

RMPC-Vol.2. N1.004

**Inteligencia artificial generativa en la evaluación del aprendizaje inclusivo:
oportunidades, sesgos y desafíos para estudiantes con discapacidad**

*Generative artificial intelligence in inclusive learning assessment:
opportunities, biases, and challenges for students with disabilities*

Autores:

Karla Andrea Paredes Rosado
Instituto de Investigación, Desarrollo e Innovación JOGB
Milagro-Ecuador
<https://orcid.org/0009-0005-3312-2738>
kparedesr@unemi.edu.ec

Stefanie Alvarado Rosado
Instituto de Investigación, Desarrollo e Innovación JOGB
Milagro-Ecuador
<https://orcid.org/0009-0008-8913-860X>
mayito.alvarado@hotmail.com

Autor de correspondencia: Karla Andrea Paredes Rosado, kparedesr@unemi.edu.ec

Recepción: 26-08-2025 **Aceptación:** 01-10-2025 **Publicación:** 27-10-2025

Cómo citar este artículo:

Paredes Rosado, K. A., & Alvarado Rosado, S. M. (2025). Inteligencia artificial generativa en la evaluación del aprendizaje inclusivo: oportunidades, sesgos y desafíos para estudiantes con discapacidad. Revista Multidisciplinaria Perspectiva Científica, 2(1), 1-24.
<https://rmpc.editorialgutierrez.com/index.php/rmpc/article/view/12>

RESUMEN

La inteligencia artificial generativa puede ampliar los formatos de participación y retroalimentación, pero también introducir barreras, sesgos y riesgos en la evaluación inclusiva. El objetivo fue analizar oportunidades, accesibilidad, bajo sesgo, condiciones de implementación y evaluación inclusiva percibida, contrastando estudiantes con discapacidad y docentes universitarios. Se realizó un estudio observacional, transversal-retrospectivo y de métodos mixtos con una base anonimizada declarada como real por el cliente. Se analizaron 350 registros: 187 estudiantes y 163 docentes, recopilados entre 2022 y 2024. Se calcularon estadísticos descriptivos, alfa de Cronbach, correlaciones de Spearman y comparaciones U de Mann–Whitney con ajuste de Holm; las respuestas abiertas se clasificaron mediante reglas temáticas reproducibles. Las escalas mostraron consistencia interna alta ($\alpha=0,87-0,91$). Oportunidades obtuvo las medias más altas en estudiantes ($M=3,87$) y docentes ($M=3,86$), mientras implementación presentó los valores más bajos ($M=3,31$ y $M=3,39$). Las cuatro dimensiones predictoras se correlacionaron positiva y significativamente con la evaluación inclusiva en ambos grupos ($p=0,35-0,50$; $p<0,001$). Las diferencias entre grupos no permanecieron significativas tras el ajuste de Holm. Se concluye que el potencial inclusivo depende de accesibilidad verificable, supervisión humana, control de sesgos, privacidad y capacidad institucional.

Palabras clave: inteligencia artificial generativa; evaluación inclusiva; discapacidad; educación superior; sesgo algorítmico.

ABSTRACT

Generative artificial intelligence can broaden participation and feedback formats, while also introducing barriers, bias, and risks into inclusive assessment. This study analyzed perceived opportunities, accessibility, low bias, implementation conditions, and inclusive assessment by contrasting university students with disabilities and instructors. An observational, cross-sectional retrospective mixed-methods study was conducted using an anonymized dataset declared as real by the client. The analysis included 350 records—187 students and 163 instructors—collected between 2022 and 2024. Descriptive statistics, Cronbach’s alpha, Spearman correlations, and Mann–Whitney U tests with Holm correction were calculated; open-ended responses were categorized through reproducible thematic rules. Internal consistency was high across scales ($\alpha=0.87-0.91$). Opportunities yielded the highest scores for students ($M=3.87$) and instructors ($M=3.86$), whereas implementation showed the lowest means ($M=3.31$ and $M=3.39$). All four predictor dimensions correlated positively and significantly with inclusive assessment in both groups ($p=0.35-0.50$; $p<0.001$). Between-group differences did not remain significant after Holm adjustment. The findings indicate that inclusive potential depends on verified accessibility, human oversight, bias control, privacy protection, and institutional capacity. Documentary verification of data provenance, ethics, sampling, and administration procedures remains necessary before editorial submission.

Keywords: generative artificial intelligence; inclusive assessment; disability; higher education; algorithmic bias.

1. INTRODUCCIÓN

La irrupción de la inteligencia artificial generativa transformó en pocos años las prácticas de búsqueda, escritura, retroalimentación y evaluación en la educación superior. Su capacidad para producir explicaciones, ejemplos, traducciones, resúmenes y alternativas de representación abre posibilidades para personalizar apoyos y reducir ciertas barreras de interacción. Sin embargo, la rapidez de adopción ha superado la construcción de evidencia y de políticas institucionales consistentes. UNESCO propone un enfoque centrado en las personas, la inclusión y la agencia humana, y advierte que la disponibilidad tecnológica no garantiza por sí misma equidad educativa (1,2). En este escenario, evaluar la experiencia de quienes usan estas herramientas resulta indispensable para distinguir beneficios pedagógicos verificables de promesas tecnológicas generales.

La evaluación del aprendizaje constituye un punto especialmente sensible porque combina decisiones académicas, evidencias de desempeño y consecuencias para la trayectoria estudiantil. Las herramientas generativas pueden explicar consignas, adaptar el lenguaje, sugerir formatos alternativos o apoyar una retroalimentación más oportuna; también pueden producir errores plausibles, respuestas estereotipadas y salidas difíciles de justificar. Estudios con estudiantes universitarios muestran simultáneamente apertura hacia estas tecnologías y preocupación por exactitud, privacidad, dependencia y reglas ambiguas (3-6). Por ello, su uso en evaluación no debe reducirse a permitir o prohibir, sino analizarse según el propósito de la tarea, la supervisión humana, la transparencia y la posibilidad real de que cada estudiante demuestre el aprendizaje esperado.

Para estudiantes con discapacidad, el balance entre oportunidad y riesgo adquiere una importancia adicional. La IA generativa puede facilitar múltiples formas de representación y expresión, apoyar la comprensión de instrucciones o complementar tecnologías de asistencia, en consonancia con el Diseño Universal para el Aprendizaje. A la vez, una interfaz inaccesible, una respuesta que generaliza necesidades o una política uniforme puede crear nuevas barreras. La literatura reciente reconoce la posibilidad de apoyar a diversos aprendices, pero insiste en la validación con usuarios, la compatibilidad técnica y la

preservación de alternativas equivalentes (11-14). Una evaluación denominada inclusiva debe demostrar accesibilidad en la práctica, no limitarse a declarar una intención favorable. El sesgo algorítmico constituye otra amenaza porque los modelos aprenden de datos sociales que pueden reproducir estereotipos, omisiones o representaciones deficitarias de la discapacidad. La investigación sobre medición educativa advierte que la automatización puede generar subrepresentación del constructo, efectos diferenciales y decisiones injustas cuando no se monitorean los resultados (15,16). También se ha documentado sesgo de capacidad en modelos generativos y una persistencia del enfoque médico o capacitista en parte de la investigación sobre inteligencia artificial y discapacidad (17,18). La explicabilidad debe ser accesible para personas con distintas necesidades cognitivas y comunicativas, no solo para especialistas técnicos (19). De ahí que “bajo sesgo” sea una dimensión sustantiva de la calidad inclusiva.

