

Integridad académica y evaluación educativa en

la era de la inteligencia artificial generativa

Molina Quiroz Marlon Josue

Laica Anchatuña Jessica Consuelo

Alcocer Yunda Carlos Daniel

Santamaría Tipantasig Mayra Germania

Semanate Semanate Doris Elizabeth

Laica Herrera Mónica Yolanda

Integridad académica y evaluación educativa en la era de la inteligencia artificial generativa

Molina Quiroz Marlon Josue
Laica Anchatuña Jessica Consuelo
Alcocer Yunda Carlos Daniel
Santamaría Tipantasig Mayra Germania
Semanate Semanate Doris Elizabeth
Laica Herrera Mónica Yolanda

ISBN: 978-9942-53-221-3
Primera edición, 2026

© **Molina Quiroz Marlon Josue**

Médico Veterinario
Clínica Veterinaria Dog y Cat
Cotopaxi - Latacunga
moljosue44@gmail.com
ORCID: 0009-0006-9928-5211

Laica Anchatuña Jessica Consuelo

Magister en Educación Inicial
U. E. Dr. Trajano Naranjo Iturralde
Cotopaxi - Latacunga
jessyl_21@hotmail.com
ORCID: 0009-0002-4659-9971

Alcocer Yunda Carlos Daniel

MSc. Educación Básica
U.E. Vicente León
Cotopaxi - Latacunga
carlosd.alcocer88@gmail.com
ORCID: 0009-0000-8836-2468

Santamaría Tipantasig Mayra Germania

Magister en Educación, Mención Innovación y
Liderazgo Educativo
U.E. Vicente León
Cotopaxi - Latacunga
mayras@hotmail.com
ORCID: 0009-0002-5925-0824

Semanate Semanate Doris Elizabeth

Magister en Innovación en Educación
U.E. Vicente León
Cotopaxi - Latacunga
dsemanate@hotmail.com
ORCID: 0009-0006-2736-6178

Laica Herrera Mónica Yolanda

Master Universitario en Formación del
Profesorado de Educación Secundaria de
Ecuador, Especialidad Física y Química
U.E. Vicente León
Cotopaxi - Latacunga
monylaica@gmail.com
ORCID: 0009-0001-5234-971X

© **Editorial Grupo Compás, 2026**

Guayaquil, Ecuador
www.grupocompas.com
<http://repositorio.grupocompas.com>

ISBN: 978-9942-53-221-3

DOI: 10.31876/9789942532213

Distribución online

Acceso abierto



Cita

Molina, M., et. Al. (2026). Integridad académica y evaluación educativa en la era de la inteligencia artificial generativa. Editorial Grupo Compás.

Primera edición, 2026

Esta obra ha sido sometida a un proceso de evaluación bajo el sistema de arbitraje doble ciego (double-blind peer review), garantizando el anonimato tanto de los autores como de los evaluadores externos. El dictamen favorable certifica que el contenido cumple con los más altos estándares de rigor científico, calidad editorial y originalidad exigidos por la comunidad académica internacional para su indexación y reconocimiento científico.

Índice

INTRODUCCIÓN	4
Capítulo 1	9
Integridad académica en la era de la inteligencia artificial generativa..	9
Concepto y fundamentos de la integridad académica	10
Plagio, fraude, autoría y responsabilidad intelectual	12
Inteligencia artificial generativa y transformación de las prácticas académicas.....	15
Uso legítimo, uso inapropiado y delegación intelectual mediante IA..	19
Desafíos de la integridad académica en América Latina y Ecuador	23
Capítulo 2.....	28
Deshonestidad académica y detección del uso de inteligencia artificial	28
Nuevas manifestaciones de deshonestidad académica mediadas por IA	29
Generación, reformulación y producción automatizada de contenidos académicos	32
Autoría, originalidad y transparencia en el uso de IA generativa	36
Detectores de contenido generado por IA: posibilidades y limitaciones	39
Prevención, evidencia y respuesta ante posibles infracciones académicas.....	43
Capítulo 3.....	50
Evaluación educativa en contextos mediados por inteligencia artificial	50

Evaluación educativa y validez de las evidencias de aprendizaje.....	51
Limitaciones de los modelos tradicionales de evaluación ante la IA generativa	54
Evaluación del producto y evaluación del proceso de aprendizaje.....	58
Evaluación auténtica, pensamiento crítico y resolución de problemas	61
Estrategias de evaluación y uso responsable de IA en educación escolar y superior	64
Capítulo 4.....	71
Cultura institucional, políticas y uso responsable de la inteligencia artificial.....	71
Políticas educativas e institucionales para el uso de IA generativa	72
Alfabetización en inteligencia artificial de docentes y estudiantes.....	77
Transparencia, declaración y criterios para el uso académico de IA	80
Equidad, privacidad, sesgos y responsabilidad ética	82
Perspectivas y desafíos para América Latina y Ecuador.....	87
CONCLUSIONES	93
BIBLIOGRAFÍA	95

INTRODUCCIÓN

La expansión de la inteligencia artificial generativa ha introducido un cambio significativo en las condiciones bajo las cuales se produce, comunica y evalúa el conocimiento en los sistemas educativos. Desde la aparición pública de herramientas capaces de generar textos, imágenes, código, explicaciones, resúmenes y respuestas complejas a partir de instrucciones formuladas en lenguaje natural, estudiantes y docentes han incorporado progresivamente estos recursos a actividades que hasta hace pocos años dependían casi exclusivamente de la producción humana. Esta transformación no constituye únicamente una innovación tecnológica: plantea interrogantes pedagógicos y éticos acerca de la autoría, la responsabilidad intelectual, la demostración del aprendizaje y el significado mismo de realizar un trabajo académico propio. UNESCO ha señalado que las capacidades de la inteligencia artificial generativa obligan a reconsiderar qué, cómo y para qué se aprende, al tiempo que requieren marcos educativos orientados por principios humanos, éticos y pedagógicos, y no solamente por la disponibilidad de nuevas herramientas (Miao & Holmes, 2023).

En este escenario, la integridad académica adquiere una relevancia renovada. Tradicionalmente, su discusión ha estado asociada a conductas como el plagio, la copia en exámenes, la falsificación de información o la utilización de ayudas no autorizadas. Sin embargo, reducirla a la ausencia de fraude resulta insuficiente. El International Center for Academic Integrity (ICAI, 2021) la concibe como un compromiso con seis valores fundamentales: honestidad, confianza, equidad, respeto, responsabilidad y valentía. Esta perspectiva desplaza el análisis desde una lógica exclusivamente sancionadora hacia la construcción de culturas educativas en las que estudiantes, docentes e instituciones comparten responsabilidades respecto de la producción y evaluación del conocimiento.

La inteligencia artificial generativa tensiona precisamente esas fronteras. La utilización de una herramienta para corregir ortografía, generar ideas, explicar un concepto o proporcionar retroalimentación no implica necesariamente el mismo problema ético que solicitarle la elaboración completa de una actividad y presentarla como producción propia. Entre ambos extremos existe un amplio conjunto de prácticas cuya aceptabilidad

depende del propósito de la tarea, las reglas establecidas, el grado de intervención de la herramienta y la transparencia con la que se reconoce su utilización. Eke (2023) advierte que la inteligencia artificial generativa puede ofrecer oportunidades relevantes para la actividad académica, pero que determinadas formas de utilización pueden comprometer la integridad si sustituyen indebidamente la contribución intelectual esperada del estudiante. Esta distinción obliga a superar clasificaciones simplistas en las que todo uso de inteligencia artificial se interpreta como fraude o, por el contrario, se considera automáticamente legítimo por tratarse de una tecnología disponible.

La dificultad se vuelve especialmente evidente cuando la evaluación educativa depende de productos que pueden ser generados total o parcialmente por sistemas automatizados. Ensayos, informes, ejercicios de programación, respuestas a preguntas abiertas y otras tareas tradicionalmente empleadas para inferir conocimientos o capacidades pueden dejar de constituir evidencia suficiente del aprendizaje cuando no es posible determinar con claridad cómo fueron producidas. El problema, por tanto, no reside únicamente en que un estudiante pueda utilizar inteligencia artificial para obtener una respuesta, sino en que ciertas prácticas evaluativas pueden perder capacidad para demostrar de manera válida qué comprende, sabe hacer o puede explicar el propio estudiante. Investigaciones recientes han señalado precisamente que la expansión de la IA generativa cuestiona la sostenibilidad de algunos formatos tradicionales de evaluación y refuerza la necesidad de valorar procesos de aprendizaje, pensamiento de orden superior y aplicaciones auténticas, más que productos finales aislados (Kizilcec et al., 2024; Lodge et al., 2023).

Esta discusión afecta tanto a la educación superior como a los niveles escolares. En contextos universitarios, estudios realizados después de la expansión de ChatGPT muestran que los propios estudiantes utilizan herramientas generativas para apoyar la elaboración de actividades evaluativas y que no siempre existe coincidencia entre estudiantes y docentes respecto de qué usos constituyen una vulneración de la integridad académica (Gruenhagen et al., 2024). En el ámbito escolar, Waltzer, Cox y Heyman (2023) encontraron que docentes de educación secundaria identificaron correctamente textos generados por ChatGPT aproximadamente en el 70 % de los casos, mientras que los estudiantes alcanzaron un 62 % de precisión. Además, la confianza de los participantes

en su capacidad para distinguir los textos no predijo su desempeño real. Estos resultados ilustran un problema relevante: la identificación intuitiva del origen de un texto no proporciona una base suficientemente robusta para establecer por sí sola una acusación de deshonestidad académica.

La búsqueda de soluciones tecnológicas para este problema ha conducido al desarrollo de detectores de contenido generado por inteligencia artificial. Sin embargo, la evidencia disponible obliga a ser cautelosos respecto de su utilización. Weber-Wulff et al. (2023), tras evaluar diferentes sistemas de detección, documentaron falsos positivos, falsos negativos y vulnerabilidad frente a modificaciones del texto, concluyendo que estas herramientas no resultaban suficientemente fiables para utilizar sus resultados como evidencia de una conducta académica indebida. Elkhatal, Elsaid y Almeer (2023) encontraron igualmente inconsistencias al clasificar textos humanos y generados mediante diferentes versiones de modelos de lenguaje. Por ello, sustituir la valoración pedagógica por mecanismos automatizados de vigilancia no resuelve el problema de fondo y puede introducir nuevos riesgos asociados con la equidad, la transparencia y la presunción injustificada de fraude.

La aparición de estas tecnologías obliga también a revisar conceptos que anteriormente parecían relativamente estables, como originalidad y autoría. Luo (2024), tras analizar políticas institucionales sobre inteligencia artificial generativa y evaluación, observa que muchas respuestas universitarias se han concentrado en proteger la originalidad del trabajo estudiantil, aunque la propia noción de originalidad se vuelve más compleja cuando la producción intelectual está mediada por tecnologías capaces de sugerir, reformular y generar contenidos. En consecuencia, la pregunta relevante ya no es únicamente si una inteligencia artificial intervino en una tarea, sino qué parte del desempeño evaluado corresponde efectivamente al estudiante, qué asistencia estaba permitida, qué decisiones intelectuales tomó y con qué grado de transparencia utilizó recursos externos.

Estas cuestiones adquieren particular importancia en América Latina, donde la incorporación de inteligencia artificial generativa ocurre dentro de sistemas educativos caracterizados por diferencias importantes en acceso tecnológico, alfabetización digital, formación docente y desarrollo institucional. La literatura regional emergente evidencia que los problemas éticos y de integridad académica se encuentran entre los temas más

frecuentes en la investigación sobre inteligencia artificial generativa y docencia universitaria, aunque buena parte de la producción continúa concentrándose en respuestas inmediatas frente al plagio, el uso instrumental de las herramientas y la necesidad de capacitación docente (Blake, 2026). En paralelo, la discusión latinoamericana comienza a desplazarse hacia cuestiones de gobernanza, transparencia, responsabilidad y regulación institucional de estas tecnologías (Kanobel, 2026).

Ecuador no permanece al margen de esta transformación. Investigaciones recientes realizadas en instituciones ecuatorianas muestran la necesidad de distinguir entre el empleo legítimo de la inteligencia artificial como apoyo al aprendizaje y la delegación no declarada de actividades intelectuales que forman parte de una evaluación. Un estudio desarrollado en educación superior virtual encontró diferencias entre estudiantes y docentes respecto de la utilización de inteligencia artificial y la interpretación de sus implicaciones para la integridad académica (Carvajal et al., 2026). Asimismo, evidencia obtenida con estudiantes de universidades públicas ecuatorianas sugiere que los patrones de utilización de IA generativa no son homogéneos y que determinados usos se relacionan con mayores niveles declarados de delegación intelectual, lo que refuerza la conveniencia de evitar políticas indiferenciadas y diseñar estrategias de alfabetización y regulación adaptadas a los contextos educativos (Torres-Díaz, 2026).

Frente a este panorama, preservar la integridad académica no significa intentar reconstruir un sistema educativo previo a la inteligencia artificial ni prohibir indiscriminadamente tecnologías que ya forman parte del entorno académico y profesional. El desafío consiste en establecer criterios suficientemente claros para diferenciar asistencia, colaboración y sustitución; diseñar evaluaciones capaces de producir evidencia válida del aprendizaje; formar a estudiantes y docentes para utilizar estas herramientas críticamente; y desarrollar políticas institucionales transparentes que combinen innovación, responsabilidad y equidad. Desde esta perspectiva, la inteligencia artificial generativa no crea todos los problemas actuales de la evaluación educativa, pero sí hace visibles algunas de sus debilidades preexistentes y obliga a interrogarlas con mayor urgencia.

Este libro aborda esa problemática articulando la integridad académica con los procesos de evaluación en contextos escolares y universitarios. Los capítulos desarrollan, de manera progresiva, los fundamentos conceptuales de la integridad académica y su transformación ante la inteligencia artificial generativa; los nuevos desafíos de autoría, plagio, transparencia y responsabilidad; las limitaciones de los mecanismos de detección y control; y las posibilidades de rediseñar la evaluación mediante estrategias que permitan valorar procesos, razonamiento, desempeño y aprendizaje auténtico. Desde una perspectiva latinoamericana, con atención particular al contexto ecuatoriano, la obra busca contribuir a una discusión que trascienda la oposición entre prohibir o permitir la inteligencia artificial y sitúe el problema donde pedagógicamente corresponde: en las condiciones necesarias para que la evaluación continúe ofreciendo evidencia válida del aprendizaje y para que la tecnología sea utilizada sin erosionar los principios que sostienen la confianza en la educación.

Capítulo 1

Integridad académica en la era de la inteligencia artificial generativa

La incorporación de la inteligencia artificial generativa en los entornos educativos ha modificado las condiciones en las que estudiantes y docentes acceden a la información, producen contenidos, resuelven problemas y desarrollan actividades académicas. La particularidad de estas tecnologías no reside únicamente en su capacidad para automatizar tareas, sino en que pueden intervenir directamente en procesos intelectuales que tradicionalmente se han considerado evidencia del aprendizaje, como redactar, argumentar, sintetizar información, generar código, formular explicaciones, proponer soluciones o estructurar productos académicos complejos.

Esta transformación ha situado a la integridad académica en el centro de una discusión que excede las formas tradicionales de plagio y fraude. Determinar si una práctica es académicamente legítima ya no depende solamente de establecer si un contenido fue copiado de una fuente identificable. También requiere valorar qué parte del proceso intelectual fue realizada por el estudiante, qué participación tuvo una herramienta generativa, si dicha intervención estaba autorizada, si fue declarada y, especialmente, si el producto presentado representa de manera válida los conocimientos o capacidades que se esperaba demostrar.

El problema tampoco puede resolverse mediante una oposición simple entre permitir o prohibir la inteligencia artificial. La misma herramienta puede desempeñar funciones pedagógicamente legítimas en una actividad y resultar incompatible con los objetivos de otra. Una aplicación generativa puede utilizarse para solicitar explicaciones adicionales, generar ejemplos de práctica, contrastar argumentos o recibir comentarios preliminares sobre un texto. Sin embargo, puede también sustituir la producción intelectual que una tarea pretendía evaluar. Por esta razón, comprender la integridad académica en la era de la inteligencia artificial exige trasladar la atención desde la existencia de la tecnología hacia las condiciones bajo las cuales se utiliza y las responsabilidades que permanecen bajo control humano.

Concepto y fundamentos de la integridad académica

La integridad académica ha sido frecuentemente interpretada como el cumplimiento de normas destinadas a impedir el plagio, la copia o diferentes formas de fraude. Aunque estas conductas forman parte del problema, reducir la integridad a la ausencia de infracciones limita considerablemente su significado educativo. La integridad académica constituye un marco de valores y prácticas que sustenta la confianza en los procesos de enseñanza, aprendizaje, evaluación, investigación y producción de conocimiento.

El International Center for Academic Integrity define la integridad académica a partir de seis valores fundamentales: honestidad, confianza, equidad, respeto, responsabilidad y valentía. Estos valores no se circunscriben al comportamiento individual del estudiante, sino que comprometen a toda la comunidad académica y deben traducirse en principios de actuación dentro de las instituciones educativas (International Center for Academic Integrity [ICAI], 2021). Esta formulación resulta especialmente útil porque transforma la integridad académica de una categoría exclusivamente disciplinaria en una condición de funcionamiento de la educación.

La honestidad supone representar de manera veraz el origen y las características del trabajo realizado. La confianza permite otorgar credibilidad a los procesos educativos y a sus resultados. La equidad exige establecer expectativas comprensibles y aplicar criterios consistentes. El respeto comprende el reconocimiento del trabajo intelectual propio y ajeno. La responsabilidad implica asumir las decisiones adoptadas durante la actividad académica y sus consecuencias. Finalmente, la valentía se relaciona con la capacidad de sostener estos valores incluso cuando existen presiones o incentivos para abandonarlos (ICAI, 2021).

Esta aproximación permite diferenciar la integridad académica de la deshonestidad académica. La primera constituye una orientación positiva de la actividad educativa, mientras que la segunda describe conductas que vulneran las normas o principios establecidos. Una institución puede contar con procedimientos rigurosos para sancionar el plagio y, al mismo tiempo,

presentar debilidades importantes en su cultura de integridad si las expectativas no son claras, los criterios varían considerablemente entre docentes o los estudiantes reciben poca formación sobre las razones que fundamentan las normas.

Bretag (2016) plantea la integridad académica como un campo multidimensional en el que intervienen factores pedagógicos, culturales, institucionales y éticos. Esta perspectiva permite comprender que las conductas académicas no se producen en un vacío. Las características de las evaluaciones, la presión por las calificaciones, la claridad de las instrucciones, la formación sobre escritura académica, las políticas institucionales y las condiciones en las que se desarrolla el aprendizaje pueden influir en la manera en que los estudiantes enfrentan sus responsabilidades académicas.

La inteligencia artificial generativa amplifica la importancia de esta concepción. Los sistemas generativos contemporáneos pueden producir resultados que anteriormente requerían una participación intelectual humana mucho más evidente. UNESCO ha señalado que la expansión de estas tecnologías obliga a reconsiderar aspectos centrales de la enseñanza y el aprendizaje, manteniendo la agencia humana como principio rector de su utilización educativa (Miao & Holmes, 2023). El problema no consiste únicamente en que una máquina sea capaz de generar contenido, sino en que dicho contenido pueda confundirse con evidencia de conocimientos o capacidades que el estudiante no necesariamente ha desarrollado.

Esta cuestión se relaciona directamente con la confianza. Las calificaciones, certificados y títulos poseen valor académico y social porque se asume que representan determinados aprendizajes. Una institución que certifica el desempeño de un estudiante afirma implícitamente que esa persona posee ciertos conocimientos y competencias. Si los productos utilizados para realizar esa inferencia fueron elaborados mediante procesos que no permiten distinguir entre el desempeño del estudiante y el de una herramienta externa, la cuestión deja de afectar exclusivamente a la conducta individual y alcanza la credibilidad de la evaluación.

La integridad académica se vincula así con la correspondencia entre lo que un estudiante presenta y aquello que realmente sabe, comprende o puede hacer. Esta relación es especialmente importante ante herramientas que pueden producir textos coherentes, resolver determinados problemas o

generar respuestas aparentemente especializadas. Cuando una actividad tiene como finalidad valorar la capacidad individual para argumentar y esa argumentación es sustancialmente generada por una herramienta, existe una discrepancia entre el propósito de la actividad y el desempeño que el producto parece representar. Sin embargo, si la actividad consiste expresamente en evaluar la capacidad para analizar críticamente una respuesta producida por inteligencia artificial, utilizar la herramienta puede ser parte legítima del aprendizaje.

La legitimidad del uso tecnológico depende, por tanto, del propósito pedagógico y no exclusivamente de las características de la herramienta. Este principio evita dos posiciones extremas. La primera identifica cualquier intervención de inteligencia artificial con una infracción académica. La segunda considera que la disponibilidad generalizada de estas tecnologías convierte automáticamente su utilización en aceptable. Ninguna de las dos posiciones permite abordar adecuadamente la diversidad de prácticas que pueden producirse.

Una cultura de integridad adaptada a estos escenarios requiere que las instituciones y los docentes definan con mayor precisión qué se espera que realice el estudiante, qué recursos están permitidos y qué formas de asistencia deben declararse. La responsabilidad institucional consiste también en evitar zonas de ambigüedad en las que una misma conducta sea aceptada en una asignatura, sancionada en otra y nunca discutida explícitamente con los estudiantes. La integridad académica pasa así de constituir únicamente un mecanismo de control a formar parte del diseño pedagógico de las experiencias de aprendizaje.

Plagio, fraude, autoría y responsabilidad intelectual

La expansión de la inteligencia artificial generativa ha provocado que términos como plagio, fraude, autoría y uso indebido de IA sean utilizados en ocasiones de manera intercambiable. Esta equivalencia es conceptualmente problemática y puede producir políticas institucionales imprecisas.

El plagio implica, en términos generales, apropiarse o presentar como propias ideas, expresiones o producciones procedentes de otros sin realizar el reconocimiento correspondiente. La deshonestidad académica

constituye una categoría más amplia y puede comprender, además del plagio, utilización de recursos no permitidos, falsificación de datos, suplantación, colaboración indebida o contratación de terceros para realizar un trabajo. La vulneración de la integridad académica, por tanto, no se limita a aquello que puede clasificarse estrictamente como plagio.

Esta distinción adquiere especial relevancia ante la inteligencia artificial generativa. Si un estudiante solicita a un sistema que produzca íntegramente una respuesta y posteriormente la presenta como evidencia de su propio razonamiento, existe un problema evidente de integridad. Sin embargo, clasificar automáticamente la conducta como plagio puede resultar conceptualmente insuficiente cuando el contenido generado no corresponde directamente a la obra identificable de otro autor. La infracción puede radicar, en cambio, en una representación falsa del proceso de producción y de la autoría intelectual.

