



Ethics, transparency and AI governance in educational decision-making

Ética, transparencia y gobernanza de la IA en la toma de decisiones educativas

Para citar este trabajo:

Sucre, A., Mitre, M., & Consuegra, D. . (2026). Ética, transparencia y gobernanza de la IA en la toma de decisiones educativas. *Star of Sciences Multidisciplinary Journal*, 3(1), 1-14. <https://doi.org/10.63969/64h3sr59>

Autores:

Antonio Sucre

Universidad de Panamá - Facultad de Economía

Los Santos – Panamá

antonio.sucre@up.ac.pa

<https://orcid.org/0009-0000-0243-277X>

María Mitre

Universidad de Panamá - Facultad de Informática

Azuero– Panamá

maria.mitre@up.ac.pa

<https://orcid.org/0009-0000-8154-025X>

Delia Consuegra

Universidad de Panamá - Facultad de Informática

Los Santos – Panamá

delia.consuegra@up.ac.pa

<https://orcid.org/0000-0002-4661-6578>

Autor de Correspondencia: Delia Consuegra, delia.consuegra@up.ac.pa

RECIBIDO: 19-Enero-2026

ACEPTADO: 02-Febrero-2026

PUBLICADO: 16-Febrero-2026



Resumen

La incorporación de la inteligencia artificial en los sistemas educativos está transformando los procesos de enseñanza, aprendizaje y gestión institucional mediante el uso de algoritmos y sistemas automatizados para la toma de decisiones. Si bien estas tecnologías ofrecen oportunidades para mejorar la eficiencia y la calidad educativa, también plantean desafíos éticos, de transparencia y de gobernanza que afectan derechos, equidad y responsabilidad institucional. El estudio examina los enfoques éticos, de transparencia y de gobernanza vinculados a la IA en decisiones educativas, a partir de un análisis documental de fuentes académicas y normativas, con el fin de identificar principios orientadores, problemáticas recurrentes y criterios analíticos relevantes.

Palabras clave: Ética; Transparencia; Gobernanza de la inteligencia artificial; Toma de decisiones educativas; Sistemas automatizados.

Abstract

The incorporation of artificial intelligence into educational systems is transforming teaching, learning and institutional management processes through the use of algorithms and automated systems for decision-making. While these technologies offer opportunities to enhance efficiency and educational quality, they also raise ethical, transparency and governance challenges that affect rights, equity and institutional accountability. This study examines ethical, transparency and governance approaches related to AI in educational decision-making, based on a documentary analysis of academic and regulatory sources, with the aim of identifying guiding principles, recurring issues and relevant analytical criteria.

Keywords: Ethics; Transparency; Artificial intelligence governance; Educational decision-making; Automated systems.



1. Introducción

La incorporación progresiva de la inteligencia artificial (IA) en los sistemas educativos está produciendo transformaciones sustantivas en los procesos de enseñanza, aprendizaje y gestión institucional. En particular, la utilización de algoritmos predictivos y sistemas automatizados para respaldar la toma de decisiones educativas —tales como la evaluación del desempeño estudiantil, la adaptación personalizada de trayectorias de aprendizaje y la optimización de la gestión académica— abre oportunidades relevantes para incrementar la eficiencia operativa y fortalecer la calidad educativa. No obstante, estas aplicaciones tecnológicas también introducen complejidades éticas, jurídicas y pedagógicas que demandan un análisis crítico y teóricamente fundamentado desde la investigación educativa.

En este escenario, los conceptos de ética, transparencia y gobernanza de la IA adquieren una centralidad estratégica, en tanto las decisiones automatizadas inciden de manera directa en los derechos fundamentales, las trayectorias formativas y las oportunidades educativas de los estudiantes. La opacidad algorítmica, la presencia de sesgos derivados de los datos y la carencia de marcos normativos y de gobernanza robustos pueden profundizar desigualdades preexistentes y vulnerar principios esenciales de justicia, equidad y responsabilidad institucional. En consecuencia, resulta indispensable examinar la literatura científica especializada para comprender cómo estos principios están siendo abordados y aplicados en el ámbito educativo contemporáneo.

El crecimiento acelerado del uso de sistemas de inteligencia artificial en la toma de decisiones educativas ha superado, en numerosos contextos, el desarrollo de marcos éticos, normativos y regulatorios capaces de orientar y controlar adecuadamente su implementación. Tanto instituciones educativas como organismos gubernamentales han incorporado soluciones algorítmicas para apoyar procesos evaluativos, administrativos y académicos sin disponer de criterios claramente definidos sobre responsabilidad institucional, explicabilidad de los modelos y mecanismos efectivos de rendición de cuentas. Esta situación favorece la generación de decisiones opacas, difíciles de comprender, auditar o cuestionar por parte de docentes, estudiantes y otros actores del sistema educativo.