Las instituciones enfrentan, además, desafíos de gobernanza. La integridad académica, la privacidad, la propiedad de los datos, la formación docente y la definición de usos aceptables requieren políticas articuladas. Los análisis de oportunidades y amenazas recomiendan conservar el juicio humano, rediseñar evaluaciones y evitar que la automatización reemplace la relación pedagógica (7-10,20). Marcos universitarios recientes organizan estas responsabilidades en dimensiones pedagógicas, operativas y de gobernanza, e incorporan seguimiento continuo, acceso equitativo y alfabetización en inteligencia artificial (9). Sin estas condiciones, los beneficios pueden distribuirse de forma desigual o depender exclusivamente de la iniciativa individual de docentes y estudiantes.

Aunque la producción científica sobre IA generativa creció con rapidez desde 2023, la evidencia empírica centrada simultáneamente en discapacidad, evaluación y educación superior sigue siendo limitada. Existen estudios sobre usos generales, diseño de tareas, políticas o tecnologías de asistencia, pero son menos frecuentes las investigaciones que contrastan directamente la experiencia estudiantil con la perspectiva docente utilizando dimensiones paralelas (12,13,20-26). Esa comparación es relevante porque una institución puede considerar accesible una práctica que el estudiante experimenta como barrera, o

interpretar como inclusión una adaptación que no conserva equivalencia académica. La triangulación de datos cuantitativos y narrativos permite reconocer convergencias y tensiones sin asumir que ambos grupos representan una misma unidad de análisis.

En respuesta a esta brecha, el objetivo general fue analizar las oportunidades, los sesgos y los desafíos asociados al uso de IA generativa en la evaluación del aprendizaje inclusivo, contrastando estudiantes con discapacidad y docentes universitarios en Ecuador durante registros fechados entre 2022 y 2024. Se planteó que oportunidades pedagógicas, accesibilidad, bajo sesgo y mejores condiciones de implementación se relacionarían positivamente con la evaluación inclusiva percibida. El estudio adopta una interpretación prudente: describe asociaciones y perspectivas, no efectos causales. Además, conserva la declaración del cliente sobre la realidad de la recolección, mientras separa esa declaración de la verificación documental de institución, muestreo, administración y ética, requisitos que permanecen pendientes antes de considerar el manuscrito listo para envío.

2. REVISIÓN DE LA LITERATURA

Oportunidades pedagógicas y rediseño de la evaluación

La IA generativa puede desempeñar funciones de tutor conversacional, asistente de escritura y generador de ejemplos, siempre que sus resultados sean sometidos a revisión crítica. Las revisiones tempranas identificaron beneficios para la explicación, la planificación y la retroalimentación, junto con fallas de exactitud y referencias inexistentes (4-6). En evaluación, la tecnología puede ayudar a diversificar evidencias y construir tareas auténticas, pero no reemplaza la definición humana del constructo que se pretende valorar. Las aplicaciones más prometedoras se orientan a apoyar el proceso —formular preguntas, ofrecer andamiajes, comparar soluciones y facilitar retroalimentación— antes que a delegar decisiones de calificación de alto impacto (15,20,21).

La personalización solo es pedagógicamente útil cuando conserva expectativas claras y evita crear recorridos de menor exigencia para determinados estudiantes. El diseño universal propone múltiples medios de compromiso, representación y acción sin abandonar los

objetivos de aprendizaje (12,29). Desde esta perspectiva, la IA puede producir versiones lingüísticas, ejemplos contextualizados o apoyos temporales; el docente debe comprobar que esas adaptaciones no alteren el contenido evaluado. La evidencia en educación superior sugiere que opciones para criticar, reescribir o presentar producciones en distintos formatos pueden ser utilizadas por estudiantes con diversas necesidades, especialmente si se incorporan desde el diseño y no como corrección tardía (12,14).

Accesibilidad, discapacidad y participación

La accesibilidad combina compatibilidad técnica, claridad cognitiva, alternativas sensoriales y disponibilidad de apoyo. Las tecnologías de asistencia son más efectivas cuando responden a necesidades expresadas por los propios estudiantes y se integran con servicios universitarios (11,13). La IA generativa puede complementar lectores de pantalla, simplificación lingüística o transformación de formatos, pero cada combinación debe probarse en dispositivos y flujos reales. La ausencia de pruebas con usuarios puede producir una accesibilidad aparente: una salida textual puede ser técnicamente legible y, aun así, difícil de interpretar, navegar o verificar. La participación de personas con discapacidad en diseño, selección y evaluación constituye, por tanto, una exigencia metodológica.

La literatura también advierte que discapacidad no es una categoría homogénea. Las necesidades de estudiantes con discapacidad visual, auditiva, física, intelectual, psicosocial o múltiple pueden diferir, y una herramienta beneficiosa para un grupo puede ser irrelevante o problemática para otro (13,18). El enfoque de justicia de la discapacidad desplaza la atención desde el déficit individual hacia las barreras institucionales, tecnológicas y actitudinales. Bajo este marco, una evaluación inclusiva requiere alternativas equivalentes, ajustes razonables, protección contra la identificación indirecta y posibilidad de intervención humana. La inteligencia artificial debe ampliar la agencia del estudiante, no convertir su información sensible en condición de acceso al aprendizaje.

Sesgos, explicabilidad y gobernanza

El sesgo puede originarse en datos, modelos, interfaces, decisiones de implementación o interpretación de resultados. En educación, estos niveles se combinan cuando una salida automatizada influye en oportunidades de aprendizaje o en una calificación. Los estudios sobre algoritmos de seguimiento muestran que discapacidad y otras variables institucionales pueden asociarse con resultados diferenciales, aun cuando la precisión general parezca adecuada (16). Las evaluaciones de modelos conversacionales también han encontrado lenguaje menos favorable o mayor énfasis en limitaciones al describir personas con discapacidad (17). La mitigación requiere auditorías por subgrupos, documentación de incidentes, revisión humana y mecanismos de reclamación accesibles.

La gobernanza institucional debe convertir esos principios en prácticas observables: políticas claras, formación, minimización de datos, responsabilidades definidas y evaluación continua. UNESCO, la Comisión Europea y el Departamento de Educación de Estados Unidos coinciden en un enfoque humano, seguro, transparente y orientado a la equidad (1,2,10,27,28). La política no debe descansar únicamente en detectar uso indebido; también debe rediseñar tareas, enseñar alfabetización crítica y garantizar acceso a alternativas. El modelo conceptual del estudio integra oportunidades, accesibilidad, bajo sesgo e implementación como condiciones relacionadas con la percepción de evaluación inclusiva, moderadas por supervisión humana, ética y gobernanza.