Eke (2023) sostiene que los sistemas generativos introducen riesgos para la integridad académica precisamente porque pueden utilizarse para producir contenidos que sustituyen actividades intelectuales que corresponderían al estudiante. El autor advierte, sin embargo, que las consecuencias dependen del modo en que se empleen estas tecnologías y de las respuestas desarrolladas por los distintos actores institucionales. De manera semejante, Cotton, Cotton y Shipway (2024) señalan que las posibilidades educativas de estas herramientas coexisten con riesgos relacionados con deshonestidad y plagio, lo que exige respuestas institucionales que no se reduzcan a la prohibición.

La cuestión de la autoría resulta central porque ser autor de un trabajo académico implica más que haber producido materialmente las palabras que aparecen en un documento. La autoría supone asumir responsabilidad por la selección de información, el razonamiento desarrollado, las fuentes utilizadas, los argumentos formulados y las conclusiones presentadas. En otras palabras, existe una dimensión de responsabilidad intelectual que no desaparece porque determinadas partes del proceso sean asistidas tecnológicamente.

Eaton (2023) propone el concepto de postplagiarism para describir un escenario en el que la interacción entre seres humanos y tecnologías avanzadas se vuelve cotidiana en la producción de conocimiento. La propuesta no pretende declarar obsoleto el concepto de integridad, sino

mostrar que las fronteras tradicionales entre producción humana y tecnológica se vuelven más complejas. Una de las consecuencias de esta transformación es que la responsabilidad humana debe mantenerse incluso cuando parte de la producción haya sido automatizada.

Esta idea permite distinguir entre asistencia tecnológica y transferencia de responsabilidad. Un estudiante puede utilizar una herramienta para revisar la claridad de una explicación, explorar posibles enfoques o identificar errores en un borrador. Sin embargo, continúa siendo responsable de verificar el contenido, determinar su pertinencia y responder por lo que finalmente presenta. Una herramienta generativa no puede asumir responsabilidad académica por una afirmación falsa, una referencia inexistente o un argumento conceptualmente incorrecto.

La capacidad de los modelos generativos para producir respuestas lingüísticamente plausibles refuerza esta exigencia. La fluidez de una respuesta no garantiza su exactitud y el estudiante no puede considerar a la herramienta como una fuente de autoridad equivalente a la evidencia científica o académica. Utilizar contenidos generados sin comprobarlos implica no solo un problema potencial de precisión, sino también una renuncia al juicio crítico que forma parte de la responsabilidad intelectual.

La transparencia constituye otro componente fundamental. Reconocer que se utilizó una herramienta no convierte automáticamente cualquier uso en legítimo, porque una actividad expresamente restringida sigue siendo incompatible con las instrucciones aunque el estudiante declare la intervención tecnológica. Sin embargo, la transparencia permite reconstruir de manera más adecuada el proceso de producción y diferenciar la asistencia declarada de la sustitución encubierta.

Luo (2024), al analizar políticas sobre inteligencia artificial generativa y evaluación en veinte universidades de reconocido posicionamiento internacional, encontró que la preocupación predominante se relacionaba con la originalidad del trabajo presentado por los estudiantes. Su análisis cuestiona que la originalidad pueda continuar interpretándose exclusivamente como una producción aislada de cualquier asistencia externa y propone reconsiderarla en contextos donde diferentes formas de mediación tecnológica participan de manera creciente en la producción académica.

Esto no implica abandonar la autoría como criterio. Por el contrario, exige definirla con mayor precisión. La pregunta relevante no consiste únicamente en determinar si hubo utilización de inteligencia artificial, sino en analizar qué aportó la herramienta y qué aportó el estudiante. También resulta necesario conocer si su intervención estaba permitida, si fue declarada, si el estudiante verificó la información generada y si puede explicar y defender el contenido presentado.

Estas preguntas permiten diferenciar prácticas que podrían parecer similares si únicamente se identificara la presencia de una herramienta. Un estudiante que utiliza inteligencia artificial para recibir comentarios sobre un borrador previamente elaborado conserva un nivel de participación intelectual diferente al de aquel que solicita la elaboración completa de la actividad. Del mismo modo, utilizar una herramienta para generar alternativas que posteriormente son analizadas, seleccionadas y modificadas no representa necesariamente el mismo nivel de delegación que aceptar sin revisión la primera respuesta obtenida.

Por ello, la discusión sobre autoría en la era de la inteligencia artificial requiere abandonar la clasificación simplista entre contenido humano y contenido generado por máquinas. En numerosos procesos contemporáneos existirá algún grado de mediación tecnológica. La cuestión central será determinar cuándo esa mediación permanece subordinada al trabajo intelectual del estudiante y cuándo comienza a sustituir precisamente aquello que la actividad pretendía desarrollar o evaluar.

Inteligencia artificial generativa y transformación de las prácticas académicas

La incorporación de inteligencia artificial en educación precede a la expansión reciente de los modelos generativos. Sistemas de recomendación, tutores inteligentes, traducción automática, reconocimiento de voz, analítica del aprendizaje y mecanismos automatizados de retroalimentación se han utilizado durante años. Sin embargo, la difusión de herramientas generativas accesibles mediante lenguaje natural modificó sustancialmente la escala y facilidad con que

estudiantes y docentes pueden recurrir a sistemas capaces de producir contenidos complejos.

La publicación masiva de herramientas conversacionales desde finales de 2022 aceleró esta transformación. UNESCO advirtió tempranamente que la velocidad de desarrollo y difusión de la inteligencia artificial generativa estaba superando la capacidad de adaptación de muchos sistemas regulatorios y educativos, generando un desfase entre las prácticas tecnológicas emergentes y las normas disponibles para orientarlas (Miao & Holmes, 2023).

Las aplicaciones académicas posibles son heterogéneas. Los estudiantes pueden solicitar explicaciones alternativas de un concepto, ejemplos adicionales, ejercicios de práctica, ayuda para organizar ideas, corrección lingüística, traducción, síntesis de información o retroalimentación preliminar. También pueden utilizar los mismos sistemas para generar una actividad completa, resolver un ejercicio sin comprender su procedimiento o elaborar un producto evaluado con una participación intelectual mínima.

Esta dualidad muestra por qué hablar genéricamente del uso de inteligencia artificial ofrece poca información sobre su valor educativo. La misma frecuencia de utilización puede corresponder a prácticas profundamente distintas. Un estudiante puede mantener una interacción intensiva con un sistema para formular preguntas, comprobar su comprensión y contrastar razonamientos, mientras otro puede emplearlo con la misma frecuencia para reducir su participación en las tareas. Medir exclusivamente cuánto se utiliza una tecnología no permite establecer si su integración favorece o perjudica el aprendizaje.

El estudio ecuatoriano de Buele et al. (2025) resulta ilustrativo. Los autores analizaron a 833 estudiantes universitarios y encontraron que las dimensiones cognitivas y afectivas se relacionaban con el comportamiento ético frente a la inteligencia artificial, mientras que la frecuencia de uso, considerada de manera aislada, no presentó una asociación estadísticamente significativa en el modelo evaluado. Aunque este estudio no permite generalizar los resultados a todo el sistema educativo ecuatoriano, aporta evidencia para cuestionar una interpretación según la cual una mayor frecuencia de uso equivaldría necesariamente a una conducta menos ética.

La transformación afecta también a la naturaleza de los productos académicos. El debate inicial se concentró especialmente en textos escritos debido a la popularización de los modelos conversacionales, pero las capacidades generativas abarcan programación, imágenes, presentaciones, audio, traducción, resolución de problemas y producción multimodal. Por ello, la discusión sobre integridad no puede restringirse a ensayos o trabajos escritos. En distintas disciplinas será necesario identificar cuáles son las competencias centrales y qué formas de automatización resultan compatibles con su desarrollo.

En educación escolar existe una consideración adicional relacionada con la etapa formativa del estudiante. Las capacidades de autonomía, evaluación crítica y reconocimiento de información incorrecta evolucionan progresivamente, por lo que las expectativas de uso responsable no pueden ser idénticas en todos los niveles educativos. El marco de competencias en inteligencia artificial para estudiantes de UNESCO propone abordar esta formación mediante doce competencias organizadas en cuatro dimensiones: una perspectiva centrada en el ser humano, ética de la inteligencia artificial, técnicas y aplicaciones de IA, y diseño de sistemas de IA. Estas competencias se estructuran en niveles progresivos de comprensión, aplicación y creación (Miao, Shiohira, & Lao, 2024).

El valor de esta propuesta reside en reconocer que la alfabetización en inteligencia artificial no equivale a dominar técnicamente una herramienta. También requiere comprender sus limitaciones, analizar sus implicaciones éticas, evaluar críticamente los resultados y mantener capacidad de decisión humana. Esta orientación resulta relevante para la integridad académica porque un estudiante difícilmente puede utilizar una herramienta de manera responsable si desconoce cómo valorar la confiabilidad de sus resultados o las consecuencias de delegar determinadas decisiones.

La disponibilidad de herramientas generativas tampoco elimina la necesidad de adquirir conocimientos disciplinares. Esta idea merece especial atención ante la posibilidad de interpretar que determinadas habilidades dejan de ser necesarias porque una tecnología puede ejecutarlas. Para evaluar críticamente una respuesta generada se requiere suficiente conocimiento para reconocer errores, omisiones y simplificaciones. Un estudiante que no comprende el contenido posee

menor capacidad para supervisar la producción automatizada y puede aceptar resultados incorrectos simplemente porque están redactados de manera convincente.

Por tanto, integrar inteligencia artificial no significa sustituir competencias fundamentales por habilidades para formular instrucciones. La alfabetización en IA debe construirse sobre conocimiento disciplinar, pensamiento crítico y capacidad de verificación. La herramienta puede ampliar determinadas capacidades humanas, pero su valor educativo depende de que el estudiante conserve una participación intelectual suficiente para comprender y controlar el proceso.

Las respuestas institucionales también han evolucionado. Moorhouse, Yeo y Wan (2023) analizaron las orientaciones de las cincuenta universidades mejor posicionadas en clasificaciones internacionales y encontraron que, en el momento de su estudio, menos de la mitad disponían de directrices públicas específicas sobre inteligencia artificial generativa y evaluación. Las orientaciones existentes se concentraban principalmente en integridad académica, diseño de evaluaciones y comunicación con los estudiantes. Este hallazgo permite observar el carácter inicialmente reactivo de las instituciones ante una tecnología cuya adopción ocurrió con mayor rapidez que la elaboración de políticas.

Las propuestas posteriores han avanzado hacia modelos más diferenciados. Perkins, Furze, Roe y MacVaugh (2024) desarrollaron la Artificial Intelligence Assessment Scale como marco para establecer diferentes niveles de participación permitida de IA en actividades evaluativas. Aunque este modelo se analizará con mayor profundidad en los capítulos posteriores, su principio general resulta relevante en este punto: no todas las actividades educativas requieren el mismo nivel de restricción tecnológica.

Esta consideración ayuda a superar una dicotomía poco productiva. La decisión educativa no consiste necesariamente en escoger entre una institución que prohíbe la inteligencia artificial y otra que permite su utilización sin restricciones. Las condiciones pueden variar según los objetivos de aprendizaje, el nivel educativo, la disciplina y el tipo de actividad. Lo importante es que esas condiciones sean explícitas y pedagógicamente justificables.

Uso legítimo, uso inapropiado y delegación intelectual mediante IA

Uno de los principales problemas introducidos por la inteligencia artificial generativa consiste en delimitar la frontera entre asistencia legítima y sustitución indebida del trabajo intelectual. La creación de listas permanentes de herramientas autorizadas o prohibidas ofrece una solución limitada, porque las funcionalidades evolucionan con rapidez y aplicaciones originalmente destinadas a otras tareas incorporan progresivamente capacidades generativas.

Un criterio más sostenible consiste en analizar la función que desempeña la tecnología dentro del proceso académico. El uso puede considerarse legítimo cuando resulta compatible con las instrucciones de la actividad, mantiene la responsabilidad intelectual del estudiante y no altera de manera engañosa la evidencia de aprendizaje que debe producirse. Dependiendo del propósito educativo, esto puede incluir solicitar explicaciones adicionales, recibir retroalimentación sobre un texto, generar ejercicios de práctica, explorar distintas aproximaciones a un problema o utilizar la IA como objeto explícito de análisis.

El uso inapropiado aparece cuando la participación de la herramienta contradice las condiciones establecidas o modifica sustancialmente la relación entre el producto entregado y el trabajo realizado por el estudiante. La conducta puede constituir una vulneración de la integridad aunque no corresponda estrictamente a una forma clásica de plagio. Presentar como razonamiento personal una argumentación que fue producida por un sistema generativo sin que esa asistencia estuviera permitida constituye un ejemplo de esta discrepancia.

Para analizar mejor estas situaciones resulta útil introducir el concepto de delegación intelectual. En este libro, se entenderá por delegación intelectual la transferencia a una herramienta generativa de operaciones cognitivas sustantivas que el estudiante debía desarrollar para alcanzar o demostrar un determinado resultado de aprendizaje. Esta definición no implica que toda delegación a una tecnología sea académicamente incorrecta. Las actividades profesionales utilizan permanentemente herramientas para automatizar procesos. La incompatibilidad aparece

cuando la tarea delegada coincide precisamente con la competencia que el estudiante debía ejercitar o demostrar.

La diferencia puede observarse en una actividad de escritura. Si el objetivo consiste en desarrollar la capacidad para construir un argumento, delegar íntegramente la formulación de la tesis, la organización de las razones y la elaboración de las conclusiones reduce considerablemente la evidencia de esa competencia. En cambio, una vez elaborado el razonamiento, solicitar comentarios sobre claridad o coherencia puede constituir una forma de asistencia compatible con el objetivo, siempre que las reglas de la actividad lo permitan.

Una situación equivalente puede producirse en programación. Solicitar una explicación sobre la causa de un error después de haber intentado desarrollar y analizar un código no representa necesariamente la misma participación cognitiva que solicitar un programa terminado, ejecutarlo y presentarlo sin comprender su funcionamiento. La diferencia relevante se encuentra en qué operaciones intelectuales realizó realmente el estudiante.

La transparencia también debe formar parte de este análisis. Declarar el uso de una herramienta permite conocer mejor las condiciones de producción del trabajo y atribuir adecuadamente las contribuciones. Sin embargo, transparencia y autorización son criterios diferentes. Informar que se utilizó inteligencia artificial no legitima una práctica expresamente prohibida, del mismo modo que una utilización autorizada puede requerir ser declarada para preservar la trazabilidad del proceso.

Otro criterio fundamental es la verificación. La responsabilidad intelectual implica revisar la exactitud de la información, comprobar las referencias, contrastar las afirmaciones relevantes y corregir errores antes de incorporar un resultado generado a un trabajo académico. La facilidad con la que los sistemas producen respuestas coherentes puede generar una confianza excesiva si el estudiante no dispone de conocimientos suficientes para evaluarlas. Por esta razón, la supervisión crítica debe considerarse parte del uso responsable y no una actividad opcional posterior.

La relación entre estos criterios puede sintetizarse mediante una matriz que no pretende clasificar automáticamente todas las situaciones, sino proporcionar preguntas consistentes para el análisis.

Tabla 1. *Criterios para valorar la intervención de inteligencia artificial generativa en una actividad académica*

Criterio	Uso compatible con la integridad académica	Situación de riesgo	Uso potencialmente incompatible
<i>Autorización</i>	La utilización corresponde a las condiciones explícitas de la actividad	Las instrucciones son ambiguas o no especifican el alcance permitido	La herramienta se utiliza en contra de una restricción explícita
<i>Contribución intelectual</i>	El estudiante conserva las operaciones cognitivas centrales de la tarea	La IA interviene de forma importante y resulta difícil diferenciar las contribuciones	La IA sustituye la competencia que debía demostrar el estudiante
<i>Transparencia</i>	La intervención de la herramienta se declara cuando corresponde	La declaración es incompleta o no permite comprender cómo fue utilizada	Se oculta deliberadamente una participación sustantiva de la IA
<i>Verificación</i>	El estudiante comprueba, contrasta y asume responsabilidad	La revisión es superficial o se acepta información cuya validez no	Se presentan resultados generados sin comprensión ni verificación

	por contenido	el se comprobado	ha
<i>Correspondencia con el aprendizaje</i>	El uso de IA apoya el proceso invalidar evidencia de aprendizaje	Existe el incertidumbre sobre cuánto del desempeño corresponde al estudiante	El producto ya no permite inferir razonablemente la competencia evaluada

Nota. Elaboración propia a partir de ICAI (2021), Eaton (2023), Luo (2024), Miao y Holmes (2023) y Perkins et al. (2024).

La utilidad de esta matriz radica en que desplaza la pregunta desde la herramienta hacia la práctica. Clasificar una aplicación como permitida o prohibida puede ofrecer claridad temporal, pero no resuelve las transformaciones constantes de las funcionalidades tecnológicas. En cambio, criterios como autorización, contribución intelectual, transparencia, verificación y correspondencia con el aprendizaje pueden aplicarse a herramientas diferentes y mantenerse relevantes aunque cambien las plataformas utilizadas.

Este planteamiento también evita que la política educativa se construya exclusivamente alrededor de la detección. Una cultura de integridad más sólida debería permitir que el estudiante conozca antes de realizar una actividad qué participación de IA resulta compatible con sus objetivos. La claridad preventiva es pedagógicamente preferible a establecer las reglas únicamente después de que aparece una sospecha.

La pregunta central, por tanto, no debería formularse simplemente como “¿se utilizó inteligencia artificial?”. Resulta más preciso preguntar para qué se utilizó, qué parte del proceso modificó, qué decisiones permanecieron bajo responsabilidad del estudiante y si el resultado sigue representando de manera válida aquello que debía demostrar.

Desafíos de la integridad académica en América

Latina y Ecuador

La discusión internacional sobre inteligencia artificial generativa no puede trasladarse automáticamente al contexto latinoamericano. La región presenta diferencias relevantes entre países e instituciones en materia de conectividad, infraestructura, formación docente, disponibilidad tecnológica, políticas institucionales y alfabetización digital. Estas condiciones influyen tanto en las posibilidades de aprovechamiento educativo como en los riesgos asociados con el uso de herramientas generativas.

La evidencia regional es todavía emergente y presenta una distribución desigual. Motta Soto et al. (2026) realizaron una revisión de alcance de estudios empíricos publicados entre 2021 y 2025 sobre adopción de inteligencia artificial generativa en educación superior latinoamericana. La revisión identificó 31 estudios y mostró un predominio de investigaciones centradas en estudiantes y en aplicaciones relacionadas con apoyo al aprendizaje y escritura académica. Los problemas vinculados con integridad y ética aparecieron con frecuencia, mientras que la evaluación, los impactos educativos y la equidad estructural recibieron una atención comparativamente menor.

Este patrón es coherente con la revisión sistemática desarrollada por Blake (2026), centrada en la investigación sobre docentes e inteligencia artificial generativa en educación superior latinoamericana publicada entre 2023 y abril de 2025. El análisis de 35 estudios encontró que la necesidad de capacitación docente y los desafíos éticos y de integridad académica aparecían cada uno en el 54 % de los trabajos revisados. El autor caracteriza esta primera producción regional como predominantemente reactiva y orientada hacia problemas inmediatos, entre ellos el plagio y el dominio instrumental de las herramientas, mientras aspectos como desigualdad tecnológica y desarrollo de políticas institucionales recibían menor atención.

Estos resultados deben interpretarse con cautela. No demuestran que los problemas identificados tengan la misma magnitud en todos los países latinoamericanos ni permiten extrapolar conclusiones entre sistemas

educativos diferentes. Más bien, muestran cuáles han sido las preocupaciones predominantes de un campo de investigación que todavía se encuentra en consolidación.

Una dificultad adicional consiste en que gran parte de la evidencia disponible se concentra en educación superior. El conocimiento sobre utilización de inteligencia artificial generativa en educación básica y bachillerato latinoamericanos es considerablemente menos desarrollado. Esta limitación es relevante para un libro que aborda ambos niveles educativos, porque impide trasladar automáticamente al alumnado escolar los hallazgos obtenidos con universitarios.

En estudiantes menores de edad intervienen además factores particulares. La autonomía para evaluar información, reconocer errores, comprender las implicaciones de privacidad y distinguir entre asistencia y sustitución de una tarea se desarrolla progresivamente. La formación en inteligencia artificial en estos niveles requiere, por tanto, una orientación ajustada a la edad, al contexto y a los propósitos curriculares. El planteamiento de UNESCO de desarrollar competencias progresivas en IA ofrece una base conceptual útil, pero su aplicación requiere adaptación a las condiciones de cada sistema educativo (Miao et al., 2024).

En Ecuador comienza a existir evidencia empírica que permite analizar el fenómeno con mayor precisión. Torres Díaz, Duarte, Rivera y Beltrán Flandoli (2025) estudiaron la relación entre utilización de inteligencia artificial generativa y plagio autoinformado en una muestra de 4.811 estudiantes presenciales pertenecientes a ocho universidades ecuatorianas. Los resultados mostraron asociaciones entre determinadas variables relacionadas con el uso y percepción de IA y los niveles autoinformados de plagio. Sin embargo, el diseño transversal y correlacional del estudio no permite establecer relaciones causales.

Esta limitación metodológica es fundamental. Los resultados no permiten afirmar que la utilización de inteligencia artificial produzca por sí misma conductas de plagio. De hecho, el estudio encontró que un mayor conocimiento de herramientas de inteligencia artificial se asociaba con menores niveles de plagio, mientras otras variables presentaban asociaciones en dirección contraria (Torres Díaz et al., 2025). El fenómeno parece, por tanto, más complejo que una relación lineal entre mayor uso tecnológico y menor integridad.

Los hallazgos de Buele et al. (2025), obtenidos con 833 estudiantes universitarios ecuatorianos, también sugieren que el comportamiento ético frente a la inteligencia artificial se relaciona con componentes cognitivos y actitudinales y no únicamente con la frecuencia de utilización. Esta evidencia fortalece la necesidad de centrar las intervenciones educativas en conocimiento, criterios de uso y reflexión ética, en lugar de asumir que la reducción del uso constituye por sí sola una estrategia suficiente de integridad.

La evidencia disponible para Ecuador continúa siendo insuficiente para describir el conjunto del sistema educativo. Los estudios mencionados corresponden a educación superior y no deberían utilizarse para afirmar que los mismos patrones se reproducen en escuelas y colegios. Esta diferencia evidencia una línea de investigación pendiente y obliga a evitar generalizaciones que la literatura todavía no puede sostener.

Otro desafío regional corresponde a la desigualdad de acceso. Las posibilidades de utilizar inteligencia artificial no son equivalentes para todos los estudiantes. Existen diferencias de conectividad, disponibilidad de dispositivos y acceso a versiones comerciales con mayores capacidades. También existen diferencias en alfabetización digital y en el conocimiento necesario para evaluar críticamente los resultados. Por ello, incorporar obligatoriamente determinadas herramientas en actividades evaluativas sin analizar las condiciones de acceso puede introducir desigualdades adicionales.

La formación docente representa una segunda dimensión crítica. No resulta razonable exigir que cada profesor resuelva individualmente los dilemas asociados con herramientas cuya evolución es continua. La respuesta necesita incluir criterios institucionales, formación pedagógica y espacios para discutir las particularidades de cada disciplina. La alfabetización del profesorado tampoco debería reducirse al aprendizaje de instrucciones o técnicas para utilizar plataformas. Requiere comprender cuestiones relacionadas con autoría, privacidad, confiabilidad, sesgos, evaluación y responsabilidad.