De igual forma, la evidencia científica señala que los sistemas de IA pueden reproducir y amplificar sesgos presentes en los datos educativos, reforzando patrones de exclusión y afectando de manera desproporcionada a poblaciones históricamente vulnerables. La ausencia de mecanismos sólidos de transparencia y supervisión limita la comprensión de los criterios que sustentan las recomendaciones o decisiones automatizadas, reduciendo la capacidad de detectar errores, corregir inequidades y garantizar procesos educativos justos, inclusivos y equitativos.

Finalmente, la gobernanza de la inteligencia artificial en el ámbito educativo se configura como un desafío de carácter estructural, al requerir la articulación coherente entre políticas públicas, regulaciones institucionales, marcos éticos y prácticas pedagógicas. La falta de enfoques integrales y coordinados dificulta la adopción responsable de estas tecnologías y debilita la confianza de la comunidad educativa en los sistemas automatizados de decisión, lo que pone de relieve la necesidad de fortalecer los fundamentos que orientan la investigación y el diseño de políticas educativas.

El desarrollo de principios orientadores para el uso responsable de la inteligencia artificial ha cobrado relevancia ante su creciente presencia en contextos educativos. Estos principios buscan garantizar que los sistemas automatizados contribuyan al bienestar humano, eviten daños, respeten la autonomía y promuevan condiciones de justicia. En el ámbito educativo, donde las decisiones inciden directamente en los procesos formativos y en las trayectorias académicas de los estudiantes, UNESCO (2020) enfatizan que estos principios deben ajustarse a las



particularidades institucionales y pedagógicas, considerando su impacto en el desarrollo integral y en la dignidad de las personas.

La incorporación de la inteligencia artificial en educación suele estar acompañada de discursos centrados en la eficiencia, la optimización y la mejora del rendimiento, lo que puede desplazar reflexiones de carácter ético y social. Este enfoque prioriza soluciones técnicas sin problematizar suficientemente sus consecuencias educativas. En este sentido, Cepeda et al. (2025) advierte que muchas aplicaciones de IA responden a una lógica tecnocrática que tiende a invisibilizar cuestiones relacionadas con la equidad, la inclusión y la justicia social, reproduciendo desigualdades ya existentes en los sistemas educativos.

El incremento sostenido de investigaciones sobre inteligencia artificial aplicada a la educación refleja un notable interés por el desarrollo de herramientas y modelos algorítmicos. Sin embargo, este crecimiento no se distribuye de manera equilibrada entre los distintos enfoques de análisis. Al examinar la producción académica del campo, Salvatierra (2024) señalan que los estudios centrados en aspectos éticos y de gobernanza siguen siendo considerablemente menores en comparación con aquellos orientados a la automatización y al rendimiento académico.

El uso de algoritmos en procesos de gestión, evaluación y seguimiento educativo está transformando las formas tradicionales de toma de decisiones institucionales. Estas transformaciones introducen nuevas dinámicas de control basadas en datos y métricas. En este contexto, Sánchez (2023) sostiene que la toma de decisiones apoyada en sistemas algorítmicos redefine las relaciones de poder dentro de los sistemas educativos, desplazando progresivamente el juicio humano y generando la necesidad de mecanismos claros de supervisión y regulación.

Las orientaciones promovidas por organismos internacionales han adquirido un papel relevante en la definición de criterios para el uso responsable de la inteligencia artificial. Estas directrices buscan asegurar que los sistemas automatizados respeten derechos fundamentales y valores educativos. En este marco, UNESCO (2024) subraya la importancia de garantizar la transparencia, la explicabilidad y la responsabilidad en los sistemas de IA utilizados en educación, con especial atención a la protección de los derechos de los estudiantes.

La expansión de sistemas de inteligencia artificial en los entornos educativos también ha generado preocupaciones vinculadas con la vigilancia, la recopilación masiva de datos y la automatización de decisiones pedagógicas. Estas prácticas pueden afectar el rol profesional del docente y limitar su capacidad de intervención. Al analizar estos riesgos, Rondon et al. (2025) advierten que la reducción de la intervención humana en procesos clave puede debilitar la autonomía docente y empobrecer la experiencia educativa.

La formulación de guías éticas para la inteligencia artificial ha dado lugar a múltiples marcos de referencia a escala internacional. No obstante, la existencia de estos documentos no garantiza su aplicación efectiva en sectores específicos como la educación. A partir de un análisis comparativo de estas guías, Sánchez et al. (2025) identifican que la transparencia y la gobernanza se reconocen de forma reiterada como principios centrales, aunque su implementación práctica en contextos educativos sigue siendo limitada.

El diseño de políticas orientadas a regular el uso de la inteligencia artificial en educación responde a la necesidad de integrar dimensiones éticas, legales y pedagógicas en un marco coherente. Estas políticas buscan equilibrar la innovación tecnológica con la protección de los valores educativos y sociales. Desde esta perspectiva, OECD (2025) destaca la importancia de desarrollar marcos de gobernanza que promuevan un uso responsable y sostenible de la IA en los sistemas educativos, fortaleciendo la legitimidad y la confianza de la comunidad educativa en estas tecnologías.