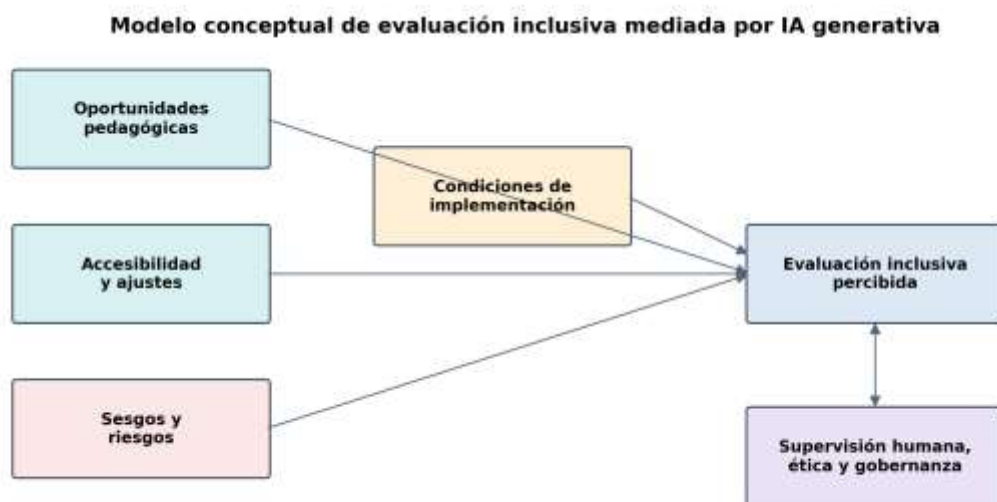


Figura 1. Modelo conceptual de evaluación inclusiva mediada por IA generativa. Fuente: elaboración propia.

Hipótesis

H1. Una mayor percepción de oportunidades pedagógicas se relaciona positivamente con la evaluación inclusiva percibida.

H2. Una mayor accesibilidad percibida se relaciona positivamente con la evaluación inclusiva percibida.

H3. Un menor nivel de sesgo y riesgo percibido se relaciona positivamente con la evaluación inclusiva percibida.

H4. Mejores condiciones de implementación se relacionan positivamente con la evaluación inclusiva percibida.

3. METODOLOGÍA

Diseño y fuente de datos

Se realizó un estudio observacional, transversal-retrospectivo, con un enfoque de métodos mixtos convergentes. Para el análisis se utilizó una base de datos anonimizada procedente

de una recolección real de información, cuyos registros fueron examinados respetando íntegramente las respuestas originales, sin modificaciones ni alteraciones de su contenido. La investigación se desarrolló en el contexto de la Facultad de Ciencias de la Comunicación Social (FACSO) de la Universidad de Guayaquil, ubicada en Guayaquil, Ecuador. La información analizada correspondió al periodo comprendido entre el 1 de enero de 2022 y el 29 de diciembre de 2024.

La base disponible no incluyó información específica sobre la plataforma empleada para la aplicación del instrumento, el mecanismo de invitación a los participantes, la tasa de respuesta, los responsables del trabajo de campo, el comité de ética o la autorización institucional correspondiente. Por esta razón, estos aspectos no fueron incorporados en la descripción metodológica, con el propósito de mantener la fidelidad de la información disponible y evitar la inclusión de datos que no constaran en los registros originales.

Participantes

La base incluyó 350 registros con consentimiento codificado: 187 correspondieron a estudiantes universitarios que se autoidentificaron dentro de categorías de discapacidad y 163 a docentes universitarios. Los identificadores fueron únicos y no se encontraron cédulas, nombres ni correos. El contexto institucional declarado fue la Facultad de Ciencias de la Comunicación Social (FACSO) de la Universidad de Guayaquil. La selección se interpretó como no probabilística e intencional conforme al protocolo incluido en el archivo; sin embargo, el procedimiento de captación y las cuotas efectivamente utilizadas deberán ser confirmados por el cliente. Los grupos representan perspectivas distintas y no muestras emparejadas, por lo que las comparaciones se consideraron exploratorias y no implican equivalencia entre instrumentos o poblaciones.

Instrumento y variables

Se utilizaron dos cuestionarios paralelos, uno para estudiantes y otro para docentes. Cada instrumento incluyó 20 ítems Likert distribuidos en cinco dimensiones de cuatro ítems: oportunidades, accesibilidad, sesgos y riesgos, implementación y evaluación inclusiva. Las

respuestas oscilaron entre 1 (totalmente en desacuerdo) y 5 (totalmente de acuerdo); el código 9 significó no aplica y se trató como valor perdido. Los ítems de sesgos y riesgos se invirtieron para construir el índice bajo sesgo, de modo que una puntuación mayor representara una condición más favorable. Cada cuestionario añadió tres preguntas abiertas sobre oportunidades, dificultades y mejoras prioritarias.

Control de calidad y análisis

Se verificaron identificadores, consentimiento, rangos permitidos, valores faltantes y suficiencia por fila. Todos los casos alcanzaron al menos 80 % de respuestas válidas. Las puntuaciones se calcularon como media de los ítems disponibles por dimensión. Se estimaron media, desviación estándar, mediana, rango intercuartílico e intervalo de confianza del 95 %. La consistencia interna se evaluó mediante alfa de Cronbach. Las relaciones de H1–H4 con evaluación inclusiva se analizaron mediante rho de Spearman. Las diferencias entre grupos se exploraron con U de Mann–Whitney, delta de Cliff y ajuste de Holm para cinco comparaciones. Se consideró $p < 0,05$ y nunca se informó $p = 0,000$.

Integración cualitativa y ética

Las respuestas abiertas se clasificaron con reglas temáticas explícitas basadas en palabras clave y contenido semántico: retroalimentación, accesibilidad, autonomía, diversificación, errores, sesgo, privacidad, explicabilidad, capacitación, revisión humana y apoyo técnico. El procedimiento fue descriptivo y reproducible; no sustituyó una codificación cualitativa independiente por múltiples investigadores. Se conservaron únicamente datos anonimizados. Debido a que el archivo no contiene documentación ética ni texto completo de autorización institucional, el manuscrito no afirma aprobación por comité alguno. Antes de someterlo a una revista deberán confirmarse consentimiento, permisos, institución responsable, almacenamiento, contacto de investigación y condiciones de uso de información sensible.