La tercera cuestión corresponde a las políticas institucionales. Una norma excesivamente general puede resultar sencilla de comunicar, pero difícil de aplicar de manera pedagógicamente coherente. Una actividad de escritura, un problema matemático, un proyecto de programación, una

producción artística y una simulación profesional no evalúan las mismas competencias y, por tanto, no necesariamente requieren idénticas restricciones. Las políticas deberían establecer principios comunes y permitir que las condiciones específicas se definan de acuerdo con los objetivos educativos.

También es necesario evitar que la preocupación por la integridad convierta toda interacción tecnológica en objeto de sospecha. Una cultura construida exclusivamente sobre vigilancia puede deteriorar la relación de confianza que la propia integridad académica intenta proteger. La claridad de las expectativas, la educación preventiva y el diseño coherente de las actividades constituyen respuestas más sostenibles que la presunción generalizada de deshonestidad.

En sentido contrario, normalizar cualquier utilización de inteligencia artificial con el argumento de que estas tecnologías estarán presentes en la vida profesional tampoco constituye una respuesta suficiente. El contexto educativo posee objetivos específicos. Determinadas actividades deben permitir al estudiante desarrollar capacidades antes de automatizarlas. La existencia de calculadoras no eliminó la necesidad de comprender principios matemáticos, del mismo modo que la disponibilidad de sistemas capaces de redactar no elimina automáticamente la necesidad de aprender a construir argumentos.

La cuestión fundamental consiste en preservar la relación entre aprendizaje, autoría y responsabilidad. La inteligencia artificial generativa puede modificar la forma en que se desarrollan numerosas actividades académicas, pero no elimina la necesidad de que los estudiantes puedan demostrar qué saben, qué comprenden y qué son capaces de realizar.

Desde esta perspectiva, la integridad académica en la era de la inteligencia artificial no debe entenderse como un intento de conservar artificialmente un sistema educativo previo a estas tecnologías. Tampoco implica aceptar de forma acrítica la automatización de los procesos académicos. Supone establecer condiciones en las que la tecnología pueda incorporarse sin distorsionar la evidencia del aprendizaje ni diluir la responsabilidad intelectual.

La inteligencia artificial generativa introduce nuevos problemas, pero también hace más visibles debilidades que preceden a su aparición.

Cuando una actividad puede resolverse mediante la generación automática de un producto sin que el estudiante comprenda su contenido, es necesario interrogar tanto la conducta del estudiante como el diseño de la propia actividad. Esta relación entre integridad y evaluación constituye precisamente el punto de transición hacia los siguientes capítulos, en los que se analizarán las nuevas manifestaciones de deshonestidad académica, los límites de las estrategias de detección y las transformaciones que la inteligencia artificial exige en los procesos evaluativos.

Capítulo 2

Deshonestidad académica y detección del uso de inteligencia artificial

La inteligencia artificial generativa ha modificado algunas de las condiciones bajo las cuales pueden producirse las infracciones académicas, pero no ha creado desde cero el problema de la deshonestidad. El plagio, la colaboración no autorizada, la falsificación de información, la suplantación y la contratación de terceros para elaborar trabajos existían antes de la expansión de los modelos generativos. Lo que estas tecnologías introducen es una reducción considerable de las barreras necesarias para externalizar determinadas actividades intelectuales y, al mismo tiempo, una mayor dificultad para establecer cómo se produjo un determinado producto académico.

Esta característica diferencia parcialmente a la inteligencia artificial generativa de formas anteriores de asistencia no autorizada. En el plagio textual convencional puede existir una fuente identificable con la cual comparar el documento presentado. En la contratación de terceros puede intentarse establecer una diferencia de autoría entre trabajos. En cambio, un sistema generativo puede producir en segundos un texto nuevo, adaptarlo a instrucciones específicas, modificar posteriormente su estilo y combinarlo con fragmentos elaborados por el propio estudiante. El resultado puede no coincidir literalmente con ninguna fuente previamente publicada y, aun así, no representar el desempeño intelectual que la actividad pretendía evaluar.

Por ello, el desafío de este capítulo no consiste exclusivamente en identificar nuevas formas de fraude. También implica analizar las limitaciones de los procedimientos mediante los cuales se intenta inferir que una infracción ocurrió. La sospecha, la detección tecnológica y la demostración de una conducta académicamente inapropiada pertenecen a momentos diferentes y no deberían tratarse como equivalentes. Esta distinción adquiere especial importancia cuando una acusación puede

generar consecuencias académicas, disciplinarias o reputacionales para un estudiante.

Nuevas manifestaciones de deshonestidad académica mediadas por IA

La deshonestidad académica comprende conductas heterogéneas cuya característica común consiste en producir una representación engañosa del aprendizaje, de la autoría o de las condiciones bajo las cuales se realizó una actividad. En este sentido, la inteligencia artificial generativa no necesariamente crea categorías completamente nuevas de mala conducta, sino que amplía y transforma las posibilidades mediante las cuales pueden producirse prácticas previamente conocidas.

Antes de la expansión de la inteligencia artificial generativa, la externalización de actividades académicas ya podía producirse mediante el denominado contract cheating, en el que un estudiante encarga a otra persona la elaboración parcial o total de una tarea para posteriormente presentarla como propia. La literatura previa a ChatGPT describía esta práctica como particularmente difícil de detectar debido a que el trabajo producido podía ser original desde el punto de vista textual y, por tanto, no necesariamente generar coincidencias en sistemas convencionales de detección de plagio (Ahsan et al., 2022). La inteligencia artificial generativa modifica esta lógica porque permite una forma de externalización inmediata, accesible y potencialmente gratuita de determinadas operaciones intelectuales.

Sin embargo, equiparar automáticamente el uso de inteligencia artificial con contract cheating tampoco resulta preciso. Un modelo generativo no es un tercero humano contratado y las formas de intervención pueden variar desde una asistencia mínima hasta la producción completa del trabajo. Perkins (2023) sostiene que determinar si el uso de un modelo de lenguaje constituye una infracción depende de las condiciones de la evaluación y de las políticas institucionales, no simplemente de que la herramienta haya sido utilizada. Esta distinción impide convertir la presencia de inteligencia artificial en sinónimo de conducta deshonestas.

La infracción aparece con mayor claridad cuando una herramienta es utilizada para realizar una actividad expresamente prohibida o para sustituir una competencia que el estudiante debía demostrar personalmente. Esto puede incluir la generación integral de un ensayo, la solución automática de ejercicios, la producción de código, la elaboración de análisis, la resolución de problemas o la creación de respuestas durante una evaluación en la que tales recursos no se encontraban autorizados. En estos casos, el problema no se limita a la procedencia del contenido, sino a la falsa representación del desempeño.

También pueden existir formas parciales de sustitución. Un estudiante puede elaborar una introducción y delegar el desarrollo argumentativo a una herramienta, solicitar la generación de una conclusión o introducir fragmentos producidos por IA en diferentes partes de un documento. Estas prácticas generan productos híbridos en los que las fronteras entre escritura humana y generación automatizada resultan menos claras. La existencia de este tipo de documentos es una de las razones por las que clasificar un trabajo únicamente como “humano” o “generado por IA” resulta cada vez menos adecuado.

La deshonestidad mediada por IA puede además extenderse más allá de la producción textual. Los modelos contemporáneos pueden generar imágenes, presentaciones, código, datos simulados, traducciones, audio y otros materiales. Dependiendo de la actividad, presentar estos productos como resultado del propio desempeño puede constituir una infracción cuando la generación automatizada sustituye la competencia que debía evaluarse. Por ello, una política centrada exclusivamente en ChatGPT o en textos escritos corre el riesgo de quedar rápidamente desactualizada.

Otra manifestación problemática corresponde a la utilización de inteligencia artificial para fabricar o alterar información. Los sistemas generativos pueden producir referencias bibliográficas inexistentes, datos ficticios, citas atribuidas incorrectamente o afirmaciones presentadas con apariencia de precisión académica. Walters y Wilder (2023) evaluaron 636 referencias producidas por GPT 3.5 y GPT 4 en revisiones breves de literatura y encontraron que el 55 % de las referencias generadas por GPT 3.5 y el 18 % de las generadas por GPT 4 eran inexistentes. Además, una proporción de las referencias reales contenía errores bibliográficos sustantivos. Estos resultados proceden de versiones tempranas de los modelos y no deben asumirse como tasas actuales, pero muestran un

mecanismo relevante: una herramienta generativa puede producir información bibliográfica plausible sin que exista una fuente académica correspondiente.

La presencia de una referencia ficticia no demuestra por sí misma que se haya utilizado inteligencia artificial. Un estudiante puede inventar deliberadamente una referencia, cometer errores al transcribirla o incorporar información incorrecta desde otra fuente. No obstante, la fabricación de referencias puede constituir una infracción académica independientemente de la tecnología empleada. La diferencia es importante porque permite investigar el problema verificable, la existencia y correspondencia de la fuente, sin necesidad de afirmar inicialmente una causa tecnológica que no ha sido demostrada.

La reformulación automatizada constituye otra modalidad relevante. Las herramientas de parafraseo preceden a la actual expansión de los modelos generativos y ya habían sido identificadas como un desafío para la integridad académica. Roe y Perkins (2022) describieron un escenario de competencia tecnológica en el que los sistemas destinados a transformar textos podían utilizarse para ocultar similitudes con las fuentes originales y dificultar la identificación mediante software de comparación textual. Con los modelos generativos actuales, este procedimiento puede ampliarse considerablemente, pues un texto puede resumirse, expandirse, cambiar de registro, traducirse o reescribirse repetidamente sin mantener necesariamente coincidencias literales con su versión inicial.

No toda reformulación constituye fraude. Utilizar una herramienta de apoyo lingüístico puede ser legítimo en determinadas condiciones, particularmente cuando la finalidad es mejorar la expresión de ideas previamente elaboradas por el estudiante. La infracción depende de qué se transforme, qué grado de intervención exista, qué normas hayan sido establecidas y si el estudiante continúa siendo responsable de las operaciones intelectuales esenciales del trabajo. La dificultad radica precisamente en que la frontera entre asistencia lingüística y generación sustantiva puede no ser evidente cuando las herramientas integran múltiples funciones en una misma interfaz.

Las investigaciones realizadas con estudiantes muestran además que las percepciones sobre estas fronteras no son homogéneas. Waltzer et al. (2023), en un estudio con 140 estudiantes y 69 docentes de educación

secundaria, encontraron que los estudiantes desaprobaban presentar como propio un ensayo generado mediante ChatGPT, aunque evaluaban diversas posibles infracciones relacionadas con la herramienta de manera menos negativa que sus profesores. La diferencia evidencia que las normas sociales sobre el uso académico de IA todavía se encuentran en formación y que no puede asumirse que estudiantes y docentes atribuyen automáticamente el mismo significado a una determinada práctica.

La situación es igualmente relevante en educación superior. En Ecuador, Torres-Díaz et al. (2025) analizaron datos de 4.811 estudiantes pertenecientes a ocho universidades y encontraron asociaciones entre determinadas formas de utilización y percepción de la inteligencia artificial y los niveles de plagio autoinformado. El diseño correlacional del estudio impide establecer causalidad, por lo que sus resultados no permiten afirmar que utilizar IA produzca plagio. En cambio, muestran que las relaciones entre uso tecnológico, actitudes, conocimiento y conductas de integridad son más complejas que una asociación directa entre frecuencia de uso y fraude.

Esta complejidad debería reflejarse en las políticas institucionales. Una regulación que considere cualquier intervención de inteligencia artificial como deshonestidad corre el riesgo de incluir prácticas legítimas, mientras que una regulación demasiado general puede dejar sin definición conductas en las que la tecnología sustituye de manera sustancial el desempeño esperado. La clasificación debe depender de la actividad realizada y no únicamente del nombre de la herramienta utilizada.

Generación, reformulación y producción automatizada de contenidos académicos

La producción de contenidos mediante inteligencia artificial puede situarse en un continuo que abarca desde modificaciones superficiales hasta la generación completa de un producto. Esta característica dificulta las clasificaciones binarias y exige analizar cómo se construyó el resultado presentado.

En un extremo se encuentra la generación prácticamente integral. El estudiante puede introducir las instrucciones de una actividad en un

modelo de lenguaje y utilizar directamente la respuesta obtenida. La posibilidad de producir textos coherentes mediante procedimientos de este tipo fue una de las primeras preocupaciones documentadas después de la expansión pública de ChatGPT. Perkins (2023) mostró que los modelos de lenguaje podían producir contenido suficientemente original en términos de coincidencia textual como para dificultar la identificación mediante procedimientos diseñados para detectar plagio convencional.

Este aspecto es fundamental para diferenciar los sistemas de comparación textual de los detectores de inteligencia artificial. Los primeros buscan coincidencias entre un documento y fuentes disponibles en determinadas bases de datos. Un texto generado por un modelo, en cambio, puede no existir previamente en ninguna fuente. Por tanto, la ausencia de coincidencias no constituye evidencia de autoría humana ni demuestra que el documento haya sido elaborado por el estudiante.

Entre la generación completa y la escritura enteramente humana existen numerosas configuraciones híbridas. El estudiante puede solicitar un esquema, redactar una parte, pedir al sistema que complete otra, modificar posteriormente el resultado y combinarlo con fragmentos propios. También puede introducir un documento previamente escrito y solicitar su reestructuración, ampliación o reducción. Estos escenarios hacen que el concepto de “porcentaje de IA” deba interpretarse con cautela, porque el origen material de una determinada secuencia de palabras no informa necesariamente sobre la distribución real de las decisiones intelectuales durante el proceso.

La reformulación automatizada añade otra capa de complejidad. Roe y Perkins (2022) explican que las herramientas de parafraseo automatizado permiten transformar vocabulario, sintaxis y organización de los textos, dificultando la identificación de la relación con el material original. La evolución de los modelos generativos amplía esa capacidad: no es necesario sustituir palabras por sinónimos, porque el sistema puede reconstruir conceptualmente una idea y expresarla mediante estructuras considerablemente distintas.

Esta capacidad puede utilizarse de manera legítima o indebida. Una persona que ha producido un argumento puede solicitar sugerencias para mejorar su claridad lingüística. En cambio, alguien puede copiar el contenido de una fuente, solicitar una reformulación suficientemente

diferente y presentar posteriormente el resultado sin reconocer el origen intelectual. En este último escenario, la modificación formal no elimina el problema de atribución. La falta académica deriva de la apropiación de ideas o contenidos y no de la coincidencia literal de palabras.

La traducción automatizada presenta un problema similar. Traducir un texto propio puede constituir una asistencia lingüística; traducir una fuente ajena y presentar el contenido resultante como elaboración original puede constituir plagio aunque no exista coincidencia superficial con la fuente en el idioma de entrega. Por tanto, la evaluación de integridad debe conservar la atención sobre la procedencia de las ideas y no limitarse a la semejanza textual.

Los sistemas generativos también permiten modificar deliberadamente características asociadas con la escritura automatizada. La literatura ha demostrado que la reformulación, el parafraseo y otras técnicas de ofuscación pueden disminuir considerablemente la capacidad de determinados detectores para reconocer contenido generado. Weber-Wulff et al. (2023), al evaluar catorce sistemas, observaron que técnicas de parafraseo y modificaciones del contenido deterioraban el desempeño de los detectores. Esto cuestiona una estrategia institucional basada en una competencia permanente entre herramientas generadoras y herramientas destinadas a identificarlas.

La producción automatizada incluye además un riesgo epistemológico que no debe confundirse con la deshonestidad intencional: la generación de información falsa. Walters y Wilder (2023) demostraron que los modelos evaluados podían producir referencias ficticias con formato aparentemente verosímil. Un estudiante puede incorporar una referencia de este tipo deliberadamente o hacerlo porque confió de manera acrítica en la respuesta. Las dos situaciones pueden tener implicaciones académicas diferentes aunque produzcan el mismo resultado documental.

Esta diferencia entre intención, conducta y resultado es importante para los procedimientos institucionales. La existencia de un error no demuestra necesariamente intención de fraude. Sin embargo, determinadas normas académicas imponen al estudiante responsabilidad sobre la información que presenta, independientemente de que el error haya sido producido inicialmente por una herramienta. Por ello, la utilización de sistemas

generativos no elimina la obligación de verificar fuentes, datos y afirmaciones.

Los productos híbridos representan probablemente uno de los principales retos para la detección contemporánea. Las evaluaciones más recientes muestran que incluso sistemas que han mejorado su capacidad para identificar textos completamente generados pueden tener dificultades cuando un documento combina contribuciones humanas y automatizadas. Hadra et al. (2026), utilizando 192 textos distribuidos entre escritura auténtica de estudiantes de inglés como lengua extranjera, escritura profesional humana, producción generada por IA y composiciones híbridas, encontraron que los dos detectores comerciales analizados presentaban un desempeño particularmente deficiente en documentos de autoría mixta.

Van Vlasselaer et al. (2026) obtuvieron resultados más favorables para algunos detectores en condiciones experimentales controladas. Su estudio comparó cuatro herramientas sobre 160 documentos humanos, generados, híbridos y posteriormente “humanizados”. Uno de los sistemas evaluados alcanzó un desempeño claramente superior a los demás y los falsos positivos sobre los textos completamente humanos fueron escasos. Sin embargo, los autores mantuvieron una conclusión prudente: incluso los sistemas de mejor desempeño deberían utilizarse como señales iniciales y no como evidencia única en decisiones de alta consecuencia.

Esta evolución de la evidencia es importante. Los primeros estudios de 2023 mostraron limitaciones muy pronunciadas y condujeron a cuestionar la viabilidad de los detectores. Los estudios posteriores indican que determinadas herramientas han mejorado en escenarios controlados. Esto impide sostener de manera rigurosa que todos los detectores sean incapaces de identificar cualquier contenido generado. Al mismo tiempo, la mejora técnica tampoco convierte sus resultados en demostraciones automáticas de una infracción académica.

La razón es conceptual además de técnica. Detectar que un segmento presenta características asociadas con generación automatizada no permite determinar por sí mismo si la utilización estaba autorizada, quién introdujo el contenido, qué transformaciones posteriores se realizaron o si la actividad permitía cierta intervención de inteligencia artificial. La

detección del posible origen tecnológico y la determinación de una infracción académica son problemas relacionados, pero distintos.

Autoría, originalidad y transparencia en el uso de IA generativa

La utilización de inteligencia artificial generativa desafía las nociones convencionales de autoría porque permite distribuir la producción de un trabajo entre intervenciones humanas y automatizadas. Sin embargo, esta complejidad no elimina la necesidad de atribuir responsabilidad por el producto final.

En contextos educativos, la autoría cumple una función adicional a la atribución intelectual. Permite vincular un producto con la competencia de la persona que lo presenta. Cuando un estudiante entrega un ensayo, una resolución o un proyecto bajo su nombre, existe una expectativa de que el documento refleja su participación dentro de las condiciones establecidas para esa actividad. Por ello, el problema de la autoría académica no puede reducirse a determinar quién escribió físicamente cada palabra.

Perkins (2023) señala que la utilización de modelos de lenguaje debe interpretarse de acuerdo con las reglas de la actividad y con las políticas institucionales. Esta posición resulta especialmente relevante porque una misma intervención tecnológica puede tener significados diferentes. Si el docente autoriza expresamente la utilización de inteligencia artificial para generar ideas y exige posteriormente una evaluación crítica de los resultados, la generación forma parte del procedimiento previsto. Si la actividad pretende evaluar producción individual sin asistencia externa y el estudiante delega la respuesta, la misma tecnología adquiere una función diferente.

La transparencia ayuda a distinguir estos escenarios. Declarar el uso de inteligencia artificial permite conocer, al menos parcialmente, las condiciones bajo las cuales se elaboró un trabajo. Sin embargo, la transparencia no debe confundirse con autorización. Un estudiante puede reconocer que utilizó una herramienta y aun así haber incumplido las condiciones de la actividad. Del mismo modo, puede existir un uso

permitido que no requiera una declaración específica cuando las normas institucionales así lo establecen.

Por esta razón, las instituciones necesitan evitar fórmulas genéricas como “el uso de IA debe declararse siempre” si no establecen qué se entiende por uso. Procesadores de texto, correctores, traductores, funciones de transcripción, motores de búsqueda y aplicaciones educativas incorporan cada vez más componentes de inteligencia artificial. Una política excesivamente amplia puede convertir actividades instrumentales en aparentes infracciones sin que exista sustitución del trabajo académico.

La distinción entre función y plataforma resulta especialmente importante. Una misma aplicación puede ofrecer transcripción, corrección lingüística, generación de imágenes, búsqueda, resumen y producción textual. Wright (2026) advierte que políticas formuladas únicamente alrededor del nombre o de la categoría tecnológica pueden confundir funciones generativas con otras formas de asistencia automatizada que no producen el contenido intelectual evaluado. Aunque su análisis se concentra en educación superior, el principio resulta aplicable de manera más amplia: la regulación debe considerar qué hizo la herramienta y no solamente qué herramienta fue utilizada.

La originalidad presenta una dificultad semejante. Un trabajo puede no contener coincidencias literales con fuentes existentes y, sin embargo, no representar una producción intelectual original del estudiante. Los modelos generativos pueden elaborar texto nuevo desde el punto de vista de su secuencia lingüística. Por ello, ausencia de plagio textual y existencia de autoría auténtica no son equivalentes.

La originalidad académica tampoco exige necesariamente crear ideas completamente inéditas. En muchos niveles educativos se espera que el estudiante comprenda, reorganice, relacione y argumente a partir de conocimiento existente. El valor de la actividad reside precisamente en que estas operaciones sean realizadas por quien está aprendiendo. Una herramienta que produce una respuesta original en términos lingüísticos puede disminuir la originalidad del proceso intelectual si reemplaza esas operaciones.

Por esta razón, la transparencia debería describir la naturaleza de la contribución tecnológica cuando esa información sea relevante para la

actividad. Una simple frase que indique “se utilizó inteligencia artificial” puede resultar insuficiente si no permite conocer para qué se utilizó. Dependiendo del contexto, puede ser más informativo distinguir entre generación de ideas, asistencia lingüística, traducción, síntesis, producción de borradores, generación de código o revisión del contenido.

Gustilo et al. (2024), al estudiar prácticas y percepciones docentes relacionadas con escritura asistida algorítmicamente, muestran que las tecnologías de escritura plantean dificultades precisamente porque las fronteras entre apoyo y sustitución pueden variar según las prácticas pedagógicas y las políticas existentes. Esta heterogeneidad refuerza la necesidad de comunicar las condiciones antes de la realización de las tareas, y no intentar definir las retrospectivamente después de detectar un posible uso.

La responsabilidad final permanece en la persona que presenta el trabajo. Esto implica que la atribución transparente de una herramienta no elimina la obligación de revisar el contenido. Si una inteligencia artificial genera una referencia inexistente, una afirmación falsa o un cálculo incorrecto, el estudiante no puede desplazar completamente la responsabilidad académica hacia el sistema. La utilización de la herramienta forma parte de las decisiones que condujeron al producto entregado.