El análisis ético de los sistemas de inteligencia artificial en educación requiere un enfoque que considere la información como un elemento central de los procesos sociales y tecnológicos. Desde esta perspectiva, la información no es neutral, sino que produce efectos directos sobre las personas y las instituciones. En este marco, OECD (2022) plantea que los procesos algorítmicos y el tratamiento de datos deben gestionarse de manera responsable, ya que una administración inadecuada de la información puede afectar la dignidad humana, especialmente cuando las decisiones automatizadas inciden en oportunidades educativas, trayectorias formativas y procesos de evaluación.

El funcionamiento de los algoritmos utilizados en sistemas automatizados presenta múltiples desafíos vinculados con su diseño y aplicación en contextos sensibles como el educativo. Entre estos desafíos se encuentran la dificultad para comprender cómo se generan las decisiones, la ausencia de responsables claramente identificables y la presencia de sesgos derivados de los datos. En este sentido, González et al. (2024) señalan que la opacidad algorítmica y la falta de mecanismos de responsabilidad constituyen riesgos significativos, particularmente cuando los sistemas de IA influyen en decisiones educativas que afectan el acceso, la permanencia y el rendimiento de los estudiantes.

La evaluación de la equidad en los sistemas de inteligencia artificial utilizados en educación exige criterios normativos que permitan analizar si las decisiones automatizadas distribuyen beneficios y cargas de manera justa. En este contexto, los aportes clásicos sobre justicia social ofrecen herramientas relevantes para examinar el impacto de la IA en la educación. Desde este enfoque, UNESCO (2025) proporciona principios que permiten analizar si las decisiones apoyadas por sistemas automatizados respetan condiciones de equidad y favorecen a los grupos más desfavorecidos dentro de los sistemas educativos.

La exigencia de transparencia en los sistemas automatizados se ha convertido en un elemento clave para garantizar la confianza y la legitimidad de las decisiones basadas en IA. La posibilidad de comprender cómo funcionan los algoritmos y qué criterios utilizan resulta fundamental en entornos educativos, donde las decisiones deben ser justificables ante estudiantes y docentes. En este sentido, Araya (2021) introduce la noción de transparencia algorítmica, destacando que la explicabilidad de los sistemas automatizados es una condición necesaria para asegurar la rendición de cuentas y la supervisión institucional.

El diseño y la implementación de sistemas de inteligencia artificial en educación plantean la necesidad de preservar el rol activo de las personas en los procesos decisionales. La automatización no debe sustituir completamente el juicio humano, especialmente en decisiones de alto impacto educativo. Desde esta perspectiva, Miao et al. (2021) sostiene que los sistemas de IA deben mantenerse bajo control humano significativo, garantizando que docentes y autoridades educativas conserven la capacidad de intervenir, evaluar y corregir las decisiones automatizadas relacionadas con el aprendizaje y la evaluación.

La integración de la inteligencia artificial en los sistemas educativos no ocurre en un vacío normativo, sino dentro de estructuras institucionales y políticas más amplias. La regulación de estas tecnologías requiere marcos claros que definan responsabilidades y promuevan la participación de los distintos actores educativos. En este contexto, Valenzuela et al. (2021) señalan que la gobernanza digital en educación debe incorporar la voz de la comunidad educativa, evitando que las decisiones algorítmicas se impongan de manera unilateral desde lógicas técnicas o comerciales.

Las orientaciones impulsadas por organismos internacionales han reforzado la necesidad de establecer políticas públicas que regulen el uso de la inteligencia artificial desde una perspectiva ética y social. Estas políticas buscan garantizar que la innovación tecnológica no entre en conflicto



con los valores educativos ni con los derechos fundamentales. En este marco, Arriagada et al. (2024) destaca la importancia de alinear el desarrollo y uso de la IA con principios como la equidad, la inclusión y el respeto a los derechos humanos en los sistemas educativos.

La legitimidad de las decisiones apoyadas por inteligencia artificial en educación depende, en gran medida, de la posibilidad de establecer procesos comunicativos transparentes y comprensibles. La aceptación social de estas decisiones requiere espacios de diálogo y justificación racional. Desde esta perspectiva, Habermas mencionado por Martínez (2023) aporta elementos clave para comprender la importancia del consenso, la comunicación y la legitimidad en los procesos de toma de decisiones, aspectos fundamentales cuando la IA interviene en ámbitos educativos que afectan directamente a las personas.

El estudio se desarrolla a partir de una revisión documental bibliográfica orientada al análisis crítico de literatura científica, informes institucionales y marcos normativos vinculados con la ética, la transparencia y la gobernanza de la inteligencia artificial en el ámbito educativo. Este enfoque metodológico permite examinar de manera ordenada los principales aportes académicos, identificar tendencias relevantes y reconocer vacíos existentes en la producción científica, proporcionando una base sólida para comprender el estado actual del conocimiento en el campo.

Examinar los enfoques éticos, de transparencia y de gobernanza de la inteligencia artificial vinculados a la toma de decisiones educativas, a partir del análisis de fuentes académicas y normativas, con el propósito de identificar principios orientadores, problemáticas recurrentes y criterios de análisis presentes en la producción científica especializada.