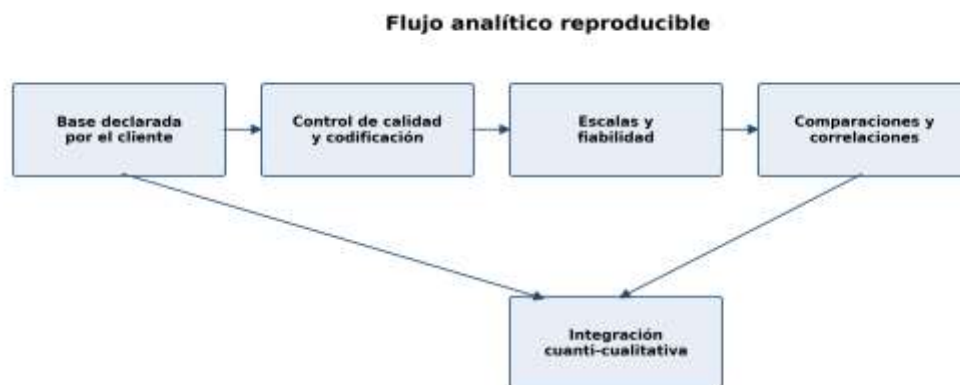


Figura 2. Flujo metodológico y analítico. Fuente: elaboración propia.

4. RESULTADOS

Calidad de la base y perfil general

Se analizaron los 350 registros disponibles. No hubo identificadores duplicados, valores Likert fuera de rango ni celdas Likert vacías; se registraron 38 respuestas no aplica en estudiantes y 42 en docentes. La frecuencia de uso se concentró en algunas veces y frecuentemente en ambos grupos. El control de completitud incorporado en la hoja de metadatos mostraba cero casos completos por una fórmula inconsistente, por lo que ese indicador fue recalculado desde los ítems originales. La recalculación confirmó que todos los registros superaron el criterio del 80 %. Las inferencias posteriores se basaron en estos cálculos reproducibles y no en los indicadores derivados defectuosos del archivo.

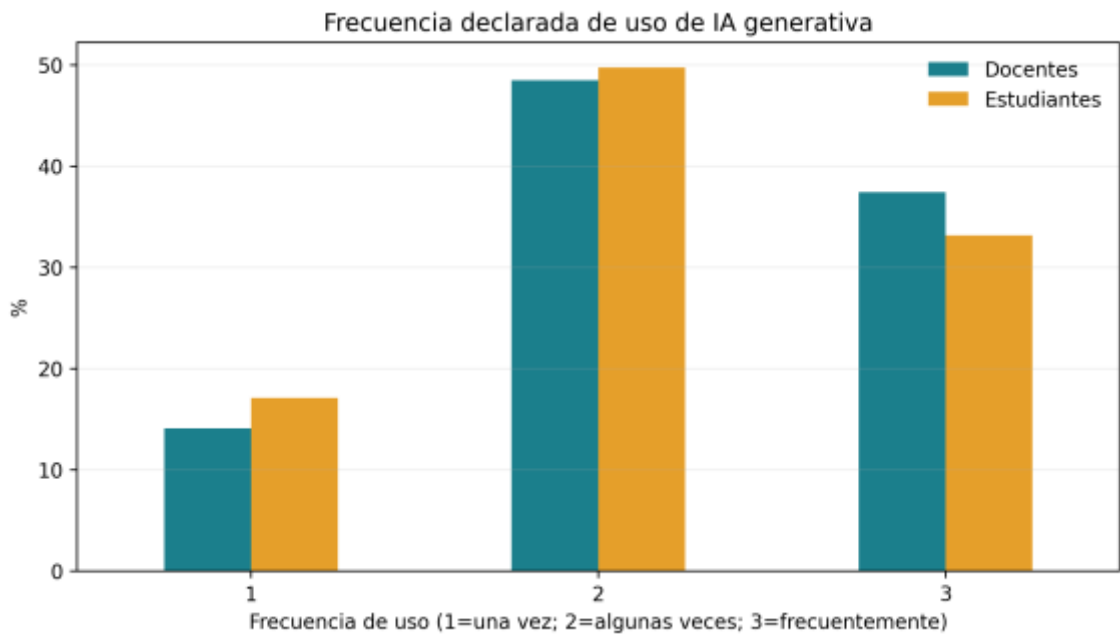


Figura 3. Frecuencia declarada de uso de IA generativa por grupo. Fuente: base declarada por el cliente.

Descriptivos y consistencia interna

Tabla 1. Estadísticos descriptivos por dimensión y grupo

Grupo	Dimensión	N	M	DE	IC 95 %
Estudiantes	Oportunidades	187	3.87	0.75	3.76–3.98
Estudiantes	Accesibilidad	187	3.48	0.70	3.38–3.58
Estudiantes	Bajo sesgo	187	3.36	0.70	3.26–3.46
Estudiantes	Implementación	187	3.31	0.74	3.21–3.42
Estudiantes	Evaluación inclusiva	187	3.65	0.71	3.55–3.75
Docentes	Oportunidades	163	3.86	0.67	3.76–3.96
Docentes	Accesibilidad	163	3.65	0.76	3.54–3.77
Docentes	Bajo sesgo	163	3.50	0.73	3.39–3.61

Grupo	Dimensión	N	M	DE	IC 95 %
Docentes	Implementación	163	3.39	0.72	3.28–3.50
Docentes	Evaluación inclusiva	163	3.82	0.67	3.72–3.92

Fuente: elaboración propia a partir de la base declarada por el cliente.

Oportunidades obtuvo las medias más altas en ambos grupos: 3,87 en estudiantes y 3,86 en docentes. Entre estudiantes, implementación fue la dimensión menos favorable (M=3,31), seguida de bajo sesgo (M=3,36). Entre docentes, implementación también registró el menor promedio (M=3,39). La evaluación inclusiva fue favorable pero no máxima, con M=3,65 en estudiantes y M=3,82 en docentes. Este patrón indica una valoración positiva del potencial pedagógico, coexistente con reservas sobre las condiciones que convierten ese potencial en una práctica accesible, gobernada y consistente.