A la inversa, tampoco debería atribuirse deshonestidad a un estudiante únicamente porque un texto presenta una estructura, vocabulario o nivel de corrección que un docente asocia intuitivamente con inteligencia artificial. La autoría no puede determinarse de forma confiable mediante impresiones estilísticas aisladas. Fleckenstein et al. (2024) encontraron que tanto docentes en formación como docentes experimentados tenían dificultades para distinguir textos generados por ChatGPT de ensayos escritos por estudiantes y que mostraban además una confianza excesiva en sus propias clasificaciones.

En educación secundaria, Waltzer et al. (2023) observaron un fenómeno semejante. Los docentes identificaron correctamente el texto generado por ChatGPT aproximadamente en el 70 % de las comparaciones, mientras que los estudiantes alcanzaron alrededor del 62 %. La experiencia previa con ChatGPT y la confianza declarada no predijeron el desempeño de manera significativa. Estos resultados muestran que la familiaridad

subjetiva con el estilo de una herramienta no constituye una base suficiente para determinar autoría.

La cuestión de la transparencia debería, por tanto, operar en ambas direcciones. Los estudiantes necesitan conocer las condiciones de uso y declarar las intervenciones tecnológicas cuando corresponda. Las instituciones, a su vez, deben ser transparentes respecto de las herramientas, criterios y procedimientos utilizados para investigar una posible infracción. La integridad no puede exigirse únicamente a una de las partes del proceso educativo.

Detectores de contenido generado por IA: posibilidades y limitaciones

La dificultad para identificar contenido producido mediante sistemas generativos impulsó rápidamente el desarrollo de herramientas diseñadas para estimar si un texto fue escrito por una persona o generado mediante inteligencia artificial. Estos sistemas han adquirido una presencia significativa en la discusión educativa porque parecen ofrecer una solución técnica a un problema que, de otro modo, requiere investigación y juicio académico.

Sin embargo, un detector de inteligencia artificial no funciona de la misma manera que un sistema de comparación de similitud textual. El software tradicional puede identificar que determinadas secuencias coinciden con una fuente concreta y permite al evaluador examinar directamente esa relación. Los detectores de IA, en cambio, generalmente analizan propiedades estadísticas y lingüísticas del texto y producen una clasificación o estimación sobre su posible origen. El resultado es inferencial y probabilístico, no una identificación directa de una fuente.

Esta diferencia tiene consecuencias importantes. Un porcentaje presentado por una herramienta no debe interpretarse automáticamente como la probabilidad de que un estudiante haya cometido una infracción. Tampoco significa necesariamente que ese porcentaje exacto del documento haya sido producido por un modelo generativo. La interpretación depende del funcionamiento específico del sistema, de sus umbrales y de la población de textos sobre la que fue evaluado.

Los primeros estudios independientes documentaron limitaciones importantes. Elkhataat et al. (2023) compararon diferentes detectores sobre contenido humano y contenido producido por GPT 3.5 y GPT 4. Encontraron variabilidad entre herramientas, mejor identificación de textos generados mediante el modelo más antiguo y clasificaciones incorrectas o inciertas sobre textos humanos. Los resultados mostraron tempranamente que el desempeño de un detector podía deteriorarse conforme evolucionaban los modelos generadores.

Weber-Wulff et al. (2023) realizaron una evaluación más amplia de doce herramientas públicas y dos sistemas comerciales. Ninguna alcanzó un desempeño que los autores consideraran suficientemente fiable para utilizar sus resultados como prueba de deshonestidad académica. El estudio documentó tanto falsos positivos, en los que contenido humano era etiquetado como generado, como falsos negativos, en los que contenido automatizado era clasificado como humano. Además, el parafraseo y otras modificaciones reducían la capacidad de detección.

Los falsos positivos poseen particular importancia cuando el resultado puede desencadenar una acusación. Liang et al. (2023) encontraron un sesgo considerable al evaluar siete detectores sobre textos escritos en inglés por hablantes no nativos. En su conjunto experimental, más de la mitad de los ensayos TOEFL escritos por personas fueron clasificados incorrectamente como generados mediante inteligencia artificial. Los autores relacionaron este fenómeno con características como menor diversidad léxica y menor perplejidad estadística, propiedades que algunos detectores podían interpretar como señales de generación automatizada.

Estos hallazgos son especialmente relevantes en contextos educativos lingüísticamente diversos. Un sistema que confunde determinadas características de escritura de estudiantes que utilizan una segunda lengua con señales de generación podría introducir una desigualdad adicional en los procesos de investigación. La precisión global de una herramienta no permite asumir que su tasa de error sea equivalente en todos los grupos de estudiantes.

Los estudios sobre juicio humano tampoco respaldan la idea de que la inspección intuitiva pueda sustituir automáticamente al detector. Fleckenstein et al. (2024) encontraron que docentes con distintos niveles

de experiencia mostraban dificultades para diferenciar ensayos humanos de textos generados y tendían a confiar más en sus juicios de lo que justificaba su precisión. Waltzer et al. (2023) obtuvieron resultados similares en educación secundaria. Esto significa que eliminar el software y confiar únicamente en la intuición del docente tampoco resuelve el problema de atribución.

La evidencia ha continuado evolucionando. Hadra et al. (2026) evaluaron dos detectores comerciales mediante un conjunto equilibrado de 192 textos que incluía trabajos auténticos de estudiantes de inglés como lengua extranjera, textos profesionales, contenido generado y documentos híbridos. Las precisiones globales fueron moderadas y ambos sistemas presentaron dificultades importantes en los textos híbridos. Los autores concluyeron que podían funcionar como indicadores complementarios, pero no como fundamento exclusivo de decisiones sobre mala conducta académica.

Van Vlasselaer et al. (2026) encontraron resultados más favorables al evaluar cuatro herramientas sobre documentos con origen conocido. Uno de los detectores mostró un rendimiento especialmente alto y los sistemas analizados clasificaron correctamente los textos completamente humanos en ese conjunto experimental. Sin embargo, las diferencias entre herramientas fueron importantes y los documentos híbridos o modificados continuaron constituyendo un desafío para varios sistemas. Los autores recomiendan utilizarlos como señales iniciales integradas en una evaluación más amplia y no como evidencia única en decisiones de alta consecuencia.

Esta evolución obliga a evitar afirmaciones absolutas. No es correcto sostener que ningún detector puede identificar contenido generado; algunos muestran una capacidad considerable bajo determinadas condiciones. Tampoco es correcto asumir que una herramienta que obtiene buenos resultados experimentales puede determinar por sí sola que un estudiante incurrió en fraude. La primera pregunta es técnica y se refiere a clasificación de texto; la segunda es académica y procedimental y requiere conocer las normas aplicables, las condiciones de la tarea y la evidencia disponible.

Además, la generación contemporánea rara vez necesita permanecer completamente automatizada. Los estudiantes pueden editar respuestas,

combinar segmentos propios con contenido generado, utilizar modelos para reformular textos o solicitar modificaciones destinadas a hacerlos menos reconocibles. Esta realidad disminuye el valor de una clasificación binaria entre humano y artificial y convierte los textos híbridos en un problema central.

La actualización continua de los modelos constituye otra limitación metodológica. Un detector puede ser evaluado frente a determinados modelos y versiones, mientras que posteriormente aparecen nuevos sistemas capaces de generar patrones diferentes. La precisión es, por tanto, una propiedad situada en un momento, conjunto de datos y configuración determinados, no una garantía permanente.

Existe además una diferencia importante entre utilizar un detector con fines exploratorios y utilizarlo con fines sancionatorios. Como mecanismo de señalización, una herramienta puede orientar al docente para revisar con mayor atención un documento, comprobar fuentes o solicitar información adicional sobre el proceso. Como evidencia autónoma para imponer una sanción, la exigencia de confiabilidad es considerablemente mayor.

Bassett et al. (2026) cuestionan precisamente la utilización de detectores en procedimientos disciplinarios debido a su naturaleza probabilística, las dificultades para verificar independientemente sus resultados y el riesgo de trasladar al estudiante la obligación de demostrar su inocencia. Los autores sostienen que una acusación académica debe fundamentarse en evidencia capaz de superar el estándar institucional aplicable y no en la mera clasificación ofrecida por un algoritmo.

Por tanto, el problema no debería formularse como una competencia destinada a encontrar el detector perfecto. La pregunta pertinente es qué función puede desempeñar una señal algorítmica dentro de un procedimiento que conserve juicio académico, contraste de evidencia y garantías para el estudiante. Esta perspectiva permite utilizar los avances tecnológicos sin otorgar a un sistema de clasificación una autoridad que sus resultados no pueden justificar por sí solos.

Prevención, evidencia y respuesta ante posibles infracciones académicas

La gestión de una posible infracción relacionada con inteligencia artificial debería comenzar por una distinción fundamental entre sospecha y evidencia. Una sospecha puede justificar una revisión adicional, pero no constituye por sí misma una demostración de que la conducta ocurrió. Esta diferencia es especialmente importante cuando las señales que originan la sospecha son probabilísticas o subjetivas.

Cambios de estilo, vocabulario inusual, respuestas excesivamente generales, ausencia de errores, utilización de estructuras repetitivas o diferencias con trabajos anteriores pueden despertar dudas legítimas, pero ninguna de estas características es exclusiva de la generación mediante inteligencia artificial. Un estudiante puede mejorar su escritura, recibir asistencia legítima, utilizar un corrector, dedicar más tiempo a una actividad o modificar deliberadamente su registro. Del mismo modo, una inteligencia artificial puede ser instruida para producir errores o imitar determinados niveles de competencia. La inferencia desde una característica estilística hacia una acusación de mala conducta es, por tanto, débil si no existe información adicional.

La investigación reciente aporta evidencia empírica especialmente relevante para esta distinción. Munoz et al. (2026) analizaron 1.162 expedientes institucionales relacionados con sospechas de uso indebido de inteligencia artificial generativa registrados entre 2023 y 2025 en una universidad australiana. A partir de 1.855 elementos de evidencia, los autores desarrollaron una taxonomía de quince tipos y evaluaron su calidad mediante tres dimensiones: relevancia, credibilidad y fuerza inferencial. Los resultados mostraron que las diferentes señales no poseen el mismo valor probatorio y que los resultados de detectores se situaban entre las evidencias más débiles, mientras que determinadas admisiones del estudiante presentaban una fuerza considerablemente mayor.

El estudio también encontró un incremento con el tiempo en la utilización de referencias fabricadas y admisiones de uso como elementos de los procedimientos, mientras disminuían las acusaciones presentadas sin apoyo documental. Estos resultados son importantes porque muestran una

evolución institucional desde la sospecha genérica hacia intentos de reunir evidencia verificable. No obstante, los autores advierten que las instituciones continúan careciendo de estándares suficientemente estructurados para determinar qué evidencia es adecuada en los casos relacionados con inteligencia artificial (Munoz et al., 2026).

La fabricación de referencias ilustra bien esta diferencia. Si una referencia citada en un trabajo no existe, su inexistencia puede comprobarse mediante fuentes bibliográficas. Sin embargo, este hallazgo demuestra que existe un problema con la referencia, no necesariamente que una inteligencia artificial la haya generado. El estudiante podría haberla fabricado directamente o haber cometido otro tipo de error. Por ello, incluso una señal verificable debe interpretarse con relación a la conducta que realmente permite demostrar.

Los historiales de versiones, borradores, notas, metadatos y registros de elaboración pueden aportar contexto adicional cuando se encuentran disponibles y su utilización respeta las normas institucionales de privacidad. Un proceso de escritura progresivo puede ser coherente con autoría humana, aunque su ausencia tampoco demuestra automáticamente generación automatizada. No todos los estudiantes trabajan mediante sistemas que conservan versiones y exigir retrospectivamente evidencias que nunca se establecieron como requisito puede trasladar injustamente la carga de la prueba.

Una conversación académica o defensa oral puede aportar información complementaria cuando existen dudas fundadas. Solicitar al estudiante que explique los argumentos, decisiones metodológicas, fuentes o procedimientos empleados permite valorar si comprende el contenido presentado. Sin embargo, esta estrategia también debe aplicarse con criterios claros. El desempeño oral puede verse afectado por ansiedad, diferencias lingüísticas, discapacidad u otras variables, por lo que una dificultad para responder una pregunta aislada no debería interpretarse automáticamente como demostración de fraude.

La respuesta institucional necesita considerar además el estándar de prueba aplicable. En diversos sistemas universitarios, las investigaciones disciplinarias utilizan el criterio de balance de probabilidades, que exige determinar si, a partir de la evidencia disponible, resulta más probable que la conducta haya ocurrido que lo contrario. La Tertiary Education Quality

and Standards Agency ha señalado expresamente que un puntaje generado por un detector de IA es insuficiente por sí solo para formular o sostener una acusación y recomienda buscar evidencia adicional, incluyendo información que pueda refutar la hipótesis inicial y reducir el sesgo de confirmación (Tertiary Education Quality and Standards Agency [TEQSA], 2026).

La tabla siguiente sintetiza el papel que distintos indicios pueden desempeñar dentro de una investigación. No constituye una escala universal de culpabilidad ni sustituye las normas institucionales; su finalidad es evitar que señales con funciones diferentes sean tratadas como si poseyeran el mismo valor.

Tabla 2. Interpretación prudente de señales relacionadas con posible uso no autorizado de inteligencia artificial

SEÑAL ELEMENTO	O QUÉ PUEDE JUSTIFICAR	PRINCIPAL LIMITACIÓN	USO RAZONABLE DENTRO DEL PROCEDIMIENT O
RESULTADO DE UN DETECTOR DE IA	Revisión adicional del trabajo	Es probabilístic o, depende de la herramienta y puede fallar en textos humanos, híbridos o modificados	Señal inicial que requiere corroboración independiente
CAMBIO NOTABLE DE ESTILO O NIVEL LINGÜÍSTICO	Comparación contextual con otros trabajos	La variación estilística tiene múltiples	Motivo para indagar, no prueba de uso de IA

		explicaciones legítimas		
REFERENCIAS INEXISTENTES O DATOS FABRICADOS		Comprobación objetiva de una irregularidad	No demuestra por sí sola quién o qué generó la información	Evidencia directa del problema documental y señal contextual sobre su posible origen
HISTORIALES DE VERSIONES, BORRADORES O REGISTROS DEL PROCESO		Reconstrucción parcial de cómo se produjo el trabajo	Pueden ser incompletos y no siempre fueron requeridos previamente	Evidencia contextual complementaria cuando se obtiene legítimamente
CAPACIDAD PARA EXPLICAR Y DEFENDER EL TRABAJO		Valoración de comprensión, decisiones y dominio de fuentes	El desempeño oral puede verse afectado por factores distintos de la autoría	Corroboración contextual dentro de un procedimiento estructurado
REGISTRO DE PROMPTS O INTERACCIONES APORTADO DE MANERA VÁLIDA		Identificación del tipo de asistencia utilizada	Puede ser parcial y requiere conocer si ese uso estaba permitido	Evidencia directa de determinadas interacciones, interpretada según las reglas de la actividad
RECONOCIMIENTO O EXPLÍCITO DE UN USO NO AUTORIZADO		Vinculación directa entre la conducta y el estudiante	Es necesario establecer exactamente qué uso	Evidencia de alta relevancia cuando la declaración es clara y obtenida

	reconoce y mediante un qué regla se procedimiento aplicaba justo
--	--

Nota. Elaboración propia a partir de Bassett et al. (2026), Hadra et al. (2026), Munoz et al. (2026), TEQSA (2026), Van Vlasselaer et al. (2026) y Weber-Wulff et al. (2023).

Una investigación adecuada debería considerar tanto la evidencia que apoya una hipótesis como aquella que puede contradecirla. Este punto es relevante para reducir el sesgo de confirmación. Una vez que un docente sospecha que un trabajo fue producido mediante inteligencia artificial, puede comenzar a interpretar características ambiguas como confirmación de su sospecha inicial. Un procedimiento más riguroso exige preguntarse también qué evidencia sería esperable si el trabajo fuera legítimo y si esa evidencia se encuentra presente.

La equidad procedimental exige igualmente que el estudiante conozca la acusación y pueda responder a la evidencia utilizada. Un sistema en el que el docente recibe una puntuación algorítmica que el estudiante no puede interpretar, reproducir ni cuestionar genera problemas de transparencia. Bassett et al. (2026) advierten que este escenario puede desplazar de facto la carga de la prueba hacia el estudiante, quien termina intentando demostrar que no utilizó una herramienta en lugar de que la institución demuestre que una infracción efectivamente ocurrió.

Los resultados de Munoz et al. (2026) refuerzan esta preocupación. Su análisis de casos reales muestra que la calidad probatoria varía sustancialmente entre tipos de evidencia y que los procedimientos institucionales todavía están desarrollando criterios consistentes para evaluarla. En consecuencia, una institución no debería limitarse a establecer qué herramientas están prohibidas; también necesita determinar cómo se investigará una posible infracción y qué condiciones deben cumplirse antes de adoptar una decisión disciplinaria.

La prevención resulta, por esta razón, más sólida que una política centrada exclusivamente en detección. Cuando las condiciones de una actividad especifican claramente qué usos están permitidos, qué usos están

prohibidos y qué formas de asistencia deben declararse, disminuye la ambigüedad tanto para el estudiante como para el docente. La claridad normativa no elimina la posibilidad de fraude, pero permite diferenciar con mayor precisión entre una práctica autorizada y una infracción.

La formación también desempeña un papel preventivo. Los estudiantes necesitan comprender que el problema de la integridad no consiste solamente en evitar ser detectados. Presentar información no verificada, inventar referencias, ocultar una contribución sustantiva de una herramienta o delegar una competencia que debía demostrarse afecta la validez de su propio aprendizaje. Esta explicación resulta pedagógicamente más sólida que una política basada únicamente en la amenaza de sanción.

Los docentes requieren, a su vez, formación sobre las capacidades y limitaciones de los sistemas generativos y de las herramientas destinadas a detectarlos. La sobreconfianza constituye un riesgo en ambos sentidos. Sobreestimar la capacidad de la inteligencia artificial puede conducir a asumir que cualquier texto bien escrito fue generado; sobreestimar los detectores puede transformar una clasificación probabilística en una aparente certeza.

Las consideraciones de equidad son especialmente relevantes en contextos educativos con diversidad lingüística. El hallazgo de Liang et al. (2023) sobre falsos positivos en textos de hablantes no nativos de inglés mostró que determinados patrones lingüísticos pueden confundirse con características atribuidas a los modelos generativos. Hadra et al. (2026) encontraron igualmente razones para mantener cautela en entornos con estudiantes de inglés como lengua extranjera. Aunque estos resultados corresponden específicamente al inglés, el principio general resulta importante para América Latina: las herramientas deben ser evaluadas en las poblaciones y lenguas en las que pretenden utilizarse y no asumir que un desempeño obtenido en otro contexto se mantiene automáticamente.

En Ecuador y otros sistemas latinoamericanos, adoptar detectores comerciales sin validación contextual podría generar una falsa sensación de seguridad. La evidencia ecuatoriana disponible muestra que las relaciones entre utilización de IA y conductas de integridad son complejas y dependen de factores relacionados con conocimiento, percepción y forma de utilización (Torres-Díaz et al., 2025). No existe fundamento para

asumir que una estrategia principalmente tecnológica resolverá un fenómeno que incluye dimensiones pedagógicas, éticas e institucionales.

El objetivo de una política de integridad tampoco debería ser demostrar la mayor cantidad posible de infracciones. Un sistema de investigación legítimo debe distinguir entre prevenir conductas indebidas, identificar situaciones que justifican revisión y determinar mediante evidencia suficiente si una norma fue vulnerada. Confundir estos niveles favorece tanto falsos negativos, en los que una infracción pasa inadvertida, como falsos positivos, en los que un estudiante legítimo es acusado incorrectamente.

La inteligencia artificial generativa obliga, por tanto, a abandonar la expectativa de una solución de detección completamente automática. La evidencia actual muestra avances técnicos importantes, pero también una realidad más compleja: la escritura puede ser humana, generada, editada, híbrida o transformada mediante múltiples herramientas. La clasificación de un texto no puede sustituir el análisis de la conducta académica que produjo ese texto.

La respuesta más consistente combina claridad previa, alfabetización, procedimientos transparentes, análisis contextual y proporcionalidad. Cuando aparece una sospecha, el objetivo inicial debería ser comprender cómo fue elaborado el trabajo y reunir evidencia verificable, no confirmar una conclusión previamente asumida. Solo después puede determinarse si la conducta observada contradice las condiciones que habían sido comunicadas al estudiante.

Este enfoque permite mantener una distinción esencial entre detección e integridad. Detectar una característica tecnológica no equivale a demostrar deshonestidad, del mismo modo que no detectar inteligencia artificial no demuestra que el trabajo sea auténtico. La integridad académica requiere evaluar la relación entre las reglas establecidas, las acciones realizadas y el aprendizaje que el producto pretende representar. Esta limitación de los enfoques centrados exclusivamente en control conduce directamente al siguiente problema del libro: si los sistemas educativos no pueden depender de la detección para garantizar la autenticidad del desempeño, resulta necesario reconsiderar cómo se diseñan las evaluaciones. Esa cuestión constituye el eje del capítulo siguiente.

Capítulo 3

Evaluación educativa en contextos mediados por inteligencia artificial

La expansión de la inteligencia artificial generativa ha colocado a la evaluación educativa ante una pregunta más profunda que la identificación del uso indebido de una herramienta: ¿qué puede afirmarse realmente sobre el aprendizaje de un estudiante a partir del producto que entrega? Esta cuestión existía antes de la aparición de los modelos generativos, pero su presencia se ha vuelto más visible cuando sistemas accesibles pueden producir ensayos, resolver problemas, generar código, sintetizar información, elaborar presentaciones o construir explicaciones que, durante décadas, fueron utilizadas como indicadores del desempeño individual.

La evaluación no consiste simplemente en asignar una calificación a un producto. Su función implica obtener información que permita formular inferencias sobre conocimientos, habilidades, procesos de razonamiento o competencias. Desde esta perspectiva, el desafío que introduce la inteligencia artificial no se limita a determinar si fue utilizada, sino a establecer si las evidencias reunidas continúan siendo suficientes para sostener las conclusiones que se extraen acerca del aprendizaje.

Este desplazamiento resulta importante porque evita convertir la evaluación en un mecanismo de vigilancia tecnológica. Si el problema se formula únicamente en términos de impedir que los estudiantes utilicen inteligencia artificial, la respuesta tenderá hacia controles, prohibiciones o estrategias de detección. Si se formula como un problema de validez de las evidencias, la pregunta cambia: se trata de diseñar situaciones en las que aquello que el estudiante hace permita conocer con suficiente fundamento qué ha aprendido, incluso en un entorno donde la IA forma parte de la realidad educativa y profesional.