A partir de las problemáticas identificadas en torno al uso de la inteligencia artificial en los procesos de toma de decisiones educativas, surge la necesidad de orientar el análisis hacia la identificación de los enfoques y principios que estructuran este campo de estudio. En este sentido, la presente revisión se articula en torno a la siguiente pregunta de investigación: ¿Cuáles son los enfoques, principios y aportes predominantes en los estudios y documentos especializados en torno a la ética, la transparencia y la gobernanza de la inteligencia artificial en la toma de decisiones educativas? Esta pregunta permite delimitar el alcance del análisis y focalizar la atención en los elementos que configuran las discusiones actuales sobre el uso responsable de la IA en el ámbito educativo.

2. Metodología

El trabajo se desarrolló a partir de un enfoque documental de carácter analítico y descriptivo, orientado al examen crítico de estudios, documentos institucionales y marcos normativos relacionados con la ética, la transparencia y la gobernanza de la inteligencia artificial en contextos educativos. Dada la naturaleza del estudio, el análisis se centró exclusivamente en la exploración, comparación e integración de aportes previamente publicados, sin recurrir a la aplicación de técnicas empíricas, levantamiento de datos primarios ni medición directa de variables.

El proceso de localización y selección de las fuentes se realizó mediante una estrategia de búsqueda estructurada en bases de datos académicas indexadas y repositorios especializados en educación, tecnologías digitales, políticas públicas y gobernanza algorítmica. Para ello, se emplearon combinaciones de términos clave vinculados con inteligencia artificial, toma de decisiones educativas, ética, transparencia, gobernanza, automatización y regulación, articulados mediante operadores booleanos con el fin de ampliar y precisar el alcance de los resultados obtenidos.



Una vez recopilados los registros iniciales, se procedió a la depuración de los documentos mediante la eliminación de duplicados y la revisión preliminar de títulos y resúmenes. En esta etapa se descartaron aquellos textos que no abordaban de forma directa la relación entre inteligencia artificial y toma de decisiones educativas, que se centraban exclusivamente en aspectos técnicos sin vinculación con dimensiones éticas o de gobernanza, o que carecían de pertinencia para el ámbito educativo. Como resultado de este proceso, se seleccionó un conjunto de documentos para su análisis en texto completo.

La evaluación detallada de los textos completos se realizó aplicando criterios de inclusión previamente definidos, tales como la claridad argumentativa, la relevancia temática y la vinculación explícita con problemáticas éticas, de transparencia o de gobernanza en educación. Durante esta fase se excluyeron aquellos documentos que presentaban debilidades analíticas o abordajes tangenciales del tema, conformándose así el corpus final de análisis.

El examen de los documentos seleccionados se llevó a cabo mediante procedimientos propios del análisis documental cualitativo. Para apoyar la organización, clasificación y sistematización de la información, se utilizó Microsoft Excel como herramienta auxiliar, permitiendo registrar características de los documentos, agrupar categorías analíticas y facilitar la comparación entre enfoques, autores y líneas de análisis. Este procedimiento contribuyó a estructurar de manera ordenada los contenidos relacionados con ética de la inteligencia artificial, transparencia algorítmica, rendición de cuentas, gobernanza educativa y toma de decisiones automatizadas.

De manera complementaria, el razonamiento inductivo permitió identificar regularidades y tendencias predominantes en los documentos analizados, mientras que el análisis comparativo facilitó el examen de coincidencias y divergencias entre enfoques normativos e institucionales. Este proceso hizo posible reconocer vacíos analíticos y tensiones recurrentes en el tratamiento de la inteligencia artificial en la toma de decisiones educativas.

La organización de los hallazgos se realizó mediante un análisis temático, integrando los aportes revisados en núcleos interpretativos relacionados con los desafíos éticos, de transparencia y de gobernanza de la inteligencia artificial en educación. Este procedimiento se apoyó en una articulación interdisciplinaria desde la educación, la ética aplicada, la gobernanza digital y las políticas públicas, garantizando una construcción analítica rigurosa y coherente con los objetivos del estudio.

Tabla 1

Estructura categorial para el análisis de la ética, transparencia y gobernanza de la IA en decisiones educativas

Categoría analítica	Subcategoría	Criterio de análisis	Propósito analítico
Ética de la inteligencia artificial	Principios éticos	Identificación de valores orientadores como dignidad humana, justicia, equidad, no discriminación y responsabilidad en el uso de sistemas automatizados en educación	Examinar cómo los estudios abordan las implicaciones éticas de la IA en decisiones educativas
Ética de la inteligencia artificial	Impacto en derechos educativos	Análisis de referencias a derechos de estudiantes y docentes, tales como igualdad de oportunidades, trato justo y protección frente a decisiones automatizadas	Reconocer preocupaciones éticas vinculadas a efectos sociales y educativos
Transparencia algorítmica	Explicabilidad	Presencia de criterios que permitan comprender el	Analizar el grado de claridad y