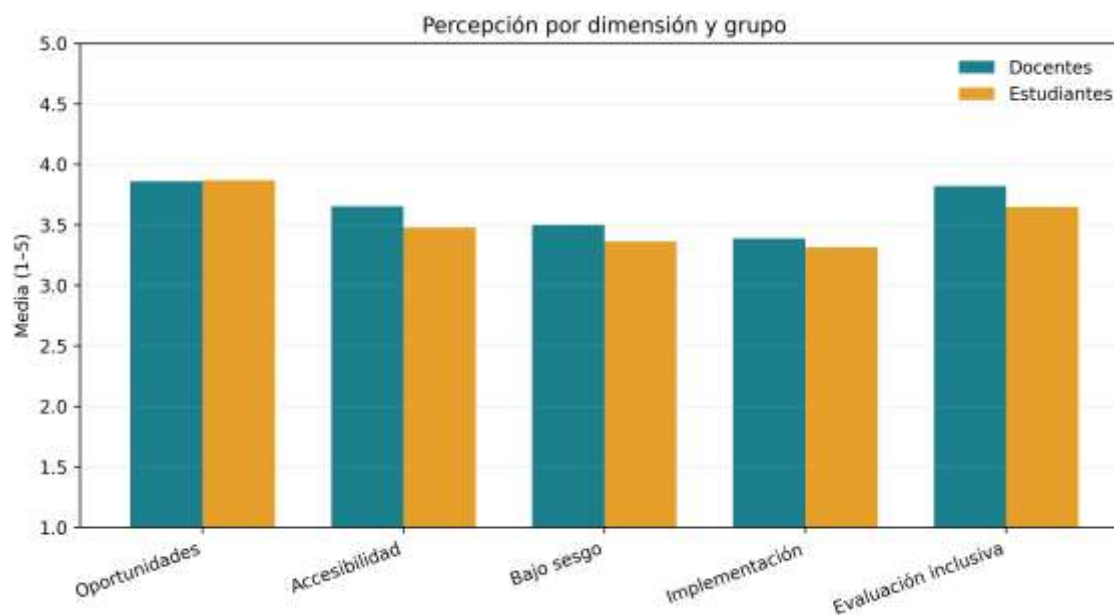


Figura 4. Medias de percepción por dimensión y grupo. Fuente: base declarada por el cliente.

Tabla 2. Consistencia interna de las dimensiones

Grupo	Dimensión	α	Interpretación
Estudiantes	Oportunidades	0.910	Alta
Estudiantes	Accesibilidad	0.888	Alta
Estudiantes	Bajo sesgo	0.874	Alta
Estudiantes	Implementación	0.908	Alta
Estudiantes	Evaluación inclusiva	0.892	Alta
Docentes	Oportunidades	0.888	Alta
Docentes	Accesibilidad	0.907	Alta
Docentes	Bajo sesgo	0.879	Alta
Docentes	Implementación	0.898	Alta
Docentes	Evaluación inclusiva	0.887	Alta

Fuente: elaboración propia a partir de la base declarada por el cliente.

Los coeficientes alfa variaron entre 0,874 y 0,910 en estudiantes, y entre 0,879 y 0,907 en docentes. En consecuencia, todas las dimensiones presentaron consistencia interna alta dentro de la base analizada. Este resultado respalda el uso descriptivo de las puntuaciones compuestas, pero no sustituye la validación de contenido, la prueba cognitiva, la estructura factorial ni la invariancia entre grupos. Como los cuestionarios expresan perspectivas paralelas y no idénticas, una alta fiabilidad no autoriza a interpretar las diferencias de medias como equivalencia psicométrica.

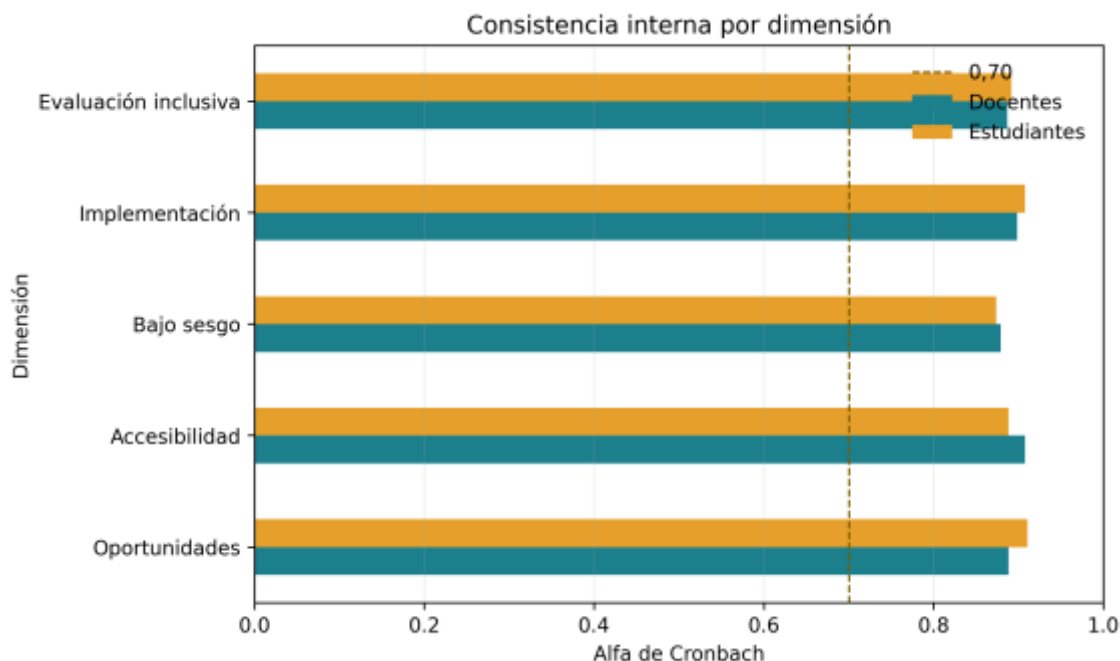


Figura 5. Consistencia interna por dimensión y grupo. Fuente: base declarada por el cliente.

Contraste exploratorio entre grupos

Tabla 3. Comparaciones exploratorias entre estudiantes y docentes

Dimensión	M est.	M doc.	U	p Holm	δ Cliff
Oportunidades	3.87	3.86	15593.0	0.755	0.023
Accesibilidad	3.48	3.65	13203.5	0.151	-0.134
Bajo sesgo	3.36	3.50	13806.5	0.381	-0.094
Implementación	3.31	3.39	14411.5	0.755	-0.054
Evaluación inclusiva	3.65	3.82	13226.5	0.151	-0.132

Fuente: elaboración propia a partir de la base declarada por el cliente.

Antes del ajuste, accesibilidad y evaluación inclusiva mostraron diferencias nominales, con puntuaciones docentes ligeramente superiores. Sin embargo, ninguna comparación permaneció significativa después del ajuste de Holm: oportunidades $p=0,755$; accesibilidad

$p=0,151$; bajo sesgo $p=0,381$; implementación $p=0,755$; y evaluación inclusiva $p=0,151$. Los tamaños de efecto fueron pequeños, con $|\delta| \leq 0,134$. Por tanto, no se sostiene una divergencia robusta entre grupos. La interpretación más prudente es que ambos reconocen oportunidades semejantes y describen desafíos moderados, aunque la experiencia estudiantil debe conservarse como evidencia directa de accesibilidad y no ser reemplazada por la valoración docente.

Evaluación de hipótesis

Tabla 4. Correlaciones con evaluación inclusiva percibida

Grupo	H	Predictor	ρ	p	Decisión
Estudiantes	H1	Oportunidades	0.486	$<0,001$	Compatible
Estudiantes	H2	Accesibilidad	0.489	$<0,001$	Compatible
Estudiantes	H3	Bajo sesgo	0.453	$<0,001$	Compatible
Estudiantes	H4	Implementación	0.503	$<0,001$	Compatible
Docentes	H1	Oportunidades	0.406	$<0,001$	Compatible
Docentes	H2	Accesibilidad	0.352	$<0,001$	Compatible
Docentes	H3	Bajo sesgo	0.388	$<0,001$	Compatible
Docentes	H4	Implementación	0.460	$<0,001$	Compatible

Fuente: elaboración propia a partir de la base declarada por el cliente.