La investigación reciente apunta en esa dirección. Lodge et al. (2023) sostienen que la inteligencia artificial generativa ha intensificado

debilidades previamente existentes en determinados modelos de evaluación y que la respuesta no debería consistir únicamente en intentar proteger tareas tradicionales frente a la tecnología. Kizilcec et al. (2024), en un estudio con 680 estudiantes y 87 docentes universitarios de Australia, Chipre y Estados Unidos, encontraron que los participantes consideraban especialmente vulnerables tareas como ensayos y programación, mientras los docentes mostraban mayor preferencia por adaptar las evaluaciones para incorporar la inteligencia artificial y enfatizar pensamiento de orden superior y aplicaciones auténticas.

El reto, entonces, no es construir evaluaciones imposibles de resolver con inteligencia artificial, objetivo que probablemente se vuelva cada vez menos sostenible conforme evolucionen los sistemas, sino preservar la capacidad de la evaluación para proporcionar evidencia válida y útil del aprendizaje.

Evaluación educativa y validez de las evidencias de aprendizaje

La evaluación educativa cumple distintas funciones. Puede certificar aprendizajes, proporcionar retroalimentación, identificar dificultades, orientar decisiones pedagógicas, permitir la autorregulación y determinar si determinados resultados educativos han sido alcanzados. Estas funciones pueden coexistir, pero no siempre utilizan la misma información ni requieren idénticos instrumentos.

Una prueba escrita, una presentación oral, un proyecto, una práctica de laboratorio o un portafolio producen evidencias diferentes. Ninguna modalidad es intrínsecamente válida por el simple hecho de haber sido utilizada durante años. Su valor depende de la relación entre aquello que se pretende evaluar, la actividad que realiza el estudiante y la interpretación que posteriormente se construye a partir de su desempeño.

La tradición contemporánea sobre validez ha insistido precisamente en que esta no constituye una propiedad fija del instrumento. Messick (1993) planteó una concepción unificada en la que la validez depende de las evidencias y fundamentos que respaldan las interpretaciones y usos derivados de una evaluación. Desde esta perspectiva, una actividad puede

estar bien construida técnicamente y, sin embargo, proporcionar información insuficiente para sostener una determinada inferencia.

Esta idea adquiere particular importancia ante la inteligencia artificial generativa. Si un estudiante presenta un ensayo correctamente estructurado, la existencia del producto permite observar la calidad del texto, pero no necesariamente permite concluir que el estudiante posee por sí mismo todas las capacidades empleadas para producirlo. La posibilidad de intervención de sistemas generativos introduce una fuente adicional de incertidumbre sobre el origen de las operaciones cognitivas reflejadas en el trabajo.

El problema no es exclusivo de la escritura. Una aplicación puede producir código funcional sin que quien lo presenta comprenda completamente su lógica, resolver determinadas operaciones matemáticas, formular diagnósticos diferenciales, traducir textos o desarrollar argumentos complejos. En cada caso, la pregunta evaluativa debe dirigirse hacia la relación entre el resultado observado y la competencia que se desea inferir.

Esta relación permite distinguir dos cuestiones que frecuentemente se confunden. Una respuesta puede ser correcta sin constituir evidencia suficiente de aprendizaje. Del mismo modo, un producto imperfecto puede revelar procesos cognitivos valiosos si permite observar cómo el estudiante identifica un problema, prueba alternativas, corrige errores o justifica una decisión. La evaluación pierde información cuando concentra todo su peso en el resultado final e ignora el proceso que permitió alcanzarlo.

Lodge et al. (2023) proponen que los sistemas de evaluación contemporáneos deben ofrecer representaciones más ricas del aprendizaje y no depender de una única tarea para asegurar que los estudiantes han alcanzado los resultados previstos. Esta idea se ha desarrollado con mayor profundidad en trabajos posteriores. Lodge et al. (2026) señalan que asegurar la calidad del aprendizaje en un entorno integrado con inteligencia artificial requiere fortalecer capacidades adaptativas, entre ellas pensamiento crítico, juicio evaluativo y razonamiento ético, porque estas permiten responder a situaciones cambiantes que no pueden resolverse únicamente mediante procedimientos rutinarios.

El concepto de juicio evaluativo ocupa aquí un lugar importante. Tai et al. (2018) lo definen como la capacidad para tomar decisiones sobre la calidad del propio trabajo y del trabajo de otros. No se limita a reconocer si una respuesta es correcta o incorrecta; supone comprender criterios de calidad, aplicarlos en situaciones concretas y ajustar el propio desempeño a partir de ellos.

La inteligencia artificial incrementa la importancia de esta capacidad porque facilita la producción de respuestas que pueden parecer correctas antes de ser examinadas críticamente. Bearman et al. (2024) argumentan que los estudiantes necesitan aprender a evaluar tanto los productos generados por IA como los procesos mediante los cuales interactúan con estos sistemas. La facilidad para producir contenido amplía la distancia entre la capacidad de generar una respuesta y la capacidad de juzgar si dicha respuesta posee calidad suficiente.

Esta diferencia altera el valor educativo de algunas actividades. Si una herramienta puede producir rápidamente una explicación plausible, la competencia relevante puede desplazarse desde la simple elaboración de la explicación hacia la capacidad de examinarla, contrastarla con evidencia, identificar supuestos, detectar errores y justificar una versión mejorada. Esto no significa que producir explicaciones deje de ser importante. Un estudiante necesita conocimiento suficiente para realizar una evaluación crítica; de lo contrario, puede aceptar una respuesta incorrecta porque su forma lingüística resulta convincente.

La evaluación formativa ofrece un marco útil para comprender este proceso. Black y Wiliam (1998) mostraron que la evaluación puede contribuir al aprendizaje cuando genera información que permite modificar la enseñanza y las acciones del estudiante. Desde esta perspectiva, evaluar no significa esperar hasta el final de una unidad para comprobar resultados, sino producir información durante el proceso que permita reconocer la distancia entre el desempeño actual y los objetivos esperados.

La presencia de inteligencia artificial puede ampliar algunas posibilidades de retroalimentación, pero no sustituye la función pedagógica de la evaluación formativa. Una respuesta automatizada puede proporcionar comentarios inmediatos, sugerir modificaciones o generar ejemplos; sin embargo, el aprendizaje depende de lo que el estudiante haga con esa información. Carless y Boud (2018) utilizan el concepto de alfabetización

para la retroalimentación para describir las capacidades necesarias para comprender los comentarios recibidos, emitir juicios y convertir la información en acciones concretas de mejora.

En un entorno con inteligencia artificial, esta capacidad cobra un significado adicional. Obtener retroalimentación se vuelve relativamente fácil; valorarla correctamente continúa siendo una actividad cognitiva exigente. Un estudiante puede recibir múltiples sugerencias generadas automáticamente sin desarrollar la capacidad de distinguir cuáles son pertinentes. La abundancia de feedback no garantiza su apropiación ni su calidad.

La validez de la evaluación debe analizarse, entonces, en relación con las evidencias que realmente produce. Si el objetivo es comprobar dominio conceptual individual, será necesario disponer de situaciones en las que ese dominio pueda observarse con una participación tecnológica limitada o suficientemente controlada. Si se pretende evaluar la capacidad para trabajar responsablemente con inteligencia artificial, impedir su utilización sería incoherente con el propio resultado de aprendizaje.

El criterio central no es maximizar o minimizar el uso de tecnología. La cuestión consiste en lograr correspondencia entre objetivo, actividad, evidencia e interpretación.

Limitaciones de los modelos tradicionales de evaluación ante la IA generativa

La aparición de inteligencia artificial generativa ha generado críticas hacia las evaluaciones denominadas tradicionales, aunque el término puede resultar demasiado amplio. No todas las pruebas escritas, ensayos o exámenes poseen los mismos objetivos ni presentan idéntica vulnerabilidad. Un error sería sustituir una respuesta simplista basada en la prohibición de IA por otra que declare obsoletos todos los instrumentos convencionales.

Los exámenes supervisados, por ejemplo, pueden seguir siendo apropiados cuando se necesita comprobar que un estudiante dispone de conocimientos o habilidades que debe ejecutar sin asistencia externa. En

medicina, matemáticas, idiomas o determinadas áreas técnicas existen conocimientos fundamentales cuya disponibilidad individual conserva valor. El problema aparece cuando un mismo instrumento se utiliza de manera rutinaria para evaluar competencias que requieren transferencia, integración, juicio o resolución de situaciones complejas.

La inteligencia artificial expone especialmente la fragilidad de actividades en las que el producto puede generarse sin necesidad de comprender el proceso. Un ensayo genérico sobre un tema ampliamente documentado, una síntesis descriptiva, una respuesta basada en definiciones o un ejercicio altamente predecible pueden ser realizados con facilidad creciente por sistemas generativos. En estos casos, intentar preservar la tarea exactamente como estaba diseñada y concentrar los esfuerzos en detectar la intervención tecnológica puede resultar menos productivo que revisar qué evidencia de aprendizaje se pretendía obtener.

Kizilcec et al. (2024) encontraron que docentes y estudiantes percibían un impacto desigual de la inteligencia artificial según el tipo de evaluación. Los ensayos y las tareas de programación fueron considerados particularmente afectados, mientras otros formatos presentaban menor vulnerabilidad. Los docentes participantes tendían a preferir adaptaciones que asumieran la existencia de la IA y desplazaran la evaluación hacia pensamiento crítico, autenticidad y procesos de aprendizaje.

Esta diferencia cuestiona la idea de una solución uniforme. Una prueba de conocimientos básicos, una investigación, un proyecto de diseño y una práctica profesional simulada no deberían responder del mismo modo a la presencia de inteligencia artificial. La decisión debe comenzar con una pregunta curricular: ¿qué aprendizaje se pretende demostrar?

La evaluación tradicional también puede presentar un problema de fragmentación. Cuando cada tarea se interpreta de manera aislada, una institución puede intentar proteger individualmente ensayos, exámenes o proyectos sin analizar si, en conjunto, el programa proporciona suficientes oportunidades para verificar los resultados de aprendizaje. Lodge et al. (2023) proponen una visión programática en la que la confianza sobre el aprendizaje se construye mediante múltiples evidencias obtenidas en distintos momentos y condiciones.

Este enfoque reduce la dependencia de una única actividad. Un estudiante puede redactar un informe con acceso autorizado a inteligencia artificial y posteriormente explicar oralmente sus decisiones, resolver una situación semejante sin asistencia o aplicar los mismos conceptos a un caso diferente. Ninguna de estas evidencias necesita ser considerada definitiva por separado. La convergencia entre ellas proporciona una representación más completa del aprendizaje.

La noción de evaluación sostenible desarrollada por Boud y Soler (2016) aporta otro argumento relevante. Los autores sostienen que la evaluación debería responder a las necesidades presentes del curso y, al mismo tiempo, preparar a los estudiantes para valorar su propio desempeño en situaciones futuras. Esta concepción adquiere especial pertinencia cuando las herramientas tecnológicas cambian con rapidez. Diseñar una evaluación únicamente para resistir la versión actual de un sistema generativo puede producir soluciones de corta duración.

La alternativa es fortalecer capacidades que mantengan valor incluso cuando cambien las herramientas. Evaluar la capacidad para justificar decisiones, reconocer limitaciones, comparar evidencia, revisar errores y valorar calidad puede resultar más sostenible que diseñar tareas alrededor de una función específica que hoy una tecnología no puede realizar pero probablemente podrá ejecutar en poco tiempo.

Esto no significa eliminar la memoria, los conocimientos fundamentales ni las habilidades básicas. La crítica a las evaluaciones tradicionales puede caer en otro extremo si interpreta que, dado que la información está disponible mediante tecnología, ya no es necesario aprenderla. El pensamiento crítico no ocurre en ausencia de conocimiento. Evaluar una respuesta de IA requiere comprender el campo suficiente para detectar errores, reconocer matices y distinguir información sólida de una formulación superficial.

Bearman et al. (2024) destacan precisamente este punto al vincular inteligencia artificial y juicio evaluativo. Cuanto mayor sea la capacidad tecnológica para producir resultados plausibles, mayor será la necesidad de que los estudiantes desarrollen criterios propios de calidad. La evaluación debe proporcionar oportunidades para construir esos criterios y no limitarse a comprobar si el producto final cumple una forma determinada.

La discusión también afecta a la calificación. Una actividad diseñada cuando la elaboración de un determinado producto implicaba necesariamente varias horas de trabajo intelectual puede cambiar de significado cuando parte sustancial del proceso puede automatizarse. Mantener idéntica rúbrica puede conducir a valorar propiedades del producto que ya no representan el mismo esfuerzo cognitivo.

Por ejemplo, otorgar una proporción elevada de la calificación exclusivamente por estructura formal, corrección lingüística o presencia de ciertos componentes puede resultar menos informativo cuando estos elementos pueden producirse automáticamente. Esto no implica que dejen de importar, sino que el peso evaluativo puede necesitar desplazarse hacia la calidad del razonamiento, la elección de evidencia, la justificación de decisiones, la capacidad de revisión y la comprensión demostrada.

Existe también un riesgo en responder a la IA mediante un retorno indiscriminado a exámenes presenciales cerrados. Estas evaluaciones pueden tener una función válida para determinados objetivos, pero convertirlas en respuesta predominante puede reducir oportunidades para evaluar competencias complejas. Un examen bajo supervisión puede demostrar que un estudiante recuerda determinados contenidos; difícilmente, por sí solo, mostrará cómo gestiona un proyecto prolongado, integra fuentes contradictorias, colabora o utiliza de forma crítica herramientas disponibles en su futura práctica profesional.

El desafío consiste en combinar formas de evaluación que permitan asegurar capacidades individuales con otras que reflejen entornos reales donde las tecnologías sí estarán disponibles. Lodge et al. (2025) describen esta necesidad mediante estrategias que combinan evidencias protegidas de aprendizaje con actividades en las que el uso de inteligencia artificial puede incorporarse deliberadamente.

Este equilibrio resulta más defensible que intentar elegir entre dos modelos absolutos. Una educación que nunca permita inteligencia artificial puede preparar insuficientemente para contextos profesionales en los que su utilización será común. Una educación que permita delegar cualquier tarea desde las primeras etapas puede debilitar el desarrollo de conocimientos y habilidades necesarios para supervisar esas mismas tecnologías.

La secuencia formativa importa. No es equivalente utilizar una herramienta después de haber desarrollado una competencia básica que emplearla para sustituir el proceso mediante el cual esa competencia debía adquirirse.

Evaluación del producto y evaluación del proceso de aprendizaje

El producto final ha ocupado históricamente un lugar central en numerosos sistemas de evaluación. Se entrega un ensayo, una presentación, un programa informático, un informe o una resolución y el docente juzga su calidad. Este procedimiento es eficiente y, en muchos contextos, razonable. Sin embargo, cuando herramientas externas pueden intervenir de manera significativa en la producción del resultado, observar únicamente el producto puede ofrecer una visión incompleta del aprendizaje.

Evaluar el proceso no significa registrar cada acción del estudiante ni implantar sistemas de vigilancia. Implica obtener evidencias de determinados momentos del razonamiento que resulten relevantes para comprender cómo fue construido el trabajo. Borradores, decisiones justificadas, revisiones, comentarios sobre errores, defensa del producto, comparación entre versiones o reflexiones sobre estrategias pueden proporcionar información que el resultado terminado oculta.

La atención al proceso tampoco surge con la inteligencia artificial. La evaluación formativa, la autorregulación y la retroalimentación han defendido durante décadas la necesidad de involucrar al estudiante en la comprensión y mejora de su propio desempeño (Black & Wiliam, 1998; Boud & Soler, 2016). La novedad consiste en que la IA aumenta el valor de esas evidencias porque un producto pulido ya no permite inferir con la misma confianza cómo se llegó hasta él.

Un enfoque centrado en procesos puede incluir la observación de decisiones críticas. En un trabajo de investigación, por ejemplo, puede resultar más informativo pedir al estudiante que justifique por qué seleccionó determinadas fuentes que valorar únicamente la cantidad de referencias. En un problema clínico, explicar qué datos modificaron la

hipótesis diagnóstica puede aportar más información sobre razonamiento que presentar simplemente la opción final. En programación, describir por qué se descartaron determinadas soluciones permite conocer mejor la comprensión del problema.

La inteligencia artificial puede formar parte de estas evidencias cuando su uso se encuentra autorizado. Un estudiante puede documentar qué tipo de asistencia solicitó, qué respuesta recibió, qué problemas identificó y qué modificaciones realizó posteriormente. El foco deja de estar en demostrar que todo fue producido sin tecnología y pasa a analizar la calidad de las decisiones humanas tomadas durante la interacción.

Este enfoque requiere cautela para evitar una burocratización excesiva. Pedir capturas de cada conversación, registros completos de prompts y reflexiones extensas en todas las tareas puede aumentar la carga de estudiantes y docentes sin garantizar mejor aprendizaje. La trazabilidad debe ser proporcional al objetivo educativo.

El principio útil consiste en seleccionar puntos del proceso que permitan observar operaciones cognitivas relevantes. Si la competencia evaluada es argumentación, puede solicitarse una justificación de la tesis y de los principales cambios realizados. Si se evalúa resolución de problemas, interesa conocer cómo se interpretó la situación y por qué se eligió determinado procedimiento. Si se permite IA, puede preguntarse qué resultado fue aceptado, rechazado o corregido y con qué fundamento.

Bearman et al. (2024) proponen que el juicio evaluativo frente a la inteligencia artificial debe dirigirse tanto a los productos como a los procesos. No basta con decidir si una respuesta generada es buena; el estudiante necesita valorar también la calidad de su interacción con la herramienta, las instrucciones utilizadas, las verificaciones realizadas y las decisiones mediante las cuales incorpora o descarta sus resultados.

La investigación de Anders y Dux Speltz (2025) ofrece una aproximación empírica a esta cuestión. En un estudio exploratorio con 38 estudiantes universitarios, los autores integraron alfabetización en IA y aprendizaje autorregulado mediante un ciclo de planificación, iteración y evaluación. Encontraron mejoras en la autoeficacia percibida para alfabetización en IA y describieron prácticas en las que los estudiantes activaban conocimiento disciplinar, establecían criterios, refinaban sus interacciones y evaluaban los

productos generados. Por el tamaño y diseño del estudio, estos resultados no deben generalizarse de manera amplia, pero muestran una vía de integración en la que la interacción con IA se convierte en objeto de reflexión en lugar de permanecer como un proceso invisible.

La evaluación del proceso puede favorecer también la metacognición. Cuando el estudiante debe explicar cómo aprendió, qué estrategias utilizó, dónde encontró dificultades o por qué cambió de decisión, se le obliga a observar su propio funcionamiento cognitivo. Esta información puede servir tanto al docente como al propio estudiante.

La reflexión, sin embargo, tampoco debe idealizarse. Un texto reflexivo puede ser superficial, producido para satisfacer una rúbrica o generado mediante inteligencia artificial. Por ello, incorporar una sección denominada “reflexión” no garantiza automáticamente evidencia del proceso. La calidad depende de las preguntas formuladas y de su conexión con decisiones concretas del trabajo.

Preguntas genéricas como “¿qué aprendiste?” suelen generar respuestas previsibles. Resulta más informativo pedir que se identifique una decisión que cambió durante el proceso, una respuesta de IA que fue rechazada y la razón de hacerlo, una fuente que modificó la interpretación inicial o un error que reveló una comprensión incorrecta.

La defensa oral puede cumplir una función complementaria. Solicitar que el estudiante explique partes seleccionadas de su trabajo permite explorar comprensión y razonamiento de manera interactiva. No debería convertirse en un interrogatorio destinado exclusivamente a descubrir fraude, porque su valor educativo es mayor cuando forma parte prevista del diseño de la evaluación.

Existen limitaciones prácticas. Las defensas individuales demandan tiempo y pueden resultar difíciles en grupos numerosos. En estos casos pueden utilizarse muestras aleatorias de estudiantes, breves conversaciones grupales, presentaciones parciales, preguntas individualizadas o actividades de aplicación posterior. Lo importante no es que todos los productos sean acompañados por una defensa extensa, sino que el sistema de evaluación disponga de más de una fuente para interpretar el aprendizaje.

La distinción entre producto y proceso tampoco implica abandonar el primero. En muchos campos, producir un resultado de calidad sigue siendo una competencia importante. El objetivo consiste en evitar que ese resultado sea tratado como representación completa del aprendizaje cuando existen razones para pensar que múltiples agentes y herramientas pudieron intervenir en su elaboración.

Una evaluación robusta puede valorar ambos niveles. El producto permite juzgar qué se consiguió; el proceso ayuda a comprender cómo se consiguió y quién tomó las decisiones que importan.

Evaluación auténtica, pensamiento crítico y resolución de problemas

La evaluación auténtica ha adquirido nueva relevancia ante la inteligencia artificial, aunque sus fundamentos son anteriores a esta tecnología. Su propósito consiste en acercar la actividad evaluativa a problemas, desempeños y estándares significativos fuera del contexto estrictamente escolar o universitario.

Villarroel et al. (2018), a partir de una revisión de literatura, organizaron la evaluación auténtica en tres dimensiones principales: realismo, desafío cognitivo y juicio evaluativo. El realismo se relaciona con la semejanza entre la tarea académica y situaciones relevantes del mundo profesional o social. El desafío cognitivo implica utilizar conocimientos para resolver situaciones que requieren operaciones de orden superior. El juicio evaluativo supone desarrollar la capacidad de reconocer criterios de calidad y aplicarlos al propio trabajo.

Estas dimensiones resultan particularmente útiles frente a la IA porque desplazan la evaluación desde la reproducción de información hacia el uso razonado del conocimiento. Villarroel y Bruna (2019), desde el contexto latinoamericano, han señalado que numerosos instrumentos universitarios tradicionales presentan limitaciones para valorar conocimiento profundo cuando privilegian reproducción o aplicación descontextualizada. La evaluación auténtica busca conectar lo que ocurre en el aula con situaciones en las que el estudiante debe movilizar diferentes saberes para actuar.

No debe concluirse, sin embargo, que una tarea se vuelve resistente a la IA simplemente por denominarse auténtica. Un sistema generativo puede intervenir también en estudios de caso, proyectos y simulaciones. La autenticidad relevante no consiste únicamente en describir un contexto realista, sino en exigir decisiones situadas que requieran integrar información, justificar criterios y asumir responsabilidad por el resultado.

Un caso complejo puede ser más útil que una pregunta genérica si contiene información incompleta, elementos contradictorios o diversas soluciones defendibles. En estas situaciones, el valor de la respuesta depende menos de encontrar una frase correcta y más de construir una decisión razonada.

El pensamiento crítico adquiere aquí una función concreta. No debería aparecer en la evaluación como una competencia abstracta que se presume presente simplemente porque la tarea es difícil. Debe observarse en operaciones identificables: comparar evidencias, cuestionar supuestos, reconocer incertidumbre, valorar alternativas, detectar errores y justificar conclusiones.

La inteligencia artificial puede convertirse incluso en un objeto para ejercitar estas capacidades. Una actividad puede proporcionar al estudiante una respuesta generada y solicitarle que la analice, identifique problemas, contraste afirmaciones con fuentes fiables y construya una versión mejor fundamentada. En esta configuración, utilizar IA deja de representar una amenaza externa a la evaluación y se integra en el problema que debe resolverse.

