra tiempo en retroalimentación formativa; ayuda a detectar patrones de error.

Buenas prácticas: La IA sugiere, el docente decide; mantener muestras doblemente calificadas para control de calidad; mostrar criterios a estudiantes (Hattie & Timperley, 2007).

Alertas: Evitar uso opaco en calificaciones finales; exigir explicabilidad y revisión humana (UNESCO, 2023; IEEE, 2022).

### **3.4. IA generativa como “copiloto” de escritura y proyectos**

Qué es: Modelos de lenguaje que borran texto, resumen, proponen ideas, guiones o esqueletos de código.

Valor: Ayuda en preescritura, revisión y planificación; facilita la diferenciación (niveles/estilos) y la traducción accesible.

Buenas prácticas: Pedagogía de pensamiento visible: diarios de proceso, trazabilidad (borradores), citación del uso de IA, consignas que pidan criterio y evidencia.

Alertas: Riesgo de dependencia y plagio por delegación. Diseñar tareas que requieran viva/defensa oral, datos locales, producción multimodal y conexión con experiencias (UNESCO, 2023).

### **3.5. Accesibilidad e inclusión (DUA con IA)**

Qué es: Subtitulado automático, reconocimiento de voz, lectores de pantalla, TTS/STT, simplificación de textos, traducción.

Valor: Reduce barreras para estudiantes con discapacidad, diversidad lingüística o conectividad limitada; mejora la participación y la autonomía (UNESCO, 2023).

Buenas prácticas: Ajustes razonables personalizados; validación con usuarios reales; opciones offline y consumo eficiente de datos.

Alertas: Cuidar privacidad de voz/imágenes y calidad de traducciones (evitar sesgos o términos ofensivos).

### **3.6. STEM: simulaciones, laboratorios virtuales y feedback en código**

Qué es: Simuladores de física/química/biología, entornos de programación con sugerencias y diagnósticos de errores.

Valor: Experimentación segura y repetible, modelos de fenómenos complejos, aprendizaje por indagación; feedback granular en programación.

Buenas prácticas: Hipótesis → experimento → análisis (ciclo científico), rubricas claras y reflexión sobre resultados.

Alertas: Explicar limitaciones del modelo/simulador; evitar que la IA oculte el proceso de resolución.

### **3.7. Aprendizaje de lenguas y oralidad asistida**

Qué es: Detección de pronunciación, diálogos simulados, corrección de gramática/vocabulario en contexto.

Valor: Práctica intensiva y segura; exposición a múltiples acentos; role-play guiado.

Buenas prácticas: Escalas CEFR, objetivos comunicativos, retroalimentación específica y registro de progresos.

Alertas: Evitar estandarizar acentos; respetar identidades lingüísticas y contextos culturales.

### **3.8. Gestión educativa y apoyo al docente**

Qué es: Organización de calendarios, generación de secuencias didácticas, creación de ítems y bancos de preguntas, analítica de carga de trabajo.

Valor: Libera tiempo para mediación, acompañamiento y feedback individual.

Buenas prácticas: Revisión crítica del material generado; co-diseño entre pares; repositorios compartidos con licencias libres.

Alertas: Evitar “comoditizar” el rol docente; promover desarrollo profesional continuo en IA.

### **3.9. Recursos abiertos y curaduría con IA**

Qué es: Asistentes que buscan, resumen y comparan recursos (OER), generando dossiers por nivel/objetivo.

Valor: Acelera la curaduría y el diseño de itinerarios; apoya la actualización docente.

Buenas prácticas: Exigir fuentes citadas y fechas; validar con criterios de calidad (relevancia, sesgo, nivel).

Alertas: Alucinaciones y desactualización: verificación obligatoria y pensamiento crítico explícito.

### **3.10. Cómo decidir si una aplicación de IA “vale la pena”**

- Alineación pedagógica: ¿Mejora el aprendizaje (evidencia), o solo la eficiencia?
- Equidad: ¿Funciona de forma comparable en subgrupos? ¿Qué sesgos puede introducir?

- Datos y privacidad: ¿Qué datos recolecta? ¿Con qué propósito y por cuánto tiempo?
- Explicabilidad: ¿Puedo explicar sus decisiones a estudiantes y familias?
- Supervisión humana: ¿Qué queda siempre en manos del docente?
- Sostenibilidad: ¿Cuánto cuesta (energía, dinero, soporte) sostenerlo?