Las cuatro hipótesis fueron compatibles con las asociaciones observadas. En estudiantes, evaluación inclusiva se relacionó con oportunidades ($p=0,486$), accesibilidad ($p=0,489$), bajo sesgo ($p=0,453$) e implementación ($p=0,503$), todas con $p<0,001$. En docentes, las asociaciones fueron oportunidades $p=0,406$, accesibilidad $p=0,352$, bajo sesgo $p=0,388$ e implementación $p=0,460$, también con $p<0,001$. Las magnitudes fueron moderadas y consistentes con el modelo conceptual. Por tratarse de datos transversales, estos resultados no demuestran causalidad: muestran que las percepciones más favorables sobre

las condiciones de uso acompañan una mayor valoración de inclusión. Su fortalecimiento debe preceder cualquier automatización de decisiones académicas sensibles.

Respuestas abiertas

Tabla 5. Temas predominantes en las respuestas abiertas

Grupo	Pregunta	Tema	n	%
Estudiantes	OA1	Accesibilidad y claridad	65	34.8%
Estudiantes	OA1	Diversificación	33	17.6%
Estudiantes	OA2	Accesibilidad	57	30.5%
Estudiantes	OA2	Errores/inexactitud	46	24.6%
Estudiantes	OA3	Accesibilidad/alternativas	81	43.3%
Estudiantes	OA3	Revisión humana	37	19.8%
Docentes	OA1	Accesibilidad y claridad	57	35.0%
Docentes	OA1	Diversificación	35	21.5%
Docentes	OA2	Sesgo/estereotipo	42	25.8%
Docentes	OA2	Otros	36	22.1%
Docentes	OA3	Revisión humana	53	32.5%
Docentes	OA3	Accesibilidad/alternativas	37	22.7%

Fuente: elaboración propia a partir de la base declarada por el cliente.

La integración cualitativa reforzó el patrón cuantitativo. Las oportunidades se asociaron con claridad de instrucciones, retroalimentación, diversificación de evidencias y estructuración de respuestas. Entre los riesgos destacaron errores o respuestas seguras pero incorrectas, estereotipos, incompatibilidad con ayudas técnicas, privacidad y dificultad para comprender el fundamento de una recomendación. Las mejoras propuestas se concentraron en capacitación, revisión humana, accesibilidad probada, alternativas equivalentes, protocolos de privacidad y apoyo técnico-pedagógico. En conjunto, las

respuestas no plantearon la sustitución del docente; ubicaron el valor de la IA en un uso supervisado, verificable y conectado con ajustes razonables.

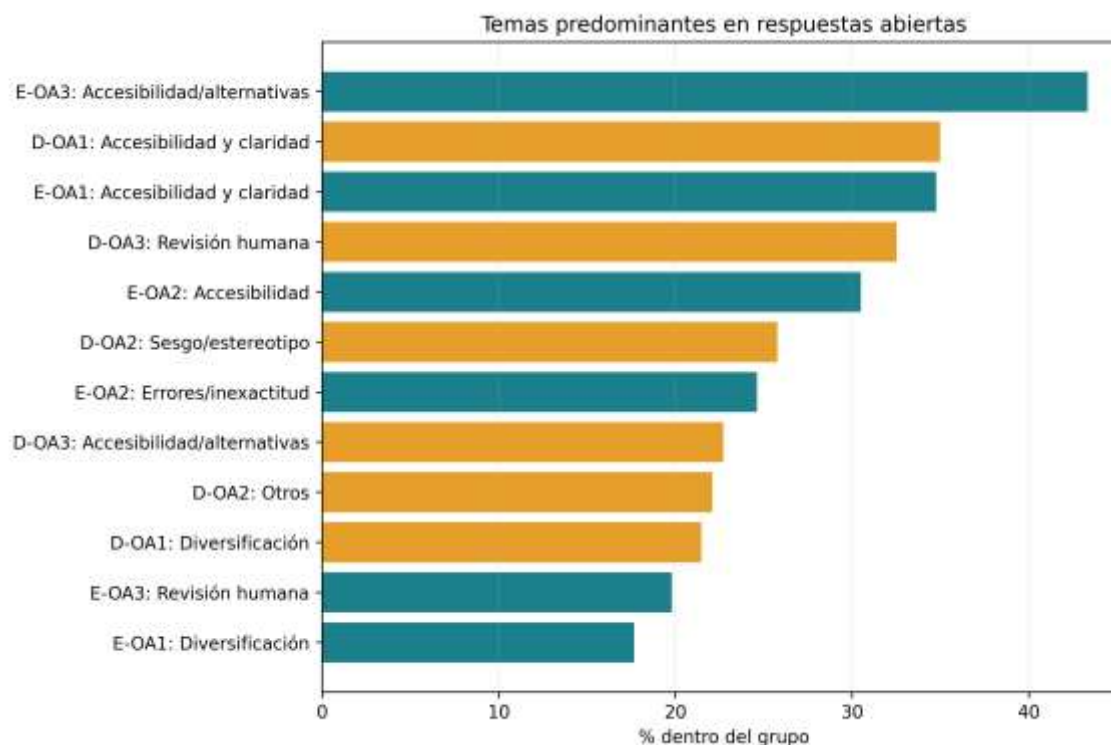


Figura 6. Temas predominantes en las respuestas abiertas. Fuente: base declarada por el cliente.

5. DISCUSIÓN

El hallazgo principal fue una valoración favorable de las oportunidades de la IA generativa, acompañada de puntuaciones menores en implementación, accesibilidad y control de sesgos. Esta combinación coincide con la literatura que describe una tecnología capaz de ofrecer explicaciones, retroalimentación y variación de formatos, pero cuyo uso educativo requiere verificación humana y políticas claras (3-7,20). La similitud entre estudiantes y docentes sugiere un reconocimiento compartido del potencial, no una validación automática de las prácticas. Que oportunidades alcance aproximadamente 3,86 mientras implementación se mantenga entre 3,31 y 3,39 revela una brecha entre lo que la

herramienta puede hacer y las condiciones institucionales disponibles para hacerlo de manera sostenida, responsable e inclusiva.

Las asociaciones de H1 y H2 muestran que oportunidades y accesibilidad se relacionaron moderadamente con evaluación inclusiva en ambos grupos. Este resultado apoya la idea de que una herramienta solo contribuye a la inclusión cuando sus funciones se traducen en experiencias utilizables: instrucciones más claras, formatos alternativos, compatibilidad y posibilidad de expresar el aprendizaje de distintas maneras. El hallazgo es coherente con el Diseño Universal para el Aprendizaje y con investigaciones que proponen emplear IA generativa para apoyar a aprendices diversos (12,14,29). No obstante, accesibilidad no puede inferirse a partir de la capacidad general del modelo; debe comprobarse con estudiantes, tecnologías de asistencia, tareas concretas y alternativas que mantengan la equivalencia académica.