Bearman et al. (2024) consideran que esta capacidad de juzgar calidad es esencial en un entorno donde resulta fácil producir contenido con apariencia de competencia. El estudiante necesita construir estándares internos que le permitan reconocer cuándo una respuesta es adecuada para un propósito determinado y cuándo requiere corrección.

El reto pedagógico consiste en evitar que estas tareas terminen evaluando solamente la capacidad de encontrar errores obvios. La crítica de una respuesta generada debe exigir conocimiento disciplinar. Una buena actividad puede incluir errores parcialmente plausibles, omisiones relevantes, fuentes de calidad desigual o decisiones que requieren ponderar consecuencias.

La resolución de problemas presenta una lógica semejante. Si la evaluación premia únicamente la respuesta final, una herramienta automatizada puede reducir sustancialmente la actividad intelectual del estudiante. Cuando se valoran representación del problema, selección de información, comparación de estrategias, justificación y revisión, aparecen más oportunidades para observar comprensión.

Esto resulta pertinente tanto en educación escolar como superior. En niveles escolares, una actividad auténtica no necesita reproducir literalmente una profesión. Puede partir de situaciones cercanas a la experiencia del estudiante que requieran aplicar conocimientos para interpretar fenómenos, tomar decisiones o explicar consecuencias. La complejidad debe adecuarse a la etapa de desarrollo.

El marco de competencias en IA para estudiantes de UNESCO propone precisamente una progresión basada en comprender, aplicar y crear, acompañada de dimensiones relacionadas con agencia humana, ética, técnicas de IA y diseño de sistemas (Miao et al., 2024). Esta estructura refuerza la idea de que las actividades escolares deben avanzar desde la comprensión básica hacia usos progresivamente más autónomos y críticos.

En educación superior, la autenticidad puede vincularse de manera más directa con prácticas profesionales. Un futuro médico puede analizar un caso y juzgar una recomendación generada; un estudiante de derecho puede contrastar argumentos y verificar jurisprudencia; uno de ingeniería puede evaluar propuestas técnicas; un estudiante de educación puede revisar un plan de clase generado y justificar sus modificaciones. En todos estos escenarios, la herramienta puede formar parte de la actividad sin reemplazar la competencia central que se pretende observar.

Existe, no obstante, un límite importante. La autenticidad profesional no significa replicar de manera acrítica cualquier práctica del mercado laboral. Que profesionales utilicen inteligencia artificial no implica que los estudiantes deban utilizarla desde el inicio para todas las tareas. La educación tiene una función formativa que precede a la automatización de determinadas operaciones.

Aprender a utilizar una herramienta profesionalmente exige primero desarrollar criterios que permitan supervisarla. Esta secuencia justifica que algunos momentos de evaluación restrinjan la IA mientras otros exijan

utilizarla. No existe contradicción si ambas decisiones responden a objetivos diferentes.

La evaluación auténtica tampoco debería convertirse en sinónimo de proyectos extensos. Un examen puede incorporar autenticidad si plantea situaciones contextualizadas que demandan decisiones complejas; un proyecto puede carecer de ella si se reduce a reunir información y presentarla de forma decorativa. La calidad del diseño depende de las operaciones cognitivas solicitadas y no del formato por sí solo.

Estrategias de evaluación y uso responsable de IA en educación escolar y superior

La respuesta educativa ante la inteligencia artificial requiere abandonar políticas de evaluación basadas en una única regla para todas las actividades. Algunas tareas necesitan demostrar desempeño individual sin asistencia; otras pueden permitir determinadas funciones; un tercer grupo puede incorporar deliberadamente la IA como parte del resultado de aprendizaje.

Perkins et al. (2024) propusieron la Artificial Intelligence Assessment Scale como marco para graduar diferentes niveles de utilización de IA en la evaluación. Su aporte principal no reside en establecer categorías permanentes, sino en reconocer que la participación tecnológica debe decidirse en función de los objetivos educativos. Esta lógica evita que la institución tenga que elegir entre prohibición general y autorización irrestricta.

Furze et al. (2024) estudiaron posteriormente la implementación del modelo en British University Vietnam. El piloto mostró una reducción de casos de mala conducta asociados con IA y cambios en las prácticas docentes, junto con un aumento de participación tecnológica en actividades diseñadas para ello. Estos resultados corresponden a una institución concreta y no deberían interpretarse como prueba de que el mismo modelo producirá idénticos efectos en cualquier contexto, pero muestran que comunicar niveles explícitos de utilización puede aportar claridad a estudiantes y docentes.

Las estrategias de evaluación pueden organizarse según la evidencia que se necesita obtener. Cuando interesa comprobar una capacidad individual básica, una actividad supervisada y sin IA puede ser adecuada. Cuando interesa observar desarrollo y autorregulación, puede ser preferible evaluar versiones sucesivas, decisiones y retroalimentación. Cuando el objetivo incluye trabajo profesional con herramientas generativas, la evaluación debería permitir o incluso exigir su utilización, pero valorar la calidad del juicio humano aplicado sobre sus resultados.

Lodge et al. (2025) plantean esta combinación mediante reformas que aseguren aprendizajes fundamentales y, simultáneamente, preparen a los estudiantes para utilizar inteligencia artificial de manera responsable. El objetivo no consiste en diseñar cada tarea para excluir la tecnología, sino en construir un sistema de evidencias suficientemente diverso para sostener las conclusiones sobre el aprendizaje.

La tabla siguiente sintetiza posibles decisiones de diseño. No representa una jerarquía en la que una estrategia sea superior a las demás. Cada una responde a un propósito distinto y puede resultar inadecuada si se utiliza para medir una competencia diferente de aquella para la que fue diseñada.

Tabla 3. Estrategias de evaluación según el propósito de aprendizaje y el papel de la inteligencia artificial generativa

<i>Estrategia</i>	<i>Evidencia principal que busca obtener</i>	<i>Papel posible de la IA</i>	<i>Ejemplo de aplicación</i>	<i>de Precaución de diseño</i>
<i>Evaluación individual supervisada</i>	Dominio personal de conocimientos o habilidades fundamentales	Restringido o no permitido	Resolución presencial de problemas, explicación conceptual, desempeño práctico	No utilizarla como única forma de evaluación de competencias complejas

<i>Evaluación del proceso</i>	Razonamiento, decisiones, revisión y autorregulación	Permitido de forma delimitada y documentada	Borradores sucesivos, bitácora de decisiones, justificación de cambios	Evitar transformar la trazabilidad en vigilancia o carga documental excesiva
<i>Evaluación auténtica</i>	Aplicación contextualizada de conocimientos y toma de decisiones	Variable según el escenario profesional o educativo	Caso, proyecto, simulación o resolución de un problema realista	El realismo superficial no garantiza desafío cognitivo
<i>Evaluación crítica de productos de IA</i>	Juicio evaluativo, verificación y pensamiento crítico	Necesario para realizar la actividad	Analizar, contrastar y corregir una respuesta generada	Requiere conocimiento disciplinar previo para que la crítica sea significativa
<i>Producción colaborativa humano IA</i>	Uso estratégico, ético y responsable de herramientas	Permitido o requerido	Crear un producto con IA y justificar las decisiones humanas realizadas	La rúbrica debe diferenciar calidad del producto y calidad del juicio del estudiante
<i>Evaluación</i>	Identificación de dificultades y	Permitido como fuente	Práctica guiada, preguntas de	La respuesta automatiza

<i>formativa con IA</i>	mejora progresiva	adicional de feedback	repaso, retroalimentación preliminar	da debe ser evaluada y no asumida como correcta
-----------------------------	----------------------	--------------------------	---	---

Nota. Elaboración propia a partir de Bearman et al. (2024), Furze et al. (2024), Lodge et al. (2023, 2025, 2026), Perkins et al. (2024) y Villarroel et al. (2018).

La combinación de estrategias puede resultar más sólida que la búsqueda de un formato único. Una asignatura podría incluir una evaluación individual destinada a comprobar conocimientos fundamentales, un proyecto en el que se permita inteligencia artificial y una defensa breve en la que el estudiante explique las decisiones adoptadas. Cada componente responde a una pregunta diferente sobre el aprendizaje.

La claridad de las instrucciones es indispensable. Una evaluación no debería limitarse a indicar “IA permitida” o “IA prohibida” cuando existen múltiples funciones posibles. Resulta más útil especificar qué tipo de asistencia puede utilizarse, qué actividad debe permanecer bajo responsabilidad del estudiante, qué información sobre la interacción debe declararse y qué criterios utilizará el docente.

Las rúbricas también necesitan revisión. Si la inteligencia artificial puede intervenir de manera legítima, los criterios deben valorar aquello que se espera del estudiante y no premiar indirectamente capacidades de la herramienta. Una rúbrica puede otorgar peso al análisis de evidencia, la justificación, la revisión crítica, la aplicación contextual o la capacidad de explicar decisiones. La corrección formal puede continuar siendo relevante, pero no necesariamente debe ocupar el mismo lugar que cuando era considerada un indicador más directo del desempeño individual.

La incorporación de inteligencia artificial a la evaluación plantea, al mismo tiempo, cuestiones de equidad. No todos los estudiantes poseen el mismo acceso a versiones comerciales, dispositivos, conectividad o experiencia de uso. Una tarea que obligue a utilizar herramientas avanzadas puede terminar midiendo diferencias de acceso en lugar de aprendizaje. UNESCO

(Miao & Holmes, 2023) ha advertido que las decisiones educativas sobre IA necesitan considerar inclusión, privacidad y capacidad institucional, no solamente las posibilidades técnicas de las aplicaciones.

En contextos escolares, estas condiciones adquieren mayor relevancia. La edad del estudiante, el nivel de autonomía, la protección de datos y el propósito curricular deben determinar cuándo y cómo se introduce una herramienta generativa. El marco de UNESCO plantea una progresión de competencias que puede orientar este proceso, pero insiste en que su aplicación debe ajustarse a la preparación local, la formación docente, los recursos disponibles y las características del sistema educativo (Miao et al., 2024).

Esto significa que una estrategia diseñada para estudiantes universitarios no debería trasladarse automáticamente a educación básica o bachillerato. Un alumno escolar puede necesitar mayor mediación docente, herramientas seleccionadas previamente y actividades más estructuradas. La autonomía en el uso de IA debe desarrollarse progresivamente, del mismo modo que otras formas de autonomía académica.

La evaluación formativa ofrece un espacio particularmente apropiado para esta progresión. Las actividades de bajo riesgo permiten experimentar, cometer errores y aprender a contrastar respuestas sin que cada interacción tenga consecuencias importantes sobre la calificación. La IA puede utilizarse para generar ejemplos, preguntas o explicaciones que posteriormente son revisadas con el docente o con los compañeros.

Este uso conserva sentido solo cuando la tecnología no desplaza la actividad que debe realizar el estudiante. Si una tarea formativa destinada a practicar una habilidad es completamente resuelta por la herramienta, la existencia de feedback inmediato resulta secundaria porque el proceso de práctica ha desaparecido.

En educación superior, la relación con el ámbito profesional adquiere mayor peso. Los programas formativos necesitan preguntarse qué usos de inteligencia artificial serán habituales o aceptables en las profesiones para las que preparan a sus estudiantes. La respuesta será distinta entre disciplinas y cambiará con el tiempo.

La preparación profesional no debería limitarse a enseñar cómo obtener respuestas eficientes. La creciente capacidad de los sistemas para producir contenido aumenta el valor de competencias como juicio evaluativo, pensamiento crítico y razonamiento ético. Bearman et al. (2024) sostienen que los seres humanos necesitan conservar su función como árbitros de calidad. Lodge et al. (2026) desarrollan una idea semejante al destacar capacidades adaptativas que permitan operar en contextos tecnológicos variables.

La integración responsable de IA en evaluación puede contribuir a formar esas capacidades si el estudiante debe tomar decisiones sobre cuándo utilizarla, cuándo desconfiar de una respuesta, cómo verificarla y cuándo resulta más apropiado resolver una tarea sin asistencia. La competencia no consiste en utilizar siempre inteligencia artificial, sino en decidir de manera fundada cuándo su utilización aporta valor.

Esta idea ayuda a superar la falsa oposición entre una evaluación tradicional completamente libre de IA y una evaluación futura dominada por herramientas generativas. Los sistemas educativos necesitan preservar espacios donde el estudiante demuestre conocimientos y habilidades propias y, al mismo tiempo, crear espacios donde aprenda a trabajar con tecnologías que probablemente encontrará fuera del aula.

El equilibrio dependerá del nivel formativo y del resultado esperado. En etapas iniciales puede ser necesario proteger determinados procesos para que se desarrollen antes de ser asistidos. Conforme aumenta el dominio, la tecnología puede incorporarse en situaciones que demanden supervisión, juicio y responsabilidad.

El rediseño tampoco debería realizarse únicamente desde cada asignatura. Si todos los docentes responden de manera independiente, los estudiantes pueden enfrentarse a normas contradictorias y las instituciones pueden terminar evaluando repetidamente algunas competencias mientras dejan otras sin comprobar. La visión programática propuesta por Lodge et al. (2023, 2025) permite distribuir las evidencias a lo largo de una trayectoria educativa y decidir dónde es imprescindible asegurar desempeño individual, dónde resulta apropiada la colaboración y dónde debe evaluarse explícitamente el uso de inteligencia artificial.

Una reforma de este tipo exige tiempo, formación y coordinación. La aparición de IA no convierte automáticamente en deficiente una práctica previa, ni obliga a modificar cada evaluación existente. El criterio debería ser diagnóstico: identificar qué inferencia se pretende realizar, examinar si la tarea proporciona todavía evidencia suficiente y modificarla cuando esa relación se haya debilitado.

Esta posición evita innovaciones meramente cosméticas. Cambiar un ensayo por un video, añadir una exposición oral o exigir una reflexión final no mejora necesariamente la evaluación si las nuevas actividades continúan midiendo superficialmente los mismos aspectos. El diseño debe comenzar por los resultados de aprendizaje y no por el formato.

La inteligencia artificial generativa ha hecho más visible un principio que la investigación sobre evaluación sostiene desde mucho antes de su aparición: obtener un producto no equivale a comprender el aprendizaje que lo produjo. La evaluación necesita reunir evidencias suficientemente ricas para conocer qué puede hacer el estudiante, cómo toma decisiones y en qué medida puede aplicar sus conocimientos en situaciones nuevas.

La respuesta educativa no consiste, entonces, en construir una barrera permanente entre estudiantes e inteligencia artificial. Tampoco en permitir que la tecnología intervenga indiscriminadamente en cada actividad. Consiste en decidir de manera pedagógicamente justificada qué debe aprenderse sin asistencia, qué puede aprenderse con ella y qué competencias surgen precisamente de la necesidad de trabajar críticamente con sistemas generativos.

En ese marco, la evaluación recupera una función que puede perderse cuando se concentra exclusivamente en calificar productos: orientar el aprendizaje. El valor de una actividad no radica solo en determinar si un estudiante merece una determinada puntuación, sino en crear condiciones para que construya conocimiento, desarrolle criterios de calidad y pueda transferirlos a situaciones futuras.

La inteligencia artificial transforma las condiciones en las que esta tarea se realiza, pero no modifica su finalidad fundamental. La calidad de la evaluación continuará dependiendo de la correspondencia entre aquello que una institución afirma enseñar, lo que solicita a sus estudiantes y la evidencia que utiliza para concluir que han aprendido.

Capítulo 4

Cultura institucional, políticas y uso responsable de la inteligencia artificial

La incorporación de inteligencia artificial generativa en educación no puede depender exclusivamente de las decisiones individuales de estudiantes y docentes. Los capítulos anteriores han mostrado que los límites entre asistencia legítima, delegación intelectual y deshonestidad académica dependen del propósito de las actividades, de las condiciones de evaluación y de la transparencia con la que se utilizan estas tecnologías. Si cada profesor establece criterios diferentes y cada estudiante interpreta por su cuenta qué prácticas son aceptables, la integridad académica queda expuesta a un sistema de normas fragmentado y, en ocasiones, contradictorio.

La respuesta institucional adquiere por ello una función decisiva. Las políticas educativas deben proporcionar orientación suficiente para proteger la integridad, la privacidad, la equidad y la calidad del aprendizaje, pero también necesitan conservar flexibilidad para responder a disciplinas, niveles educativos y objetivos pedagógicos diferentes. Una regulación demasiado general puede resultar poco útil en la práctica; una normativa excesivamente específica puede volverse obsoleta conforme cambian las capacidades de las herramientas.

La gobernanza de la inteligencia artificial supone un desafío más amplio que redactar una política de uso. Incluye definir quién toma decisiones, qué criterios orientan la incorporación de sistemas automatizados, qué datos pueden introducirse en ellos, cómo se protege a estudiantes y docentes, qué procedimientos existen cuando se producen errores, quién supervisa los resultados y de qué manera se revisan las políticas conforme aparecen nuevas evidencias y tecnologías.

Chan (2023) propone abordar este problema mediante un marco de política educativa de inteligencia artificial organizado en tres dimensiones: pedagógica, de gobernanza y operativa. La primera se concentra en los efectos sobre enseñanza y aprendizaje; la segunda comprende privacidad, seguridad y responsabilidad; la tercera considera infraestructura,

capacitación y recursos necesarios para implementar las decisiones institucionales. La utilidad de esta división reside en reconocer que disponer de principios éticos sin condiciones de aplicación resulta insuficiente, del mismo modo que contar con infraestructura tecnológica sin orientación pedagógica no garantiza una incorporación educativa adecuada.

Desde una perspectiva internacional, UNESCO (2021) sostiene que la gobernanza de la inteligencia artificial debe preservar derechos humanos, dignidad, diversidad, inclusión, transparencia, supervisión humana, privacidad y responsabilidad. Estos principios adquieren un significado particular en educación porque las decisiones tecnológicas afectan a personas que se encuentran en procesos formativos y, en los niveles escolares, a niños y adolescentes que requieren niveles superiores de protección.

El objetivo de una cultura institucional responsable no consiste, por tanto, en maximizar la adopción de inteligencia artificial ni en restringirla sistemáticamente. Consiste en crear condiciones para que su utilización responda a fines educativos identificables, pueda ser cuestionada cuando resulte problemática y mantenga a las personas como responsables de las decisiones que afectan el aprendizaje.

Políticas educativas e institucionales para el uso de IA generativa

Las primeras respuestas institucionales ante la expansión de la inteligencia artificial generativa fueron heterogéneas. Algunas instituciones establecieron prohibiciones, otras incorporaron advertencias en sus políticas de integridad académica y un tercer grupo comenzó a elaborar orientaciones sobre usos permitidos. Moorhouse et al. (2023), al analizar las directrices de universidades situadas en posiciones destacadas de clasificaciones internacionales, encontraron una disponibilidad todavía desigual de orientaciones institucionales durante la primera etapa de adopción de estas herramientas.

La rapidez con la que evolucionaron las tecnologías puso en evidencia una limitación de las políticas construidas exclusivamente alrededor de

plataformas concretas. Una norma que menciona únicamente ChatGPT puede perder vigencia cuando aparecen nuevos modelos o cuando herramientas que originalmente cumplían otras funciones incorporan generación de texto, imágenes, audio o código. Resulta más sostenible regular funciones y conductas que nombres comerciales.

Una política institucional debería comenzar por establecer principios comunes. Entre ellos se encuentran la primacía de los objetivos de aprendizaje, la responsabilidad humana sobre las decisiones, la transparencia en los usos relevantes de inteligencia artificial, la protección de datos, la equidad de acceso y la obligación de verificar información antes de incorporarla a actividades académicas. Estos principios pueden mantenerse relativamente estables aunque cambien las plataformas utilizadas.

Las reglas específicas deben situarse en niveles más cercanos a las actividades. No todas las asignaturas pueden utilizar las mismas condiciones. Una actividad destinada a comprobar dominio individual puede restringir determinadas herramientas, mientras otra orientada al desempeño profesional podría requerir explícitamente su utilización. Esta flexibilidad no equivale a ausencia de regulación. Exige que la institución defina qué decisiones corresponden al nivel central, cuáles pertenecen a programas o carreras y qué aspectos deben especificarse en cada asignatura.

Chan (2023), a partir de información obtenida de 457 estudiantes y 180 docentes y miembros del personal universitario, plantea que una política educativa de IA debe integrar dimensiones pedagógicas, de gobernanza y operativas. El estudio también destaca la necesidad de incluir a los estudiantes en los procesos de elaboración e implementación de políticas. Esta participación resulta relevante porque las normas pueden fracasar cuando son redactadas exclusivamente desde estructuras administrativas que desconocen las prácticas reales de uso.

Las políticas necesitan además distinguir entre orientación, obligación y sanción. Una recomendación para utilizar IA críticamente no tiene la misma naturaleza que una regla que prohíbe introducir información personal en un servicio externo. Tampoco debería tener las mismas consecuencias incumplir una recomendación pedagógica que vulnerar deliberadamente una condición evaluativa claramente comunicada.

La claridad de estas diferencias favorece la legitimidad institucional. Cuando las reglas son ambiguas, cambiantes o contradictorias, resulta más difícil exigir a los estudiantes que interpreten correctamente los límites. La institución tiene una responsabilidad previa a la sanción: comunicar de forma comprensible las condiciones aplicables.

Una política efectiva tampoco debería limitarse a definir lo que los estudiantes pueden hacer. Los docentes y las propias instituciones utilizan herramientas de inteligencia artificial para elaborar materiales, analizar información, producir retroalimentación o apoyar tareas administrativas. Si se exige transparencia a los estudiantes, resulta coherente establecer criterios equivalentes cuando la tecnología interviene en decisiones que los afectan.

La gobernanza requiere identificar responsabilidades. No todas las decisiones pueden recaer en el profesor de una asignatura ni todas deben centralizarse en autoridades institucionales. La siguiente tabla organiza una posible distribución.

Tabla 4. Niveles de responsabilidad para la gobernanza educativa de la inteligencia artificial

<i>Nivel</i>	<i>Responsabilidad principal</i>	<i>Decisiones que debería asumir</i>	<i>Evidencia de implementación</i>
<i>Institución o sistema educativo</i>	Establecer principios comunes, garantías y límites institucionales	Protección de datos, seguridad, contratación de herramientas, integridad académica, transparencia, accesibilidad y procedimientos ante incidentes	Política institucional, protocolos, evaluación periódica y mecanismos de rendición de cuentas

<i>Facultad, carrera o programa</i>	Adaptar los principios a las características disciplinares	Competencias en IA que deben desarrollarse, actividades profesionales en las que su uso resulta relevante y estrategias de evaluación	Acuerdos curriculares, mapas de competencias y orientaciones disciplinares
<i>Asignatura o docente</i>	Definir condiciones concretas de uso según resultados de aprendizaje	Qué herramientas o funciones se permiten, restringen o exigen y qué debe declarar el estudiante	Instrucciones de actividades, rúbricas y comunicación previa de criterios
<i>Estudiante</i>	Utilizar las herramientas dentro de las condiciones establecidas	Decidir cuándo utilizarlas, verificar sus resultados, reconocer su intervención y mantener responsabilidad sobre el trabajo presentado	Declaraciones de uso, productos académicos y capacidad de justificar decisiones
<i>Unidades técnicas, jurídicas y de protección de datos</i>	Evaluar riesgos tecnológicos y regulatorios	Seguridad, privacidad, tratamiento de datos, licencias, accesibilidad y cumplimiento normativo	Evaluaciones de riesgo, contratos, auditorías y protocolos de protección

<i>Comunidad educativa</i>	Participar en la revisión y legitimación de las políticas	Identificar problemas emergentes, desigualdades y efectos no previstos	Consultas, comisiones, revisión participativa y actualización de normas
----------------------------	---	--	---

Nota. Elaboración propia a partir de Chan (2023), Kanobel (2026), Miao y Holmes (2023) y UNESCO (2021).