Categoría analítica	Subcategoría	Criterio de análisis	Propósito analítico
		funcionamiento y los resultados de los sistemas de IA utilizados en contextos educativos	comprensibilidad atribuido a los sistemas automatizados
Transparencia algorítmica	Acceso a la información	Menciones a la disponibilidad de información sobre datos, criterios y procesos de decisión algorítmica	Evaluar la apertura informativa en torno a la toma de decisiones automatizadas
Gobernanza de la IA	Rendición de cuentas	Identificación de responsabilidades institucionales y mecanismos de supervisión sobre el uso de IA en educación	Examinar cómo se distribuye la responsabilidad frente a decisiones automatizadas
Gobernanza de la IA	Regulación y políticas	Referencias a marcos normativos, lineamientos institucionales y orientaciones de política pública	Analizar el tratamiento regulatorio del uso de IA en sistemas educativos
Toma de decisiones educativas	Control humano	Presencia de criterios que destaquen la intervención, supervisión o validación humana en decisiones apoyadas por IA	Identificar el rol atribuido a docentes y autoridades educativas
Toma de decisiones educativas	Riesgos y limitaciones	Identificación de sesgos, errores, exclusiones o efectos no deseados asociados al uso de IA	Reconocer alertas y vacíos señalados en los documentos analizados

Nota. La estructura presentada fue utilizada como guía para organizar y examinar la información contenida en los documentos analizados. Las categorías, subcategorías y criterios permitieron orientar el proceso de lectura, comparación e interpretación de los aportes identificados, facilitando una comprensión sistemática de los enfoques relacionados con la ética, la transparencia y la gobernanza de la inteligencia artificial en la toma de decisiones educativas.

3. Resultados

El análisis de los documentos considerados permitió identificar que la gobernanza de la inteligencia artificial en el ámbito educativo se caracterizó por una estructura fragmentada, en la que principios éticos, disposiciones normativas y orientaciones institucionales coexistieron sin una articulación clara. Esta falta de coherencia dificultó la consolidación de criterios compartidos para orientar la toma de decisiones educativas apoyadas por sistemas automatizados y dio lugar a prácticas heterogéneas entre instituciones. Asimismo, se evidenció que la debilidad de los marcos de gobernanza incidió negativamente en la confianza de la comunidad educativa, especialmente cuando las decisiones algorítmicas afectaron trayectorias académicas y procesos de evaluación.

El abordaje del contenido permitió reconocer la presencia reiterada de principios orientadores asociados al bienestar humano, la prevención de daños, el respeto a la autonomía y la búsqueda de condiciones de justicia en el uso de la inteligencia artificial en educación. No obstante, dichos principios se presentaron de forma general, con escasa contextualización pedagógica e institucional, lo que limitó su aplicación en situaciones educativas concretas. Esta distancia entre la formulación normativa y su puesta en práctica constituyó una dificultad persistente en la toma de decisiones mediadas por algoritmos.

Se observó que la incorporación de la inteligencia artificial en educación estuvo estrechamente vinculada a discursos centrados en la eficiencia, la optimización de procesos y la mejora del



rendimiento institucional. Este énfasis técnico tendió a desplazar consideraciones éticas y sociales, reduciendo la atención sobre aspectos relacionados con la equidad, la inclusión y la justicia educativa. Como resultado, se identificaron riesgos asociados a la reproducción de desigualdades estructurales mediante sistemas automatizados que no consideraron suficientemente la diversidad de contextos educativos.

El análisis evidenció una marcada concentración de trabajos orientados al desarrollo de modelos técnicos y herramientas algorítmicas, frente a una menor presencia de aportes enfocados en la ética, la transparencia y la gobernanza. Esta distribución desigual de enfoques limitó la comprensión integral del impacto de la inteligencia artificial en educación y restringió la construcción de marcos interpretativos que integraran de manera equilibrada dimensiones técnicas, pedagógicas y sociales.

El uso de sistemas algorítmicos en procesos de gestión, evaluación y seguimiento educativo transformó progresivamente las formas tradicionales de toma de decisiones institucionales. Estas transformaciones introdujeron lógicas de control basadas en datos y métricas que tendieron a desplazar el juicio profesional y la deliberación pedagógica. En este escenario, se identificó la necesidad de preservar espacios de intervención humana que permitieran contextualizar, validar y, en caso necesario, corregir las decisiones apoyadas por inteligencia artificial.

En relación con la transparencia, se constató que la explicabilidad de los sistemas de inteligencia artificial constituyó un elemento central para la legitimidad de las decisiones educativas. La dificultad para comprender los criterios, datos y procesos que sustentaron las decisiones algorítmicas limitó la rendición de cuentas institucional y redujo la posibilidad de cuestionamiento informado por parte de estudiantes y docentes, afectando la confianza en estos sistemas.

De manera recurrente, se identificaron preocupaciones vinculadas con la vigilancia, la recopilación intensiva de datos y la automatización de decisiones pedagógicas. Estas prácticas generaron tensiones en torno al rol profesional del docente, particularmente cuando la intervención humana se vio restringida en procesos educativos clave. La reducción de la autonomía docente se asoció con un posible empobrecimiento de la experiencia educativa y con una redefinición limitada de la práctica pedagógica.