### **Desafíos éticos y cómo gestionarlos**

En cada punto: riesgo → por qué importa en educación → cómo mitigarlo (criterios operativos para escuelas y universidades).

#### **4.1. Sesgo algorítmico y discriminación**

Riesgo. Conjuntos de datos desbalanceados o mal rotulados reproducen estereotipos (género, etnia, territorio, discapacidad), afectando recomendaciones, alertas de riesgo y calificaciones.

Por qué importa. Puede estigmatizar a estudiantes, asignar apoyos de forma inequitativa y reforzar brechas de logro.

Mitigación. Auditorías de equidad por subgrupos (paridad de error/falsa alarma), data sheets/model cards, curaduría de datos con muestreo estratificado y evaluaciones de impacto algorítmico previas al despliegue. Documentar supuestos y límites (OECD, 2019; UNESCO, 2021, 2023; IEEE, 2022).

#### **4.2. Privacidad, consentimiento y gobernanza de datos**

Riesgo. Recolección excesiva, reutilización no pedagógica, transferencias a terceros y fugas de información (incluidos prompts y voces/imágenes).

Por qué importa. Se vulnera el derecho a la educación sin vigilancia y se desincentiva la participación honesta en plataformas.

Mitigación. Minimización y finalidad pedagógica explícita; bases legales claras; anonimización/pseudonimización; retención limitada y opt-out razonable. Políticas de acceso/rectificación y transparencia para estudiantes y familias. Incluir consentimiento informado diferenciado para IA generativa y reconocimiento de voz/imagen (UNESCO, 2023; GDPR-like; OECD, 2019).

#### **4.3. Opacidad y explicabilidad (black box)**

Riesgo. Modelos opacos dificultan explicar alertas, notas o recomendaciones.

Por qué importa. Sin explicaciones comprensibles, no hay debido proceso educativo ni posibilidad de corregir errores.

Mitigación. Preferir modelos explicables en decisiones de alto impacto; model cards y reportes legibles (“qué datos, para qué, cómo decide”). Mantener bitácoras y canales de corrección/impugnación (UNESCO, 2021; IEEE, 2022).

#### **4.4. Supervisión humana significativa y rendición de cuentas**

Riesgo. Automatizar evaluación sumativa, disciplina o promoción sin control docente/institucional.

Por qué importa. Se desplaza la responsabilidad profesional y se multiplican errores difíciles de revertir.

Mitigación. Políticas de human-in-the-loop: la IA sugiere y el docente decide. Prohibir automatizar decisiones críticas; comités de ética/tecnología educativa, protocolos de reparación cuando el sistema cause daño (UNESCO, 2023; OECD, 2019).

#### **4.5. Integridad académica con IA generativa**

Riesgo. Delegación del trabajo intelectual, plagio, contract cheating.

Por qué importa. Erosionan el aprendizaje profundo y la evaluación justa.

Mitigación. Rediseñar tareas: portafolios con trazabilidad, defensas orales, datos locales, proyectos colaborativos, diarios de proceso y citación responsable del uso de IA. Evaluación formativa y rúbricas que valoren razonamiento y fuentes (UNESCO, 2023).

#### **4.6. Desprofesionalización docente y dependencia tecnológica**

Riesgo. Convertir al docente en operador de plataforma; currículos dictados por lo disponible en la herramienta.

Por qué importa. Se empobrece la mediación pedagógica y el juicio profesional.

Mitigación. IA como copiloto: liberar tareas repetitivas y fortalecer planificación, feedback y acompañamiento socioemocional. Plan de desarrollo profesional docente en IA y comunidades de práctica (Luckin et al., 2016).

#### **4.7. Sostenibilidad y huella ambiental**

Riesgo. Costos energéticos, hardware obsoleto, lock-in con proveedores.

Por qué importa. Incoherencia con educación para el desarrollo sostenible y presupuestos limitados.

Mitigación. Priorizar modelos eficientes, opciones on-premise o edge cuando proceda, y recursos abiertos. Evaluar TCO (coste total de propiedad) y plan de renovación responsable (OECD, 2019).