Esta distribución evita uno de los errores más frecuentes en la integración educativa de tecnologías: trasladar toda la responsabilidad al docente. Un profesor puede decidir cómo utilizar IA en una actividad, pero difícilmente puede evaluar por sí solo las políticas de tratamiento de datos de cada proveedor, negociar condiciones institucionales de uso o garantizar igualdad de acceso a herramientas comerciales.

La institución tampoco debería determinar desde un nivel central qué porcentaje exacto de IA puede utilizarse en cada actividad. Ese tipo de decisión requiere conocimiento del objetivo formativo y del contexto disciplinar. La gobernanza funciona mejor cuando establece límites y principios comunes mientras permite decisiones pedagógicas contextualizadas.

Otro elemento necesario es la revisión periódica. Las políticas de IA no deberían redactarse como documentos permanentes que solo se modifican cuando ocurre una crisis. La velocidad de desarrollo tecnológico exige mecanismos programados para revisar definiciones, riesgos y procedimientos. Kanobel (2026) plantea precisamente la necesidad de una gobernanza adaptativa, capaz de responder a cambios técnicos sin abandonar principios relacionados con transparencia, protección de datos y participación de la comunidad educativa.

Esta adaptabilidad no significa modificar las reglas continuamente. La inestabilidad normativa también genera problemas. Lo que debería permanecer relativamente constante son los principios; lo que puede cambiar son las formas concretas mediante las cuales se aplican.

Alfabetización en inteligencia artificial de docentes y estudiantes

Regular una tecnología que sus usuarios no comprenden tiene límites evidentes. Las políticas pueden indicar qué está permitido, pero no garantizan que estudiantes y docentes posean los conocimientos necesarios para evaluar una respuesta generada, reconocer un sesgo, identificar información falsa o comprender las implicaciones de introducir datos personales en una plataforma.

La alfabetización en inteligencia artificial constituye por ello una condición de la gobernanza y no un complemento opcional. Ng et al. (2021), a partir de una revisión exploratoria de la literatura, organizaron el concepto de alfabetización en IA alrededor de cuatro componentes: conocer y comprender la inteligencia artificial, saber utilizarla y aplicarla, evaluar críticamente sus resultados y comprender sus implicaciones éticas. Esta formulación resulta relevante porque evita reducir la competencia al manejo instrumental de herramientas.

Saber formular una instrucción eficaz puede aumentar la calidad de una respuesta, pero no permite por sí mismo decidir si esa respuesta es verdadera, éticamente aceptable o apropiada para una determinada actividad académica. Una formación centrada exclusivamente en ingeniería de prompts corre el riesgo de preparar usuarios eficientes sin desarrollar usuarios críticos.

UNESCO ha ampliado esta perspectiva mediante marcos diferenciados para estudiantes y docentes. El marco dirigido a estudiantes establece doce competencias distribuidas en cuatro dimensiones: orientación centrada en el ser humano, ética de la IA, técnicas y aplicaciones, y diseño de sistemas. Estas dimensiones se desarrollan progresivamente desde comprensión hacia aplicación y creación (Miao et al., 2024).

El marco para docentes adopta una estructura relacionada pero ajustada a las responsabilidades profesionales. Miao y Cukurova (2024) identifican quince competencias agrupadas en cinco dimensiones: mentalidad centrada en el ser humano, ética de la IA, fundamentos y aplicaciones, pedagogía de la IA y utilización de IA para el desarrollo profesional. Las

competencias se organizan en tres niveles progresivos: adquirir, profundizar y crear.

La diferencia entre ambos marcos es significativa. No basta con que el profesor sepa utilizar las mismas herramientas que sus estudiantes. Debe ser capaz de decidir cuándo una aplicación tiene valor pedagógico, reconocer riesgos, diseñar actividades adecuadas, explicar sus decisiones y supervisar el uso por parte de aprendices con diferentes niveles de autonomía.

La necesidad de formación docente está respaldada por evidencia empírica. Mah y Groß (2024) estudiaron a 122 docentes de instituciones de educación superior alemanas. Entre los principales desafíos identificados se encontraba la falta de alfabetización en IA de estudiantes y profesores. El 78,5 % manifestó interés en actividades de formación relacionadas con enseñanza y aprendizaje mediante inteligencia artificial. Aunque estos resultados corresponden a un contexto específico y no pueden extrapolarse directamente a América Latina, muestran que la incorporación efectiva de IA no depende solo de disponer de acceso tecnológico.

La formación debería abordar conocimientos conceptuales, usos pedagógicos, evaluación crítica, privacidad, sesgos, derechos de autor, integridad académica, evaluación educativa y responsabilidades institucionales. La profundidad necesaria variará según el rol. Un docente no requiere convertirse en especialista en aprendizaje automático para reconocer que un modelo puede producir información incorrecta o que introducir datos identificables de estudiantes en una plataforma externa puede generar problemas de privacidad.

La alfabetización de los estudiantes debe superar una lógica de advertencias. Informarles que “la IA puede equivocarse” tiene un efecto limitado si no aprenden procedimientos concretos para comprobar una respuesta. La verificación necesita convertirse en práctica: localizar la fuente original, contrastar afirmaciones, distinguir fuentes primarias de comentarios secundarios, revisar cálculos y reconocer cuándo una pregunta excede la confiabilidad de la herramienta.

Kasneci et al. (2023) plantean que la incorporación responsable de modelos de lenguaje requiere competencias que permitan comprender sus limitaciones, supervisar los resultados y realizar verificación sistemática

de la información. La necesidad de estas capacidades aumenta cuando las respuestas adoptan un lenguaje convincente que puede ocultar errores.

En Ecuador, los resultados de Buele et al. (2025) aportan un argumento adicional. El estudio realizado con 833 estudiantes universitarios encontró que las dimensiones cognitivas y afectivas se relacionaban con el comportamiento ético frente a la inteligencia artificial, mientras la frecuencia de utilización por sí sola no mostró una asociación significativa en el modelo analizado. Esto sugiere que utilizar más o menos IA no explica por sí mismo la calidad ética de la conducta. El conocimiento y la reflexión parecen tener un papel más relevante.

Este resultado cuestiona políticas que buscan resolver el problema simplemente limitando la frecuencia de uso. Un estudiante puede utilizar poco una herramienta y hacerlo de manera inadecuada, o utilizarla con frecuencia manteniendo procedimientos de verificación y transparencia. La formación debe dirigirse hacia criterios de decisión.

En educación escolar, el desarrollo de estas competencias requiere una progresión más cuidadosa. Los estudiantes más jóvenes no poseen necesariamente la misma capacidad para valorar riesgos, verificar información o comprender políticas de privacidad que un adulto universitario. La mediación docente, la selección institucional de herramientas y la adaptación a la edad son condiciones importantes.

La alfabetización tampoco debe presentarse como una responsabilidad exclusiva de los estudiantes. Los docentes necesitan poder explicar por qué una determinada práctica se permite o se restringe. Cuando las normas se comunican sin fundamento pedagógico, pueden ser percibidas simplemente como mecanismos de control.

Una política madura debería conseguir que el estudiante comprenda no solo qué regla debe cumplir, sino qué aprendizaje intenta proteger esa regla. Esta comprensión transforma la integridad académica desde obediencia externa hacia responsabilidad intelectual.

Transparencia, declaración y criterios para el uso académico de IA

La transparencia es uno de los principios más utilizados en los marcos éticos de inteligencia artificial, pero su significado puede volverse impreciso si no se traduce en prácticas concretas. En educación, transparentar el uso de IA implica identificar aquellas intervenciones tecnológicas que modifican de manera relevante las condiciones de producción, evaluación o interpretación de un trabajo.

No todas las funciones automatizadas requieren necesariamente la misma declaración. Un corrector ortográfico integrado en un procesador de texto no tiene el mismo impacto sobre la autoría que una herramienta que genera la estructura argumentativa de un ensayo. Establecer una obligación de declarar cada función automatizada utilizada durante el proceso puede producir registros extensos con poco valor pedagógico.

La declaración debería centrarse en contribuciones materiales para comprender cómo se elaboró el producto. Dependiendo de la actividad, puede resultar pertinente indicar la herramienta utilizada, su finalidad, las etapas en las que intervino, el tipo de contenido producido y las estrategias empleadas para comprobar o modificar los resultados.

La transparencia cumple varias funciones. Permite atribuir con mayor precisión la contribución humana, ayuda al docente a interpretar el trabajo dentro de las condiciones reales de producción y obliga al estudiante a reflexionar sobre las decisiones tomadas. También puede reducir la incertidumbre sobre prácticas permitidas cuando existe un formato común de declaración.

Sin embargo, declarar no equivale a legitimar. Si una evaluación prohíbe expresamente determinada forma de asistencia, informar posteriormente que se utilizó no elimina la infracción. Autorización y transparencia son dimensiones relacionadas pero distintas.

La regla inversa también merece atención. Una actividad puede permitir inteligencia artificial y, aun así, requerir declaración si conocer la forma en que fue utilizada tiene relevancia para interpretar el resultado. Por ejemplo, en un proyecto donde se evalúa la capacidad para trabajar críticamente

con IA, la descripción del proceso forma parte de la propia evidencia de aprendizaje.

Luo (2024) advierte que las políticas universitarias desarrolladas ante la inteligencia artificial han tendido a enfatizar la originalidad del trabajo. El problema es que la originalidad se vuelve más difícil de definir en procesos híbridos donde las herramientas pueden participar en generación, traducción, revisión y reformulación. En lugar de exigir una ficción de producción completamente aislada, las instituciones pueden establecer criterios que permitan distinguir contribuciones y responsabilidades.

La transparencia debe operar también desde el docente y la institución. Si un sistema de inteligencia artificial participa en la producción de retroalimentación, evaluación preliminar, selección de recursos o análisis de datos de estudiantes, las personas afectadas deberían conocer, cuando resulte relevante, que existe esa intervención tecnológica.

Este principio se vuelve más importante cuando la herramienta influye en decisiones de consecuencias significativas. La Recomendación sobre la Ética de la Inteligencia Artificial de UNESCO insiste en la necesidad de supervisión humana, transparencia, explicabilidad y posibilidad de atribuir responsabilidad a personas o entidades identificables (UNESCO, 2021). Delegar una decisión educativa en un sistema automatizado no elimina la responsabilidad institucional sobre sus efectos.

Un docente que utiliza IA para elaborar comentarios puede ahorrar tiempo, pero continúa siendo responsable de que la retroalimentación sea correcta, adecuada y proporcional. Una institución que emplea sistemas automatizados para identificar posibles infracciones académicas debe disponer de procedimientos humanos de revisión y permitir que el estudiante comprenda y cuestione la evidencia utilizada.

La transparencia también requiere evitar declaraciones institucionales que exageren las capacidades de los sistemas. Presentar una herramienta como objetiva, personalizada o inteligente no constituye evidencia de que sus decisiones sean imparciales o pedagógicamente adecuadas. Los criterios utilizados para adoptar tecnologías deberían ser accesibles y susceptibles de revisión.

En contextos de investigación, la declaración adquiere otra dimensión. Las políticas editoriales y científicas han avanzado hacia la exigencia de reconocer usos sustantivos de inteligencia artificial y mantener la responsabilidad humana sobre el contenido. Aunque estas normas no son idénticas a las de una actividad escolar, comparten un principio: la intervención de una herramienta no elimina la obligación de responder por el resultado.

La transparencia, por tanto, no debe convertirse en un formulario burocrático. Su valor reside en hacer visibles aquellas intervenciones necesarias para interpretar adecuadamente quién realizó qué, bajo qué condiciones y con qué responsabilidad.

Equidad, privacidad, sesgos y responsabilidad ética

La discusión educativa sobre inteligencia artificial suele concentrarse en integridad académica, pero una política institucional limitada al fraude deja sin resolver riesgos importantes. Una herramienta puede utilizarse dentro de las reglas de una actividad y, sin embargo, producir problemas relacionados con privacidad, discriminación, desigualdad de acceso o tratamiento inadecuado de información.

UNESCO (2021) establece que la gobernanza ética de la inteligencia artificial debe proteger la privacidad, prevenir discriminación, conservar supervisión humana y garantizar mecanismos de responsabilidad. Estas exigencias son especialmente pertinentes en educación porque las instituciones gestionan información personal, académica y, en ocasiones, sensible de sus estudiantes.

La privacidad constituye uno de los primeros aspectos que deben examinarse antes de adoptar una plataforma. Introducir nombres, expedientes académicos, evaluaciones, información clínica, características personales o documentos identificables en servicios externos puede implicar transferencia de datos hacia sistemas cuyas condiciones de almacenamiento y reutilización no son controladas por la institución.

Miao y Holmes (2023) recomiendan que las instituciones establezcan criterios específicos para proteger información personal y desarrollen capacidades para evaluar las condiciones de privacidad de las

herramientas generativas. La responsabilidad no debería trasladarse al usuario mediante la simple aceptación de términos de servicio que pocos estudiantes o docentes leen en detalle.

Este problema aumenta en educación escolar. Los menores requieren protecciones adicionales y las instituciones deben considerar requisitos de edad, consentimiento y tratamiento de datos antes de exigir el uso de una aplicación. No sería adecuado convertir una herramienta comercial en requisito curricular sin comprobar previamente si sus condiciones son compatibles con las obligaciones de protección aplicables.

Los sesgos representan un segundo riesgo. Los modelos aprenden a partir de grandes conjuntos de datos que contienen desigualdades, estereotipos y representaciones culturales desequilibradas. Sus respuestas pueden reproducir esas características incluso cuando el lenguaje utilizado parece neutral. Kasneci et al. (2023) incluyen el sesgo entre los principales desafíos de los grandes modelos de lenguaje en educación y plantean la necesidad de supervisión humana y formación crítica para interpretarlos.

El problema no desaparece mediante una advertencia genérica de que “la IA puede tener sesgos”. Los estudiantes necesitan aprender a reconocer cuándo determinadas perspectivas están sobrerrepresentadas, qué voces aparecen ausentes y qué consecuencias podría tener utilizar una respuesta sin examinar su contexto.

En América Latina adquiere especial relevancia la representación cultural y lingüística. Los sistemas entrenados predominantemente con información producida en ciertos idiomas y contextos pueden presentar respuestas menos precisas o menos matizadas sobre realidades locales. Una herramienta capaz de generar un texto fluido en español no garantiza que comprenda con igual profundidad las particularidades sociales, jurídicas o educativas de Ecuador.

La equidad de acceso constituye una tercera dimensión. Muchas plataformas ofrecen versiones gratuitas y comerciales con diferencias de capacidad, velocidad, contexto disponible o acceso a modelos más avanzados. Cuando una actividad evaluativa permite o exige estas herramientas, esas diferencias pueden convertirse en ventajas académicas asociadas con capacidad económica.

La desigualdad tampoco se limita a las suscripciones. Depende de conectividad, disponibilidad de dispositivos, alfabetización digital, tiempo para experimentar y apoyo institucional. El problema es particularmente importante en sistemas educativos donde persisten brechas urbano-rurales.

El diagnóstico de preparación para inteligencia artificial realizado por UNESCO en Ecuador identificó diferencias importantes de conectividad entre hogares urbanos y rurales, junto con limitaciones de infraestructura y formación especializada en ética y gobernanza de IA (UNESCO, 2025). Este contexto obliga a interpretar con cautela cualquier propuesta de incorporación obligatoria de herramientas generativas en instituciones ecuatorianas.

La equidad no implica necesariamente proporcionar la misma herramienta a todos. Requiere evitar que diferencias tecnológicas no relacionadas con el resultado de aprendizaje determinen el desempeño evaluado. Una institución puede ofrecer licencias comunes, alternativas equivalentes o actividades que no dependan de funciones disponibles solo mediante pago.

La accesibilidad presenta una dimensión distinta. La inteligencia artificial puede reducir algunas barreras mediante transcripción, traducción, generación de texto alternativo o adaptación de materiales. Mah y Groß (2024) encontraron que los docentes participantes percibían la posibilidad de mayor equidad como uno de los beneficios de la IA educativa. Sin embargo, este potencial no debe interpretarse como garantía de inclusión. Las herramientas pueden facilitar determinadas actividades y simultáneamente generar nuevas barreras si las interfaces, los datos o los resultados no contemplan diversidad de usuarios.

La responsabilidad ética exige identificar estos riesgos antes de que se materialicen. La siguiente matriz resume áreas prioritarias de vigilancia institucional.

Tabla 5. *Riesgos institucionales asociados con el uso educativo de inteligencia artificial y posibles salvaguardas*

Riesgo	Manifestación educativa	Salvaguarda institucional	Responsable principal
<i>Privacidad y protección de datos</i>	Introducción de información identificable o sensible en plataformas externas	Evaluación previa de proveedores, minimización de datos, reglas sobre información prohibida y formación de usuarios	Autoridades institucionales, protección de datos y unidades técnicas
<i>Sesgo y discriminación</i>	Respuestas que reproducen estereotipos o perjudican a determinados grupos	Supervisión humana, evaluación contextual, alfabetización crítica y mecanismos para reportar resultados problemáticos	Docentes, institución y proveedores
<i>Desigualdad de acceso</i>	Diferencias entre versiones gratuitas y pagadas, conectividad o disponibilidad de dispositivos	Acceso institucional, alternativas equivalentes y análisis previo de las condiciones tecnológicas de la población	Institución y autoridades educativas
<i>Información incorrecta</i>	Incorporación de datos, referencias o	Verificación obligatoria, formación en fuentes y	Docentes y estudiantes

	explicaciones falsas	responsabilidad humana sobre los productos	
<i>Automatización excesiva</i>	Sustitución de decisiones pedagógicas o evaluativas por resultados algorítmicos	Supervisión humana significativa y límites explícitos para decisiones de alta consecuencia	Docentes y autoridades académicas
<i>Falta de transparencia</i>	Uso institucional o estudiantil de IA sin información suficiente sobre su intervención	Políticas de declaración, documentación y comunicación de usos relevantes	Toda la comunidad educativa
<i>Dependencia de proveedores</i>	Cambios de condiciones, precios, disponibilidad o tratamiento de datos fuera del control institucional	Evaluación periódica, interoperabilidad y alternativas tecnológicas	Autoridades y unidades técnicas
<i>Uso por menores</i>	Exposición a contenido, tratamiento de datos o servicios incompatibles con su edad	Selección institucional de herramientas, mediación adulta y cumplimiento de requisitos de edad y consentimiento	Sistema educativo, institución y familias según corresponda

Nota. Elaboración propia a partir de Chan (2023), Kasneci et al. (2023), Miao y Holmes (2023), UNESCO (2021, 2025) y Kanobel (2026).

La utilidad de una matriz de este tipo no consiste en eliminar todo riesgo antes de utilizar una herramienta, objetivo difícilmente alcanzable. Permite identificar quién debe responder por cada dimensión y evitar que la decisión de utilizar IA se base únicamente en facilidad de acceso o popularidad.

La responsabilidad ética requiere además proporcionalidad. No todas las aplicaciones producen el mismo nivel de riesgo. Utilizar una herramienta para generar preguntas de práctica con información no identificable plantea problemas diferentes a enviarle expedientes académicos completos para producir recomendaciones sobre estudiantes.

La evaluación del riesgo debe corresponder a la naturaleza de la información, el impacto potencial de la decisión y la población afectada. Cuanto mayores sean las consecuencias, mayor debe ser la supervisión humana.

Perspectivas y desafíos para América Latina y

Ecuador

América Latina se encuentra en una situación caracterizada por una rápida adopción de inteligencia artificial y un desarrollo institucional más lento de políticas, estructuras de supervisión y mecanismos de evaluación. Esta diferencia entre uso real y capacidad de gobernanza constituye uno de los principales desafíos de la región.

Valentini (2026) aporta una de las aproximaciones regionales más amplias disponibles hasta el momento. La autora analizó información de 200 instituciones de educación superior pertenecientes a 19 países de América Latina y el Caribe mediante una encuesta realizada entre agosto y octubre de 2025. Los resultados muestran que el 73,5 % reportaba utilización de IA en enseñanza y aprendizaje y el 57 % en investigación. La adopción era menor en administración, con 34,1 %, y en servicio a la comunidad, con 20 %.

La infraestructura de gobernanza aparecía considerablemente rezagada frente al uso. Solo el 26 % de las instituciones contaba con una estrategia formal de inteligencia artificial, el 18,5 % disponía de políticas institucionales amplias, el 8 % había asignado un presupuesto específico y apenas el 9 % reportaba mecanismos formales de evaluación (Valentini, 2026).

Estos datos muestran que el principal problema regional no consiste en la ausencia de inteligencia artificial dentro de las instituciones. La tecnología ya está siendo utilizada. La brecha se encuentra entre esa adopción y la capacidad institucional para orientarla, supervisarla y evaluar sus consecuencias.

Una política creada después de que las prácticas estén ampliamente extendidas enfrenta una dificultad adicional: no comienza sobre un espacio vacío. Los estudiantes y docentes ya han construido formas de uso, expectativas y hábitos. La gobernanza debe reconocer esas prácticas reales en lugar de asumir que una regulación puede simplemente sustituirlas.

Kanobel (2026) analiza este problema desde el contexto de las universidades latinoamericanas y propone una gobernanza responsable organizada alrededor de transparencia algorítmica, autoría aumentada, protección de datos, sostenibilidad digital y participación democrática. La propuesta resulta pertinente porque incorpora dimensiones frecuentemente ausentes de las primeras políticas, que tendían a concentrarse en plagio y prohibición.

La participación democrática merece particular atención. La inteligencia artificial afecta docencia, evaluación, investigación, administración y relaciones laborales. Diseñar políticas únicamente desde autoridades centrales puede producir normas técnicamente correctas pero con baja legitimidad práctica. Estudiantes, docentes, investigadores, personal técnico y unidades jurídicas necesitan participar en la identificación de riesgos y necesidades.

El contexto latinoamericano presenta además una tensión entre dependencia tecnológica y autonomía institucional. La mayor parte de las herramientas generativas utilizadas en educación pertenece a proveedores internacionales cuyas decisiones sobre precios, disponibilidad,

procesamiento de datos y características de los modelos no son controladas por los sistemas educativos locales.

Esta dependencia no implica que las instituciones deban desarrollar sus propios modelos, algo poco realista para la mayoría. Sí exige evitar que una plataforma específica se vuelva indispensable para cumplir los objetivos curriculares. La capacidad de cambiar de proveedor, utilizar alternativas y conservar materiales educativos fuera de ecosistemas propietarios forma parte de una estrategia de sostenibilidad tecnológica.