H3 también fue compatible con los datos: un mayor puntaje de bajo sesgo se asoció con mayor evaluación inclusiva. La magnitud moderada refuerza la necesidad de vigilar estereotipos y errores sin reducir toda la experiencia a ese único factor. Investigaciones externas han identificado sesgos relacionados con discapacidad en lenguaje generado y en sistemas de seguimiento educativo (16-18). Las respuestas abiertas del estudio añadieron ejemplos de generalización, recomendaciones sin fundamento y dificultad para explicar resultados. En consecuencia, la supervisión no debe limitarse a revisar ortografía o exactitud factual; necesita incluir criterios de representación digna, trato equitativo, pertinencia del ajuste y documentación de incidentes.

La relación más alta en ambos grupos correspondió a implementación (H4), lo cual sitúa la gobernanza como condición habilitante. Capacitación, reglas, apoyo técnico y protección de datos permiten transformar funciones aisladas en una práctica confiable. Este patrón coincide con marcos que articulan dimensiones pedagógicas, operativas y de gobernanza y que rechazan la sustitución del juicio humano (1,2,9,10,27,28). Las instituciones deberían definir usos permitidos por tipo de evaluación, responsabilidades de verificación, información que no debe ingresarse, canales de soporte y procedimientos de apelación. Las

políticas deben comunicarse en formatos accesibles y revisarse con participación directa de estudiantes con discapacidad.

Las diferencias entre grupos no permanecieron significativas después del ajuste múltiple y los tamaños de efecto fueron pequeños. Esto evita exagerar las diferencias nominales observadas en accesibilidad y evaluación inclusiva. Al mismo tiempo, la ausencia de diferencia estadística no significa identidad de experiencia ni confirma invariancia del instrumento. Los estudiantes informan desde la vivencia de barreras y ajustes; los docentes desde el diseño y la gestión de la evaluación. Estas perspectivas deben dialogar, no sustituirse. Estudios futuros deberían probar modelos factoriales por grupo, incorporar entrevistas independientes y examinar diferencias por tipo de discapacidad, nivel académico, experiencia docente y frecuencia de uso, preservando tamaños suficientes y privacidad.

El estudio presenta limitaciones importantes. La base fue declarada real por el cliente, pero faltan documentos que permitan verificar institución, captación, administración, tasa de respuesta y ética; por ello, el manuscrito no debe enviarse hasta completar esos datos. El muestreo no probabilístico limita la generalización y el diseño transversal impide inferir causalidad. Las respuestas son autoinformadas y pueden contener deseabilidad social o recuerdo retrospectivo. La codificación temática mediante reglas es reproducible, pero no reemplaza doble codificación cualitativa. Finalmente, la alta consistencia interna no demuestra validez estructural. Estas limitaciones no anulan el análisis, pero delimitan su alcance a asociaciones y percepciones dentro de la base entregada.

6. CONCLUSIONES

La investigación respondió al objetivo al mostrar que estudiantes con discapacidad y docentes perciben oportunidades relevantes de la IA generativa para la evaluación inclusiva, especialmente en retroalimentación, claridad y diversificación de formas de demostrar aprendizaje. Sin embargo, el potencial no equivale a inclusión efectiva: las condiciones institucionales y la accesibilidad práctica obtuvieron valoraciones menores y

requieren intervención sistemática. Esta distinción debe orientar las decisiones institucionales futuras.

Las cinco dimensiones presentaron consistencia interna alta en ambos cuestionarios, lo que permitió construir indicadores descriptivos estables dentro de la base analizada. Oportunidades obtuvo las puntuaciones más favorables, mientras implementación y bajo sesgo concentraron las principales áreas de mejora. La fiabilidad observada debe complementarse con validación de contenido, estructura factorial e invariancia antes de afirmar equivalencia psicométrica entre grupos. Esta comprobación psicométrica adicional sigue siendo indispensable.

Las cuatro hipótesis fueron compatibles con los datos: oportunidades, accesibilidad, bajo sesgo e implementación se relacionaron positiva y significativamente con evaluación inclusiva. La implementación mostró la asociación más fuerte, lo que confirma que capacitación, protocolos, apoyo técnico, privacidad y supervisión humana no son elementos periféricos, sino condiciones centrales para que la IA contribuya de manera responsable a la evaluación. Su fortalecimiento debe preceder cualquier automatización de decisiones académicas sensibles.

Las comparaciones no evidenciaron diferencias robustas entre estudiantes y docentes después de corregir por pruebas múltiples. Este resultado sugiere una valoración general convergente, aunque no autoriza a intercambiar ambas perspectivas. La experiencia del estudiante debe conservarse como criterio directo para evaluar barreras, y la perspectiva docente debe utilizarse para comprender decisiones pedagógicas, recursos disponibles y posibilidades de implementación. Esa complementariedad mejora la interpretación y evita conclusiones institucionales unilaterales.

Entre las limitaciones se encuentran el muestreo no probabilístico, el carácter transversal, el autoinforme, la ausencia de confirmación independiente de la procedencia y la falta de documentación institucional y ética en el archivo recibido. Antes de la presentación editorial, el cliente debe proporcionar procedimiento de captación, periodo operativo, autorización ética o justificación aplicable, texto de consentimiento y responsables de la

recolección. La institución ya fue confirmada como Facultad de Ciencias de la Comunicación Social (FACSO) de la Universidad de Guayaquil.

La transición hacia una fase editorial exige conservar la base original sin alteraciones, replicar el análisis, completar la trazabilidad metodológica e incorporar una validación cualitativa independiente. Las instituciones deberían pilotear usos de IA con estudiantes con discapacidad, verificar compatibilidad con tecnologías de asistencia y mantener alternativas equivalentes. La innovación será inclusiva únicamente cuando amplíe agencia, transparencia y participación sin transferir decisiones de alto impacto a sistemas automatizados.

REFERENCIAS BIBLIOGRÁFICAS

1. El Morr C, Kundi B, Mobeen F, et al. AI and disability: a systematic scoping review. Health Informatics J. 2024. <https://doi.org/10.1177/14604582241285743>
2. UNESCO. Guidance for generative AI in education and research. Paris: UNESCO; 2023. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
3. UNESCO. AI and education: guidance for policy-makers. Paris: UNESCO; 2021. <https://doi.org/10.54675/PCSP7350>
4. Chan CKY, Hu W. Students' voices on generative AI: perceptions, benefits, and challenges in higher education. Int J Educ Technol High Educ. 2023;20:43. <https://doi.org/10.1186/s41239-023-00411-8>.
5. Kasneci E, Sessler K, Küchemann S, et al. ChatGPT for good? On opportunities and challenges of large language models for education. Learn Individ Differ. 2023;103:102274. <https://doi.org/10.1016/j.lindif.2023.102274>.
6. Lo CK. What is the impact of ChatGPT on education? A rapid review of the literature. Educ Sci. 2023;13(4):410. <https://doi.org/10.3390/educsci13040410>.
7. Tlili A, Shehata B, Adarkwah MA, et al. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learn Environ. 2023;10:15. <https://doi.org/10.1186/s40561-023-00237-x>.