#### **4.8. Inclusión y accesibilidad reales (más allá del “checkbox”)**

Riesgo. Herramientas que no contemplan diversidad lingüística, discapacidad o conectividad.

Por qué importa. La IA puede ensanchar brechas si solo funciona para quien ya está incluido.

Mitigación. Diseño Universal para el Aprendizaje (DUA), pruebas con usuarios diversos, modos offline y bajo consumo de datos, soporte multilingüe y lectura fácil (UNESCO, 2023).

#### **4.9. Seguridad de sistemas y contenidos**

Riesgo. Prompt injection, data poisoning, deepfakes, alucinaciones.

Por qué importa. Impacta reputación, confianza y resultados de aprendizaje.

Mitigación. Guardrails, listas de exclusión de datos sensibles, verificación de hechos, content filters, planes de contingencia y alfabetización en IA para docentes/estudiantes.

#### **4.10. Lista de verificación ética para adopción institucional**

1. Propósito pedagógico y evidencia de impacto.
2. Datos: minimizados, protegidos, con consentimiento.
3. Equidad: auditorías por subgrupos y metas de paridad.
4. Explicabilidad: informes comprensibles y vías de reclamo.
5. Supervisión humana: no automatizar decisiones críticas.
6. Gobernanza: comité, políticas públicas y mecanismos de reparación.
7. Sostenibilidad: costo energético/económico y alternativas abiertas.
8. Formación: desarrollo profesional docente y alfabetización en IA.

## Conclusiones

1. La IA ya agrega valor pedagógico, especialmente en tutoría inteligente, analítica del aprendizaje, accesibilidad y apoyo a la escritura/código. Su mayor virtud no es “automatizar”, sino personalizar y amplificar la retroalimentación formativa, liberar tiempo docente y abrir oportunidades de inclusión.

2. Sin gobierno ético, el valor se erosiona. Los riesgos —sesgo, opacidad, mal uso de datos, desprofesionalización, integridad académica— no son inevitables: dependen de cómo se diseñan y despliegan los sistemas. La evidencia internacional converge en principios de equidad, transparencia y supervisión humana significativa (UNESCO/OCDE/IEEE).

3. El docente es insustituible. La IA eficaz opera como copiloto pedagógico: potencia la mediación didáctica y socioemocional, pero no la reemplaza. Decidir qué, cómo y por qué aprender sigue siendo una tarea humana, situada y ética.

4. El aprendizaje auténtico exige rediseño. Con IA generativa disponible, las tareas deben evaluar proceso, criterio y fuentes, no solo producto final. Portafolios con trazabilidad, defensas orales y proyectos con datos/contextos locales son claves.

5. La inclusión es la prueba decisiva. Una IA educativa que funciona solo para quien ya estaba incluido amplía brechas. Diseño Universal, pruebas con usuarios diversos, opciones offline y gobernanza de datos centrada en estudiantes son condiciones de justicia educativa.

## Recomendaciones

### A. Para instituciones (escuelas y universidades)

- Política de IA educativa con principios claros (propósito pedagógico, equidad, privacidad, explicabilidad, supervisión humana, sostenibilidad) y gobernanza (comité de ética/tecnología, evaluación de impacto algorítmico, vías de queja y reparación).
- Marco de datos: minimización y finalidad pedagógica; consentimiento informado granular; anonimización/pseudonimización; retención limitada; transparencia para estudiantes/familias.
- Adopción gradual con pilotos: pruebas controladas, métricas de impacto (aprendizaje, equidad por subgrupos, carga docente), revisión a los 90/180 días.
- Desarrollo profesional docente en IA (diseño de prompts, evaluación auténtica, mitigación de sesgos, privacidad y seguridad).

## B. Para docentes

- IA como coproductora de feedback: usarla para generar explicaciones alternativas, rúbricas y ejemplos, pero validar y adaptar al contexto del grupo.
- Rediseñar evaluación: portafolios con borradores, diarios de proceso, debates/defensas orales, trabajos con fuentes verificables y datos locales; citación responsable del uso de IA.
- Protección de datos: no introducir datos sensibles de estudiantes en servicios externos; preferir herramientas con contratos y DPA compatibles con la normativa.