También se evidenció que, si bien existieron orientaciones éticas y lineamientos de alcance internacional, su incorporación en contextos educativos presentó dificultades significativas. La ausencia de mecanismos claros de implementación, seguimiento y evaluación limitó su efectividad como guías para la toma de decisiones institucionales, lo que redujo su impacto en las prácticas educativas mediadas por inteligencia artificial.

La legitimidad de las decisiones educativas apoyadas por inteligencia artificial se encontró estrechamente vinculada con la existencia de procesos comunicativos claros y comprensibles. La posibilidad de justificar las decisiones, promover el diálogo institucional e incorporar la participación de los distintos actores educativos se configuró como un elemento clave para fortalecer la aceptación social y la coherencia ética del uso de estas tecnologías en los sistemas educativos.

Para organizar los hallazgos identificados y facilitar su comprensión integral, se elaboró una síntesis estructurada que permitió agrupar los principales patrones, tendencias y tensiones emergentes en torno a la ética, la transparencia y la gobernanza de la inteligencia artificial en la toma de decisiones educativas. Esta sistematización contribuyó a visualizar de manera comparativa las dimensiones analíticas abordadas, así como los vacíos y desafíos recurrentes



señalados en los documentos considerados, sirviendo como apoyo interpretativo al desarrollo narrativo de los resultados.

Tabla 2

Síntesis de hallazgos sobre ética, transparencia y gobernanza de la IA en la toma de decisiones educativas

Dimensión analítica	Hallazgos predominantes	Tendencias observadas	Vacíos o tensiones identificadas	Implicaciones para la toma de decisiones educativas
Ética de la IA	Presencia reiterada de principios orientadores generales	Predominio de formulaciones normativas amplias	Baja contextualización pedagógica e institucional	Riesgo de aplicación superficial de principios éticos
Transparencia	Reconocimiento de la importancia de la explicabilidad	Énfasis declarativo más que operativo	Limitada claridad sobre funcionamiento algorítmico	Dificultades para la rendición de cuentas
Gobernanza	Referencias a políticas y lineamientos	Enfoques fragmentados y poco articulados	Ausencia de mecanismos de implementación y seguimiento	Debilitamiento de la confianza institucional
Control humano	Valoración de la intervención humana	Uso creciente de decisiones automatizadas	Reducción del juicio pedagógico	Tensión entre automatización y autonomía docente
Impacto educativo	Uso de IA para gestión y evaluación	Centralidad de métricas y datos	Escasa problematización de efectos sociales	Posible reproducción de desigualdades
Confianza y legitimidad	Reconocimiento de la necesidad de aceptación social	Importancia del diálogo institucional	Falta de espacios participativos	Resistencia de la comunidad educativa

Nota. La información presentada sintetizó los hallazgos derivados del análisis de los documentos considerados, lo que permitió identificar patrones, tendencias y tensiones recurrentes en torno a la ética, la transparencia y la gobernanza de la inteligencia artificial en la toma de decisiones educativas. Esta organización facilitó la comparación entre las dimensiones analíticas y apoyó la interpretación de los resultados expuestos.

4. Discusión

La gobernanza de la inteligencia artificial en el ámbito educativo se configura como un desafío de naturaleza estructural, en la medida en que exige la articulación coherente entre políticas públicas, regulaciones institucionales, principios éticos y prácticas pedagógicas. La ausencia de enfoques integrales y coordinados limita la capacidad de los sistemas educativos para orientar el uso responsable de estas tecnologías y debilita la confianza de la comunidad educativa en los procesos de toma de decisiones apoyados por sistemas automatizados. Esta situación pone de manifiesto la necesidad de fortalecer los marcos que orientan tanto la producción académica como el diseño de políticas educativas relacionadas con la inteligencia artificial.

En este escenario, la formulación de principios orientadores para el uso responsable de la inteligencia artificial adquiere una relevancia central ante su expansión en contextos educativos.



Dichos principios buscan asegurar que los sistemas automatizados promuevan el bienestar humano, prevengan daños, respeten la autonomía de las personas y contribuyan a condiciones de justicia. Sin embargo, en el ámbito educativo, donde las decisiones influyen directamente en los procesos formativos y en las trayectorias académicas, estos principios requieren ser interpretados y aplicados considerando las particularidades institucionales, pedagógicas y sociales que caracterizan a los sistemas educativos.

La incorporación de la inteligencia artificial en educación suele estar acompañada de discursos centrados en la eficiencia, la optimización de recursos y la mejora del rendimiento académico e institucional. Este énfasis técnico tiende a desplazar reflexiones de carácter ético y social, priorizando soluciones instrumentales sin problematizar de manera suficiente sus implicaciones educativas. Como consecuencia, se corre el riesgo de invisibilizar cuestiones relacionadas con la equidad, la inclusión y la justicia social, lo que puede contribuir a la reproducción de desigualdades ya existentes en los sistemas educativos.