Ecuador refleja varios de estos desafíos. El informe de preparación para inteligencia artificial presentado por UNESCO en 2025 constituye el primer diagnóstico integral del país respecto de la implementación de la Recomendación sobre la Ética de la Inteligencia Artificial. El proceso incluyó diez talleres nacionales de consulta con 91 participantes y reuniones con instituciones como el Ministerio de Telecomunicaciones y de la Sociedad de la Información, la Secretaría de Educación Superior, Ciencia, Tecnología e Innovación y la Superintendencia de Protección de Datos Personales (UNESCO, 2025).

El diagnóstico encontró avances, entre ellos la creación de un Comité de Inteligencia Artificial en 2024 y un crecimiento del debate nacional sobre regulación. Al mismo tiempo, señaló que Ecuador todavía no contaba con una estrategia nacional unificada de IA, lo que había generado iniciativas fragmentadas y coordinación limitada. El informe identificó también brechas de infraestructura y conectividad, escasez de formación en ética tecnológica y ausencia de mecanismos suficientes para evaluar impactos ambientales y diversidad en los equipos relacionados con IA (UNESCO, 2025).

Las recomendaciones formuladas incluyen desarrollar un marco jurídico integral, fortalecer mecanismos de gobernanza con múltiples actores, mejorar infraestructura digital, incorporar una perspectiva de género y promover transparencia algorítmica. Aunque estas recomendaciones se dirigen al ecosistema nacional de inteligencia artificial y no exclusivamente al sector educativo, tienen consecuencias directas para escuelas y universidades.

Las instituciones educativas ecuatorianas no deberían esperar necesariamente a que todos los marcos nacionales estén finalizados para

establecer criterios internos. Pueden desarrollar políticas compatibles con la legislación existente sobre protección de datos, con principios internacionales de derechos humanos y con las particularidades de sus programas académicos. Estas políticas tendrán que ajustarse conforme avance la regulación nacional.

La evidencia obtenida con estudiantes universitarios ecuatorianos refuerza la necesidad de una respuesta formativa. Buele et al. (2025) encontraron que el conocimiento y las actitudes hacia la IA se relacionaban con el comportamiento ético de los estudiantes de manera más clara que la frecuencia de utilización. Este resultado sugiere que una estrategia institucional basada únicamente en restricciones puede desaprovechar el componente educativo de la gobernanza.

Ecuador necesita, por tanto, evitar dos velocidades distintas: una adopción rápida impulsada por usuarios y una respuesta institucional que aparezca únicamente cuando se produce una infracción. La gobernanza debería acompañar el proceso desde antes, proporcionando formación, criterios y protección.

En educación escolar, el desafío es mayor por la menor disponibilidad de evidencia regional específica. Muchas investigaciones latinoamericanas sobre inteligencia artificial generativa se concentran todavía en educación superior. Las políticas escolares no deberían copiar directamente los modelos universitarios. Necesitan considerar edad, desarrollo cognitivo, protección de menores, mediación docente, participación de familias y condiciones de infraestructura.

Esta diferencia también afecta la formación. Un estudiante universitario puede asumir mayor responsabilidad sobre la selección y verificación de herramientas. En educación básica, la institución debe ejercer un papel mucho más activo en determinar qué aplicaciones pueden utilizarse y con qué condiciones.

La región necesita investigación que no se limite a medir percepciones sobre inteligencia artificial. Resulta necesario evaluar efectos sobre aprendizaje, diferencias entre grupos de estudiantes, consecuencias de distintas políticas, impacto de programas de alfabetización, funcionamiento de herramientas en español y capacidad de las instituciones para proteger datos y garantizar equidad.

La gobernanza debería apoyarse progresivamente en esa evidencia. Una política no puede considerarse exitosa únicamente porque exista un documento institucional. Necesita evaluarse mediante preguntas concretas: ¿los estudiantes comprenden las reglas?, ¿los docentes las aplican de manera consistente?, ¿se reducen las ambigüedades?, ¿se protege la información personal?, ¿las decisiones son equitativas?, ¿la integración tecnológica mejora o deteriora el aprendizaje?

La ausencia de evaluación constituye precisamente una de las debilidades identificadas por Valentini (2026) en América Latina y el Caribe. Que solo una pequeña proporción de las instituciones encuestadas reportara mecanismos formales de evaluación indica que muchas políticas pueden estar siendo implementadas sin conocer sus efectos reales.

Una gobernanza educativa sostenible necesita cerrar ese ciclo. Las políticas deben formularse, implementarse, evaluarse y revisarse. La experiencia de docentes y estudiantes debe formar parte del proceso de revisión junto con evidencia empírica y cambios regulatorios.

Esta lógica permite superar la oposición entre innovación y regulación. Regular no significa impedir el desarrollo tecnológico. La ausencia de reglas también genera costos: incertidumbre, decisiones inconsistentes, exposición de datos, acusaciones injustificadas de fraude y desigualdad entre quienes conocen mejor las herramientas y quienes carecen de acceso o formación.

La regulación responsable busca establecer condiciones en las que sea posible innovar sin trasladar los riesgos hacia estudiantes o docentes. El principio rector debería ser que la tecnología se incorpore cuando exista una justificación educativa y no simplemente porque esté disponible.

La inteligencia artificial generativa seguirá cambiando. Es probable que funciones que hoy parecen avanzadas se integren progresivamente en procesadores de texto, buscadores, plataformas de aprendizaje y sistemas institucionales hasta volverse menos visibles para el usuario. Una política basada únicamente en identificar herramientas externas tendrá cada vez menor capacidad para describir las prácticas reales.

Por ello, los principios de responsabilidad, transparencia, protección de datos, equidad, juicio crítico y supervisión humana poseen mayor

estabilidad que cualquier listado de aplicaciones autorizadas. También permiten mantener coherencia entre integridad académica y gobernanza institucional.

Los cuatro capítulos de esta obra convergen en esta idea. La integridad académica no puede reducirse a evitar plagio; la detección automática no puede sustituir procedimientos justos; la evaluación necesita producir evidencias válidas del aprendizaje; y las instituciones tienen la responsabilidad de crear las condiciones que permitan utilizar la inteligencia artificial sin debilitar esos principios.

La cuestión central no consiste en determinar cuánto espacio debe concederse a la inteligencia artificial dentro de la educación. Consiste en decidir qué funciones deben permanecer bajo responsabilidad humana, qué competencias necesitan desarrollar estudiantes y docentes y qué mecanismos institucionales permiten utilizar nuevas tecnologías sin comprometer la confianza sobre la que se sostiene el proceso educativo.

Una cultura responsable de inteligencia artificial será aquella en la que las reglas puedan explicarse pedagógicamente, las decisiones tecnológicas puedan cuestionarse, los riesgos se distribuyan de manera justa y la innovación permanezca subordinada a los propósitos de la educación. Bajo estas condiciones, la integridad académica deja de funcionar como una respuesta defensiva frente a la inteligencia artificial y se convierte en un principio para orientar la transformación educativa.

CONCLUSIONES

La incorporación de la inteligencia artificial generativa en los procesos educativos ha obligado a revisar conceptos que durante años parecían suficientemente delimitados, entre ellos la autoría, la originalidad, la responsabilidad académica y la propia evidencia del aprendizaje. Su presencia no elimina los principios de la integridad académica, pero sí modifica las condiciones en las que estos deben aplicarse, pues un producto elaborado con apoyo tecnológico puede combinar contribuciones humanas y automatizadas difíciles de separar mediante criterios tradicionales.

En este escenario, resulta insuficiente reducir el problema al plagio o asumir que cualquier utilización de inteligencia artificial constituye una conducta indebida. La cuestión central reside en determinar qué grado de intervención resulta compatible con el propósito formativo, qué decisiones deben permanecer bajo responsabilidad del estudiante y de qué manera debe transparentarse la participación tecnológica. La integridad académica adquiere así un sentido más amplio, asociado con la coherencia entre el trabajo presentado, el aprendizaje realmente alcanzado y la responsabilidad intelectual asumida por quien lo entrega.

Esta transformación también muestra los límites de una respuesta basada principalmente en vigilancia y detección. Las herramientas destinadas a identificar contenido generado mediante inteligencia artificial pueden aportar indicios útiles, pero su desempeño depende del tipo de texto, del modelo utilizado, de las modificaciones posteriores y de la presencia de contenidos híbridos, por lo que sus resultados no constituyen por sí solos una demostración de deshonestidad académica. Cuando una sospecha se convierte automáticamente en acusación, se debilitan tanto la equidad del procedimiento como la confianza que la integridad académica pretende proteger. Por esta razón, las instituciones necesitan diferenciar con claridad entre una señal que justifica una revisión y la evidencia necesaria para establecer que una norma fue vulnerada. La prevención, la comunicación previa de las condiciones de uso y los procedimientos de investigación transparentes ofrecen una respuesta más consistente que la búsqueda de mecanismos tecnológicos capaces de decidir automáticamente sobre la autoría o la conducta del estudiante.

El problema conduce inevitablemente hacia la evaluación educativa, porque la posibilidad de automatizar parte de una tarea obliga a preguntarse si el producto final continúa ofreciendo información suficiente sobre aquello que el estudiante sabe y puede hacer. La respuesta no consiste en abandonar los instrumentos tradicionales ni en convertir toda actividad en una evaluación mediada por inteligencia artificial, sino en revisar la relación entre los resultados de aprendizaje y las evidencias utilizadas para demostrarlos. Algunas situaciones requerirán desempeño individual sin asistencia, mientras otras podrán valorar procesos de razonamiento, toma de decisiones, revisión crítica o utilización responsable de herramientas generativas. De este modo, producto y proceso dejan de funcionar como alternativas excluyentes y pasan a proporcionar evidencias complementarias que permiten interpretar el aprendizaje con mayor precisión. La evaluación conserva así su función esencial cuando no se limita a comprobar la calidad superficial de una respuesta, sino que permite reconocer el conocimiento, el juicio y la capacidad de actuación que existen detrás de ella.

Estas decisiones no pueden depender únicamente de iniciativas individuales del profesorado, ya que la incorporación responsable de inteligencia artificial requiere políticas institucionales, formación y condiciones de acceso que otorguen coherencia al sistema educativo. Para América Latina y Ecuador, esta necesidad adquiere especial relevancia debido a las diferencias existentes en conectividad, alfabetización digital, formación docente, capacidad regulatoria y acceso a herramientas tecnológicas, factores que pueden ampliar desigualdades si no son considerados desde el diseño de las políticas. Una gobernanza adecuada deberá combinar protección de datos, transparencia, equidad, alfabetización crítica y supervisión humana, sin quedar ligada a plataformas específicas cuya evolución puede volver obsoletas las normas en poco tiempo. El desafío no consiste en mantener a la educación al margen de la inteligencia artificial, sino en incorporarla sin desplazar los procesos humanos que justifican el valor de aprender, evaluar y certificar conocimientos. Bajo esta perspectiva, la integridad académica, la evaluación y la gobernanza dejan de ser problemas separados y se convierten en componentes de una misma tarea: preservar la confianza en el aprendizaje en un entorno educativo profundamente transformado por nuevas formas de producción y mediación del conocimiento.

BIBLIOGRAFÍA

- Ahsan, K., Akbar, S., & Kam, B. (2022). Contract cheating in higher education: A systematic literature review and future research agenda. *Assessment & Evaluation in Higher Education*, 47(4), 523–539. <https://doi.org/10.1080/02602938.2021.1931660>
- Anders, A. D., & Dux Speltz, E. (2025). Developing generative AI literacies through self-regulated learning: A human-centered approach. *Computers and Education: Artificial Intelligence*, 9, Article 100482. <https://doi.org/10.1016/j.caeai.2025.100482>
- Bassett, M. A., Bradshaw, W., Bornsztejn, H., Hogg, A., Murdoch, K., Pearce, B., & Webber, C. (2026). Heads we win, tails you lose: AI detectors in education. *Journal of Higher Education Policy and Management*. Advance online publication. <https://doi.org/10.1080/1360080X.2026.2622146>
- Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 49(6), 893–905. <https://doi.org/10.1080/02602938.2024.2335321>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Blake Angulo, J. (2026). Docencia e inteligencia artificial generativa en educación superior en América Latina: Focos y vacíos de la investigación emergente (revisión sistemática 2023–2025). *Revista Realidad Educativa*, 6(1), 153–189. <https://doi.org/10.38123/rre.v6i1.600>
- Boud, D., & Soler, R. (2016). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, 41(3), 400–413. <https://doi.org/10.1080/02602938.2015.1018133>
- Bretag, T. (Ed.). (2016). *Handbook of academic integrity*. Springer. <https://doi.org/10.1007/978-981-287-098-8>

- Buele, J., Sabando-García, Á. R., Sabando-García, B. J., & Yáñez-Rueda, H. (2025). Ethical use of generative artificial intelligence among Ecuadorian university students. *Sustainability*, 17(10), Article 4435. <https://doi.org/10.3390/su17104435>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315–1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Carvajal, H., Tusa, F., Samaniego, R., & Célleri-Pacheco, J. (2026). Artificial intelligence and academic integrity in virtual higher education: A descriptive-comparative study of student and faculty perceptions in Ecuador. *Trends in Higher Education*, 5(3), Article 67. <https://doi.org/10.3390/higheredu5030067>
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20, Article 38. <https://doi.org/10.1186/s41239-023-00408-3>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>
- Eaton, S. E. (2023). Postplagiarism: Transdisciplinary ethics and integrity in the age of artificial intelligence and neurotechnology. *International Journal for Educational Integrity*, 19, Article 23. <https://doi.org/10.1007/s40979-023-00144-1>
- Eke, D. O. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? *Journal of Responsible Technology*, 13, Article 100060. <https://doi.org/10.1016/j.jrt.2023.100060>
- Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *International Journal for Educational Integrity*, 19, Article 17. <https://doi.org/10.1007/s40979-023-00140-5>

- Fleckenstein, J., Meyer, J., Jansen, T., Keller, S. D., Köller, O., & Möller, J. (2024). Do teachers spot AI? Evaluating the detectability of AI-generated texts among student essays. *Computers and Education: Artificial Intelligence*, 6, Article 100209. <https://doi.org/10.1016/j.caeai.2024.100209>
- Furze, L., Perkins, M., Roe, J., & MacVaugh, J. (2024). The AI Assessment Scale (AIAS) in action: A pilot implementation of GenAI-supported assessment. *Australasian Journal of Educational Technology*, 40(4), 38–55. <https://doi.org/10.14742/ajet.9434>
- Gruenhagen, J. H., Sinclair, P. M., Carroll, J.-A., Baker, P. R. A., Wilson, A., & Demant, D. (2024). The rapid rise of generative AI and its implications for academic integrity: Students' perceptions and use of chatbots for assistance with assessments. *Computers and Education: Artificial Intelligence*, 7, Article 100273. <https://doi.org/10.1016/j.caeai.2024.100273>
- Gustilo, L., Ong, E., & Lapinid, M. R. (2024). Algorithmically-driven writing and academic integrity: Exploring educators' practices, perceptions, and policies in AI era. *International Journal for Educational Integrity*, 20, Article 3. <https://doi.org/10.1007/s40979-024-00153-8>
- Hadra, M., Cambridge, K., & Mesbah, M. (2026). Evaluating the accuracy and reliability of AI content detectors in academic contexts. *International Journal for Educational Integrity*, 22, Article 4. <https://doi.org/10.1007/s40979-026-00213-1>
- International Center for Academic Integrity. (2021). *The fundamental values of academic integrity* (3rd ed.). https://academicintegrity.org/images/pdfs/20019_ICAI-Fundamental-Values_R12.pdf
- Kanobel, M. C. (2026). Hacia una gobernanza responsable de la inteligencia artificial generativa en la educación superior. *Revista Educación Superior y Sociedad*, 37(2), 342–371. <https://doi.org/10.54674/ess.v37i2.1204>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for

- education. *Learning and Individual Differences*, 103, Article 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kizilcec, R. F., Huber, E., Papanastasiou, E. C., Cram, A., Makridis, C. A., Smolansky, A., Zeivots, S., & Radulescu, C. (2024). Perceived impact of generative AI on assessments: Comparing educator and student perspectives in Australia, Cyprus, and the United States. *Computers and Education: Artificial Intelligence*, 7, Article 100269. <https://doi.org/10.1016/j.caeai.2024.100269>
- Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns*, 4(7), Article 100779. <https://doi.org/10.1016/j.patter.2023.100779>
- Lodge, J. M., Bearman, M., Dawson, P., Gniel, H., Harper, R., Liu, D., McLean, J., Ucnik, L., & Associates. (2025). *Enacting assessment reform in a time of artificial intelligence*. Tertiary Education Quality and Standards Agency. <https://www.teqsa.gov.au/sites/default/files/2025-09/enacting-assessment-reform-in-a-time-of-artificial-intelligence.pdf>
- Lodge, J. M., de Barba, P., Ainscough, L., Brazil, J. R., Broadbent, J., Ebbert, D., Frankland, S., Gabriel, F., Gašević, D., Hennicke, T., Lim, L.-A., Male, S. A., Mirriahi, N., Oliveira, E. A., Pacitti, H., Raković, M., Russell, J., Taylor-Griffiths, D., & Yang, S. (2026). *Assuring quality learning in a gen AI-integrated future: The role of adaptive capabilities*. Tertiary Education Quality and Standards Agency. <https://www.teqsa.gov.au/sites/default/files/2026-06/assuring-quality-learning-in-a-gen-AI-integrated-future.pdf>
- Lodge, J., Howard, S., Bearman, M. L., Dawson, P., & Associates. (2023). *Assessment reform for the age of artificial intelligence*. Tertiary Education Quality and Standards Agency. <https://www.teqsa.gov.au/sites/default/files/2023-09/assessment-reform-age-artificial-intelligence-discussion-paper.pdf>
- Luo, J. (2024). A critical review of GenAI policies in higher education assessment: A call to reconsider the “originality” of students’ work. *Assessment & Evaluation in Higher Education*, 49(5), 651–664. <https://doi.org/10.1080/02602938.2024.2309963>

- Mah, D.-K., & Groß, N. (2024). Artificial intelligence in higher education: Exploring faculty use, self-efficacy, distinct profiles, and professional development needs. *International Journal of Educational Technology in Higher Education*, 21, Article 58. <https://doi.org/10.1186/s41239-024-00490-1>
- Messick, S. (1993). Foundations of validity: Meaning and consequences in psychological assessment. *ETS Research Report Series*, 1993(2), i–18. <https://doi.org/10.1002/j.2333-8504.1993.tb01562.x>
- Miao, F., & Cukurova, M. (2024). *AI competency framework for teachers*. UNESCO. <https://doi.org/10.54675/ZJTE2084>
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. <https://doi.org/10.54675/EWZM9535>
- Miao, F., Shiohira, K., & Lao, N. (2024). *AI competency framework for students*. UNESCO. <https://doi.org/10.54675/JKJB9835>
- Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, 5, Article 100151. <https://doi.org/10.1016/j.caeo.2023.100151>
- Motta-Soto, B. P., Orbegoso-Leiva, F. C. L., Hernández-García, E. D., Carrillo-Atahualpa, R. L., Mostacedo-Villca, M., & Chafloque-Capuñay, F. (2026). Adopción de inteligencia artificial generativa en educación superior latinoamericana: Revisión de alcance. *Prohominum*, 8(1), 688–701. <https://doi.org/10.47606/ACVEN/PH0469>
- Munoz, A., Hinchcliff, M., Langfield, C., & Rogerson, A. (2026). How strong is the evidence in generative AI-related academic misconduct allegations? A mixed-methods analysis. *International Journal for Educational Integrity*, 22, Article 26. <https://doi.org/10.1007/s40979-026-00235-9>
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and*

Education: Artificial Intelligence, 2, Article 100041.
<https://doi.org/10.1016/j.caeai.2021.100041>

Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching & Learning Practice*, 20(2).
<https://doi.org/10.53761/1.20.02.07>

Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching & Learning Practice*, 21(6).
<https://doi.org/10.53761/q3azde36>

Roe, J., & Perkins, M. (2022). What are automated paraphrasing tools and how do we address them? A review of a growing threat to academic integrity. *International Journal for Educational Integrity*, 18, Article 15. <https://doi.org/10.1007/s40979-022-00109-w>

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, 76(3), 467–481.
<https://doi.org/10.1007/s10734-017-0220-3>

Tertiary Education Quality and Standards Agency. (2026). *Role-specific guide to promoting academic integrity, and managing and investigating academic misconduct*. Australian Government.
<https://www.teqsa.gov.au/sites/default/files/2026-05/role-specific-guide-to-promoting-academic-integrity.pdf>

Torres-Díaz, J. C. (2026). Generative AI adoption and self-reported academic integrity risk: A student taxonomy in Latin American higher education. *Education Sciences*, 16(9), Article 1397.
<https://doi.org/10.3390/educsci16091397>

Torres-Díaz, J. C., Duarte, J., Rivera, D., & Beltrán Flandoli, A. (2025). Artificial intelligence and academic integrity: Exploring plagiarism in Ecuadorian universities. *International Journal for Educational*

Integrity, 21, Article 35. <https://doi.org/10.1007/s40979-025-00209-3>

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>

UNESCO. (2025). *Ecuador: Artificial intelligence readiness assessment report*. <https://unesdoc.unesco.org/ark:/48223/pf0000396465>

Valentini, A. (2026). Adopción y gobernanza de la IA en la educación superior de América Latina y el Caribe: Resultados de una encuesta regional. *Revista Educación Superior y Sociedad*, 37(2), 372–397. <https://doi.org/10.54674/ess.v37i2.1278>

Van Vlasselaer, M., Van Droogenbroeck, F., & Spruyt, B. (2026). Who wrote this? Evaluating the reliability of AI detection tools in higher education. *International Journal for Educational Integrity*, 22, Article 16. <https://doi.org/10.1007/s40979-026-00226-w>

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, 43(5), 840–854. <https://doi.org/10.1080/02602938.2017.1412396>

Villarroel, V., & Bruna, D. (2019). ¿Evaluamos lo que realmente importa? El desafío de la evaluación auténtica en educación superior. *Calidad en la Educación*, 50, 492–509. <https://doi.org/10.31619/caledu.n50.729>

Walters, W. H., & Wilder, E. I. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. *Scientific Reports*, 13, Article 14045. <https://doi.org/10.1038/s41598-023-41032-5>

Waltzer, T., Cox, R. L., & Heyman, G. D. (2023). Testing the ability of teachers and students to differentiate between essays generated by ChatGPT and high school students. *Human Behavior and Emerging Technologies*, 2023, Article 1923981. <https://doi.org/10.1155/2023/1923981>

Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P., & Waddington, L. (2023).

Testing of detection tools for AI-generated text. *International Journal for Educational Integrity*, 19, Article 26. <https://doi.org/10.1007/s40979-023-00146-z>

Wright, C. (2026). Transcription is not generation: Distinguishing non-generative AI tool use from academic misconduct in higher education assessment. *International Journal for Educational Integrity*, 22, Article 24. <https://doi.org/10.1007/s40979-026-00234-w>

ISBN: 978-9942-53-221-3



Compás
capacitación e investigación