## C. Para áreas técnicas / edtech

- Auditorías de equidad: métricas de parity (precisión/errores por género, etnia, discapacidad, condición socioeconómica), model cards y documentación de datos.
- Explicabilidad: reportes legibles de cómo decide el sistema; bitácoras y posibilidad de impugnar decisiones.
- Seguridad: guardrails, listas de exclusión (PII), pruebas de prompt injection y data poisoning; planes de contingencia.
- Sostenibilidad: evaluar TCO/huella energética; explorar modelos eficientes, edge/on-premise y recursos abiertos cuando proceda.

## D. Para estudiantes y familias

- Alfabetización en IA: qué puede y qué no puede la herramienta; cómo citar su uso; riesgos de sesgo y privacidad.
- Autonomía y ética: aprovechar IA para aprender mejor (planificar, revisar, practicar), no para evadir el esfuerzo cognitivo.
- Participación informada: conocer derechos sobre sus datos y canales de reclamo.

## Ruta de implementación en 90 días (esquema práctico)

- Días 0–30: política institucional mínima; inventario de herramientas; formación docente inicial; piloto acotado (una materia).
- Días 31–60: medición de impacto (aprendizaje, tiempo docente, equidad), ajustes de privacidad y explicabilidad; rediseño de 2–3 evaluaciones por curso.
- Días 61–90: expansión a más asignaturas; comité de ética activo; publicación de métricas y lecciones aprendidas; plan anual de mejora.

## Referencias

1. UNESCO. (2021). Recomendación sobre la Ética de la Inteligencia Artificial. París: UNESCO.
2. UNESCO. (2023). Orientaciones sobre la IA generativa en la educación y la investigación. París: UNESCO.
3. UNESCO. (2019). Consenso de Beijing sobre la inteligencia artificial y la educación. París: UNESCO.
4. OCDE. (2019). Recomendación del Consejo sobre Inteligencia Artificial (Principios de la OCDE sobre IA). París: Organización para la Cooperación y el Desarrollo Económicos.
5. IEEE. (2021). IEEE 7000-2021: Model Process for Addressing Ethical Concerns During System Design. Piscataway, NJ: IEEE Standards Association.
6. IEEE. (2021). IEEE 7001-2021: Transparency of Autonomous Systems. Piscataway, NJ: IEEE Standards Association.
7. UNICEF. (2021). Orientación sobre políticas de IA para la infancia. Nueva York: Fondo de las Naciones Unidas para la Infancia.
8. HLEG on AI (UE). (2019). Directrices éticas para una IA fiable. Bruselas: Comisión Europea.
9. Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. En J. A. Larson & M. C. Yudelson (Eds.), *Encyclopedia of Education and Information Technologies* (pp. 1–7). Springer.
10. Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence Unleashed: An argument for AI in Education*. London: Pearson.
11. VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221.
12. Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5), 30–40.
13. Pane, J. F., Steiner, E. D., Baird, M. D., & Hamilton, L. S. (2015). *Continued Progress: Promising Evidence on Personalized Learning*. Santa Monica, CA: RAND Corporation.

14. Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
15. Holmes, W., Porayska-Pomsta, K., & Holstein, K. (2022). Ethics of AI in Education: Towards a Community-Wide Framework. *British Journal of Educational Technology*, 53(4), 658–675.
16. Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223–235.
17. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on AI applications in higher education. *International Journal of Educational Technology in Higher Education*, 16(1), 39.
18. Seldon, A., & Abidoeye, O. (2018). *The Fourth Education Revolution: Will Artificial Intelligence liberate or infantilise humanity?* University of Buckingham Press.
19. Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education.* Polity Press.
20. OCDE. (2021). *Inteligencia artificial, integridad académica y evaluación en educación superior (Policy Brief).* París: OCDE.
21. UNESCO. (2022). *Reimaginar juntos nuestros futuros: Un nuevo contrato social para la educación (ed. esp.).* París: UNESCO.
22. OECD. (2021). *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development (Working Paper).* París: OCDE.
23. Williamson, B. (2017). *Big Data in Education: The digital future of learning, policy and practice.* SAGE.
24. Kizilcec, R. F., & Lee, H. (2020). Algorithmic fairness in education. *AI Magazine*, 41(4), 37–48.
25. Custaud, M., Holmes, W., Porayska-Pomsta, K., & Griffiths, M. (2023). Generative AI in education: Opportunities, risks, and policy responses. *Computers & Education: Artificial Intelligence*, 4, 100154.