El crecimiento sostenido de estudios y desarrollos vinculados con la inteligencia artificial aplicada a la educación refleja un interés significativo por el diseño de herramientas, modelos y soluciones algorítmicas. No obstante, este avance no se distribuye de forma equilibrada entre los distintos enfoques de análisis, lo que limita la comprensión integral de los efectos de estas tecnologías. La menor atención a dimensiones como la ética, la transparencia y la gobernanza reduce la capacidad del campo educativo para anticipar y gestionar los impactos sociales y pedagógicos de la automatización.

El uso de algoritmos en procesos de gestión, evaluación y seguimiento educativo transforma las formas tradicionales de toma de decisiones institucionales, introduciendo dinámicas de control basadas en datos, métricas e indicadores de desempeño. Estas transformaciones reconfiguran las relaciones de poder dentro de los sistemas educativos y tienden a desplazar el juicio profesional y la deliberación pedagógica. En este contexto, se vuelve necesario reflexionar sobre los mecanismos de supervisión y regulación que permitan equilibrar el uso de datos con criterios pedagógicos y humanos.

Las orientaciones de alcance internacional orientadas al uso responsable de la inteligencia artificial aportan marcos de referencia relevantes para la protección de derechos fundamentales y valores educativos. No obstante, la existencia de estos lineamientos no garantiza su aplicación efectiva en contextos educativos específicos. La falta de mecanismos claros de implementación, seguimiento y evaluación limita su impacto en las prácticas institucionales y reduce su capacidad para orientar la toma de decisiones educativas mediadas por inteligencia artificial.

La expansión de sistemas automatizados en entornos educativos también genera preocupaciones relacionadas con la vigilancia, la recopilación intensiva de datos y la automatización de decisiones pedagógicas. Estas prácticas plantean tensiones en torno al rol profesional del docente, particularmente cuando la intervención humana se ve restringida en procesos clave como la evaluación y el acompañamiento del aprendizaje. La reducción del margen de acción docente puede afectar la autonomía profesional y empobrecer la experiencia educativa.

Asimismo, la transparencia se consolida como un elemento central para la legitimidad de las decisiones apoyadas por inteligencia artificial en educación. La posibilidad de comprender los criterios, datos y procesos que sustentan las decisiones algorítmicas resulta fundamental para garantizar la rendición de cuentas y permitir el cuestionamiento informado por parte de estudiantes y docentes. La opacidad de estos sistemas limita la supervisión institucional y debilita la confianza en su uso educativo.



La preservación del control humano en los procesos de toma de decisiones se presenta como un aspecto clave frente a la creciente automatización. La inteligencia artificial debe concebirse como un apoyo al juicio profesional y no como un sustituto del mismo, especialmente en decisiones de alto impacto educativo. Mantener la capacidad de intervención, evaluación y corrección por parte de las personas resulta esencial para asegurar prácticas educativas éticas y contextualizadas.

La integración de la inteligencia artificial en los sistemas educativos se desarrolla dentro de marcos institucionales y políticos más amplios que requieren regulaciones claras y participativas. La incorporación de la voz de la comunidad educativa en la definición de estos marcos contribuye a evitar la imposición unilateral de decisiones algorítmicas y favorece una gobernanza más democrática e inclusiva.

En este sentido, la legitimidad de las decisiones educativas mediadas por inteligencia artificial se vincula estrechamente con la existencia de procesos comunicativos claros, comprensibles y abiertos al diálogo. La posibilidad de justificar las decisiones, promover la deliberación institucional y construir consensos se configura como un elemento fundamental para la aceptación social y la coherencia ética del uso de estas tecnologías en educación.

5. Conclusión

La ética, la transparencia y la gobernanza de la inteligencia artificial en la toma de decisiones educativas se configuran como dimensiones centrales para comprender el impacto de los sistemas automatizados en los procesos formativos y en la gestión institucional. La inteligencia artificial actúa como un agente que influye de manera directa en las trayectorias académicas, en los procesos de evaluación y en la organización educativa, lo que exige un abordaje crítico que trascienda la dimensión técnica y considere sus implicaciones normativas, pedagógicas y sociales.

Los enfoques éticos identificados muestran una convergencia en torno a principios orientadores como la dignidad humana, la justicia, la equidad, la autonomía y la prevención de daños. No obstante, estos principios suelen presentarse de forma general, con escasa traducción en criterios operativos aplicables a contextos educativos específicos. Esta brecha entre la formulación ética y su aplicación práctica limita la capacidad de orientar de manera efectiva la toma de decisiones educativas apoyadas por inteligencia artificial y pone en evidencia la necesidad de avanzar hacia marcos más contextualizados.

En relación con la transparencia, se reconoce su papel fundamental para la legitimidad y la rendición de cuentas en los procesos decisionales mediados por sistemas algorítmicos. La dificultad para comprender el funcionamiento de los algoritmos, los criterios utilizados y el tratamiento de los datos restringe el cuestionamiento informado por parte de docentes y estudiantes, afectando la confianza institucional. A su vez, la preservación del control humano se presenta como un criterio clave para evitar la sustitución del juicio pedagógico y garantizar decisiones educativas responsables.