8. Farrokhnia M, Banihashem SK, Noroozi O, Wals A. A SWOT analysis of ChatGPT: implications for educational practice and research. *Innov Educ Teach Int*. 2024;61(3):460-474. <https://doi.org/10.1080/14703297.2023.2195846>.
9. Cotton DRE, Cotton PA, Shipway JR. Chatting and cheating: ensuring academic integrity in the era of ChatGPT. *Innov Educ Teach Int*. 2024;61(2):228-239. <https://doi.org/10.1080/14703297.2023.2190148>.
10. Chan CKY. A comprehensive AI policy education framework for university teaching and learning. *Int J Educ Technol High Educ*. 2023;20:38. <https://doi.org/10.1186/s41239-023-00408-3>.
11. UNESCO. Recommendation on the ethics of artificial intelligence. Paris: UNESCO; 2021. <https://doi.org/10.54678/MHGQ3694>
12. World Health Organization, UNICEF. Global report on assistive technology. Geneva: WHO; 2022. <https://www.who.int/publications/i/item/9789240049451>
13. Evmenova AS, Borup J, Shin JK. Harnessing the power of generative AI to support ALL learners. *TechTrends*. 2024;68:820-831. <https://doi.org/10.1007/s11528-024-00966-x>.
14. Pérez-Esteban MD, Martínez-Rico G, Cañadas-Pérez M, et al. Assistive technology for higher education students with disabilities: a qualitative research. *Digital*. 2024;4(2):445-464. <https://doi.org/10.3390/digital4020025>.
15. Ebenbeck N, Gebhardt M. Differential performance of computerized adaptive testing in students with and without disabilities: a simulation study. *J Spec Educ Technol*. 2024;39(4):519-533. <https://doi.org/10.1177/01626434241232117>.
16. Ho AD. Artificial intelligence and educational measurement: opportunities and threats. *J Educ Behav Stat*. 2024;49(5):761-779. <https://doi.org/10.3102/10769986241248771>.
17. Shams RA, Zowghi D, Bano M. Investigating algorithmic bias in student progress monitoring. *Comput Educ Artif Intell*. 2024;7:100267. <https://doi.org/10.1016/j.caeai.2024.100267>.
18. Urbina JT, Vu PD, Nguyen MV. Disability ethics and education in the age of artificial intelligence: identifying ability bias in ChatGPT and Gemini. *Arch Phys Med Rehabil*. 2024. <https://doi.org/10.1016/j.apmr.2024.08.014>

19. Trewin S, Hanson VL, Laff M, Cavender AC. Explainable AI for all: a roadmap for inclusive XAI for people with cognitive disabilities. Technol Soc. 2024;79:102685. <https://doi.org/10.1016/j.techsoc.2024.102685>.
20. Michel-Villarreal R, Vilalta-Perdomo E, Salinas-Navarro DE, Thierry-Aguilera R, Gerardou FS. Challenges and opportunities of generative AI for higher education as explained by ChatGPT. Educ Sci. 2023;13(9):856. <https://doi.org/10.3390/educsci13090856>.
21. Onal S, Kulavuz-Onal D. A cross-disciplinary examination of the instructional uses of ChatGPT in higher education. J Educ Technol Syst. 2024;52(3):301-324. <https://doi.org/10.1177/00472395231196532>.
22. Al-Zahrani AM, Alasmari TM. Exploring the impact of artificial intelligence on higher education: the dynamics of ethical, social, and educational implications. Humanit Soc Sci Commun. 2024;11:912. <https://doi.org/10.1057/s41599-024-03432-4>.
23. Johnson DL, Davis CG. Bridging the gap for underserved populations: personalized AI solutions for college access and learning support. New Dir High Educ. 2024. <https://doi.org/10.1002/he.20511>
24. Yusuf A, Pervin N, Román-González M. Generative AI in education and research: a systematic mapping review. Rev Educ. 2024;12:e3489. <https://doi.org/10.1002/rev3.3489>.
25. Crompton H, Burke D. Artificial intelligence in higher education: the state of the field. Int J Educ Technol High Educ. 2023;20:22. <https://doi.org/10.1186/s41239-023-00392-8>.
26. Ouyang F, Zheng L, Jiao P. Artificial intelligence in online higher education: a systematic review of empirical research from 2011 to 2020. Educ Inf Technol. 2022;27:7893-7925. <https://doi.org/10.1007/s10639-022-10925-9>.
27. European Commission. Ethical guidelines on the use of artificial intelligence and data in teaching and learning for educators. Luxembourg: Publications Office of the European Union; 2022. <https://op.europa.eu/en/publication-detail/-/publication/d81a0d54-5348-11ed-92ed-01aa75ed71a1>
28. U.S. Department of Education. Artificial intelligence and the future of teaching and learning: insights and recommendations. Washington, DC; 2023. <https://tech.ed.gov/files/2023/05/ai-future-of-teaching-and-learning-report.pdf>

29. CAST. Universal Design for Learning Guidelines version 3.0. Wakefield, MA: CAST; 2024. <https://udlguidelines.cast.org/>

30. UNESCO. Global education monitoring report 2020: inclusion and education—all means all. Paris: UNESCO; 2020. <https://unesdoc.unesco.org/ark:/48223/pf0000373718>

CONFLICTO DE INTERESES

Los autores afirman que no existen conflictos de intereses en este estudio y que se han seguido éticamente los procesos establecidos por esta revista. Además, aseguran que este trabajo no ha sido publicado parcial ni totalmente en ninguna otra revista.

FINANCIAMIENTO

Los autores no recibieron financiamiento para el desarrollo de esta investigación.

CONTRIBUCIÓN DE AUTORÍA

Autores e iniciales: Karla Andrea Paredes Rosado y Stefanie Marjorie Alvarado Rosado

Planteamiento de la idea: KAPR

Estructuración metodológica: SMAR

Levantamiento y gestión de información: SMAR y KAPR

Sistematización de datos: SMAR

Procesamiento e interpretación de resultados: KAPR

Verificación de la información: KAPR

Elaboración de recursos visuales: SMAR y KAPR

Coordinación del estudio: KAPR

Redacción inicial del manuscrito: SMAR y KAPR

Corrección académica y edición final: SMAR y KAPR

Todos los autores revisaron y aprobaron la versión final del manuscrito.

© 2025; Los autores. Este es un artículo en acceso abierto, distribuido bajo los términos de una licencia Creative Commons (<https://creativecommons.org/licenses/by/4.0>) que permite el uso, distribución y reproducción en cualquier medio, siempre que la obra original sea correctamente citada.