La gobernanza de la inteligencia artificial en educación evidencia la necesidad de marcos integrales que articulen políticas públicas, regulaciones institucionales y prácticas pedagógicas, incorporando la participación de la comunidad educativa. La ausencia de enfoques coordinados favorece implementaciones fragmentadas y amplifica problemáticas recurrentes como los sesgos, la vigilancia y la automatización excesiva de decisiones. En este sentido, los criterios identificados constituyen una base analítica relevante para orientar futuras investigaciones y el diseño de políticas educativas que sitúen la inteligencia artificial al servicio de los fines educativos y del desarrollo humano.



Referencias Bibliográficas

- Araya, P. C. (2021). Transparencia algorítmica ¿un problema normativo o tecnológico? *CUHSO (Temuco)*, <http://dx.doi.org/10.7770/cuhso-v31n2-art2196> .
- Arriagada, B. G., & Arias, J. (2024). ¿Cómo integrar la ética de la inteligencia artificial en el currículo? Análisis y recomendaciones desde el feminismo de la ciencia y de datos. *Revista de filosofía*, <http://dx.doi.org/10.5354/0718-4360.2024.74905> .
- Cepeda, C. E., Durán, N. Y., & Ocaña, O. A. (2025). Usos y perspectivas de la inteligencia artificial en la comunidad de profesores de la Universidad de Guayaquil. *Ñawi: arte diseño comunicación*, <https://doi.org/10.37785/nw.v9n1.a9> .
- González, A. M., Ordoñez, E. H., & González, S. J. (2024). ¿Tienen Derechos los Algoritmos? *Praxis & Saber*, <https://doi.org/10.19053/22160159.v15.n43.2024.16779> .
- Martínez, D. (2023). Moralidad y legitimidad en la teoría crítica de Jürgen Habermas: normatividad, reconocimiento y solidaridad. *Izquierdas*, <http://dx.doi.org/10.4067/S0718-50492022000100214> .
- Miao, F., & al, e. (2021). Inteligencia artificial y educación: guía para las personas a cargo de formular políticas. *UNESCO*, <https://unesdoc.unesco.org/ark:/48223/pf0000379376>.
- OECD. (2022). Uso estratégico y responsable de la inteligencia artificial en el sector público de América Latina y el Caribe. *OECD*, https://www.oecd.org/es/publications/uso-estrategico-y-responsable-de-la-inteligencia-artificial-en-el-sector-publico-de-america-latina-y-el-caribe_5b189cb4-es/full-report/component-4.html.
- OECD. (2025). Gobernar con la inteligencia artificial. *OECD*, https://www.oecd.org/es/publications/2025/06/governing-with-artificial-intelligence_398fa287/full-report/how-artificial-intelligence-is-accelerating-the-digital-government-journey_d9552dc7.html.
- Rondon, M. R., Pacotaípe, D. R., Alarcón, N. E., & Yopez, S. P. (2025). El Impacto de la Inteligencia Artificial en la Formación Docente. *Revista Tecnológica-Educativa Docentes 2.0*, <https://doi.org/10.37843/rted.v17i2.566> .
- Salvatierra, F. (2024). Construir el futuro: la IA en las políticas educativas. *UNESCO*, <https://unesdoc.unesco.org/ark:/48223/pf0000391418>.
- Sánchez, O. I. (2023). Inteligencia artificial en la educación superior: un análisis bibliométrico. *UNESCO*, <https://unesdoc.unesco.org/ark:/48223/pf0000390904>.
- Sánchez, T. M., González, d. R., & Suárez, P. M. (2025). Concepciones, estrategias y emociones docentes frente al uso de IA. *Revista de estudios y experiencias en educación*, <http://dx.doi.org/10.21703/rexe.v24i55.2991> .
- UNESCO. (2020). ODS y educación superior: una mirada desde la función de investigación. *UNESCO*, <https://unesdoc.unesco.org/ark:/48223/pf0000378615.locale=es>.



UNESCO. (2024). Ética de la inteligencia artificial. UNESCO, <https://www.unesco.org/es/artificial-intelligence/recommendation-ethics>.

UNESCO. (2025). Ecuador fortalece su gobernanza de la Inteligencia Artificial con la implementación de la RAM de la UNESCO. UNESCO, <https://www.unesco.org/es/articles/ecuador-fortalece-su-gobernanza-de-la-inteligencia-artificial-con-la-implementacion-de-la-ram-de-la>.

Valenzuela, F. L., Collantes, I. Z., & Durand, H. E. (2021). Sobre la gobernanza digital, política digital y educación. *Revista eleuthera*, <https://doi.org/10.17151/elev.2020.22.2.6>.

Conflicto de Intereses: Los autores declaran que no tienen conflictos de intereses relacionados con este estudio y que todos los procedimientos seguidos cumplen con los estándares éticos establecidos por la revista. Asimismo, confirman que este trabajo es inédito y no ha sido publicado, ni parcial ni totalmente, en ninguna otra publicación.