Evaluación universitaria después de la inteligencia artificial generativa

Joseph Valero, Carlos Castillo, David Báez,
Darwin Pozo & Omar Riofrio

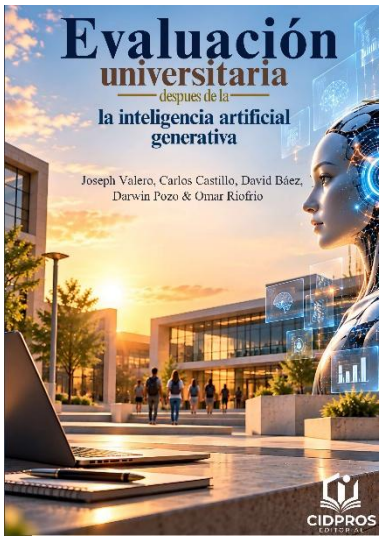
CIDPROS
EDITORIAL



Evaluación universitaria después de la inteligencia artificial generativa

Joseph Michael Valero Rodríguez, Carlos Fernando Castillo Montenegro, David Israel Báez Herrera, Darwin Paul Pozo Pilacuán & Omar Aníbal Riofrio Olaya

CIDPROS
EDITORIAL



Datos bibliográficos

ISBN	978-9907-848-17-5
Título del libro	Evaluación universitaria después de la inteligencia artificial generativa
Autores	Joseph Michael Valero Rodríguez Carlos Fernando Castillo Montenegro David Israel Báez Herrera Darwin Paul Pozo Pilacuán Omar Aníbal Riofrio Olaya
Editorial	CIDPROS EDITORIAL
Materia	370 - Educación
Público objetivo	Profesional / académico
Año	2026
Número de edición	1
Tamaño	2.91 MB
Soporte	Libro digital descargable
Formato	PDF (.pdf)
Idioma	Español
DOI	https://doi.org/10.70577/r6870402

Hecho en Ecuador / Made in Ecuador

MSc. Joseph Michael Valero Rodríguez

Instituto Superior Tecnológico de Formación Profesional, Administrativa y Comercial
joseph.valero@formacion.edu.ec
https://orcid.org/0000-0001-9120-3126
Salinas, Santa Elena, Ecuador

Semblanza



Joseph Michael Valero Rodríguez, Licenciado en Educación Básica por la Universidad Estatal Península de Santa Elena, y Magíster en Tecnología Educativa por el Instituto Tecnológico y de Estudios Superiores de Monterrey, México. Su sólida preparación académica se fortalece en la actualidad con sus estudios de Doctorado en Educación en la Universidad Nacional de Rosario, Argentina, en innovación pedagógica y la integración de la inteligencia artificial en los procesos formativos.

En su trayectoria profesional, ha ocupado el rol de Vicerrector Académico, sumado a una amplia experiencia en la enseñanza pública y privada. Actualmente, se desempeña como docente en el Instituto Superior Tecnológico de Formación Profesional Administrativa y Comercial y en el departamento de Tecnologías de la Información y Comunicación del Distrito Educativo 24D02. Su desempeño abarca la impartición de asignaturas en bachillerato —tales como Biología, Química, Física y Matemáticas—, así como la docencia en la carrera de Desarrollo Infantil y TICs a nivel técnico superior. Ha formado parte de proyectos de transformación digital, campañas de aprendizaje tecnológico e investigaciones en redes de IA aplicada a la educación, respaldo complementado con certificaciones en neurodidáctica, gamificación y liderazgo educativo.

Con una visión centrada en la evolución del pensamiento crítico y la ética en la Era Digital, su trabajo académico aborda los desafíos pedagógicos que plantea la inteligencia artificial generativa en la educación superior. A través del diseño de marcos para la evaluación auténtica y la preservación de la honestidad académica, busca orientar a docentes e instituciones hacia una convergencia responsable entre innovación tecnológica, evaluación del aprendizaje e impacto social.

MSc. Carlos Fernando Castillo Montenegro

Investigador independiente
carlosfernandocastillo@hotmail.com
https://orcid.org/0009-0008-2236-9059
Quito, Pichincha, Ecuador

Semblanza



Carlos Fernando Castillo Montenegro es un docente ecuatoriano, nacido el 1 de mayo de 1995 en la ciudad de Tulcán, provincia del Carchi, y residente en Quito, donde ha desarrollado gran parte de su trayectoria profesional. Es Licenciado en Ciencias de la Educación, mención Ciencias Naturales y del Ambiente, Biología y Química, por la Universidad Central del Ecuador, y Magíster en Educación de Bachillerato, con mención en Pedagogía de las Ciencias Naturales, por la Universidad Estatal de Milagro. Su formación académica y profesional está orientada a la enseñanza de las Ciencias Naturales, Biología y Química, con especial interés en la formación integral y el aprendizaje significativo de los estudiantes.

A lo largo de su trayectoria, ha consolidado una sólida experiencia docente en instituciones educativas públicas y particulares, principalmente en las áreas de Ciencias Naturales, Biología y Química. Su perfil se fortalece mediante una formación continua en metodologías activas, innovación educativa, tecnologías aplicadas a la enseñanza, evaluación digital y estrategias didácticas. En el ámbito académico, es coautor de publicaciones científicas relacionadas con las habilidades profesionales en la era de la transformación digital y la influencia de las metodologías activas en el aprendizaje significativo de estudiantes de Educación Superior, reflejando su interés por la innovación educativa y la investigación.

Se distingue por su compromiso con la educación, la actualización permanente y la búsqueda de estrategias innovadoras que contribuyan al fortalecimiento de los procesos de enseñanza-aprendizaje. Su trayectoria integra docencia, formación académica, capacitación e investigación, con una visión centrada en el desarrollo de competencias científicas y la formación integral de los estudiantes. Desde esta perspectiva, orienta su labor hacia el fortalecimiento de la calidad educativa y la generación de aportes que contribuyan a la innovación y transformación de la educación.

MSc. David Israel Báez Herrera

Investigador Independiente
dbmasternet@gmail.com
https://orcid.org/0009-0009-2487-7214
Quito, Pichincha, Ecuador

Semblanza



David Israel Báez Herrera es un investigador y profesional ecuatoriano residente en Quito, especializado en tecnologías de la información, transformación digital, gestión de proyectos e innovación educativa. Cuenta con un Bachelor of Arts (BA) en Redes y Sistemas Operativos, una maestría en Telecomunicaciones y una maestría en Innovación Educativa.

Actualmente cursa un Doctorado en Dirección de Proyectos, en el que desarrolla una investigación orientada al fortalecimiento de los procesos de evaluación interna y acreditación de las instituciones de educación superior mediante un modelo apoyado en una plataforma digital para la gestión, validación y análisis de evidencias. Posee más de diez certificaciones internacionales y acreditaciones de formación especializada en tecnologías, metodologías y estándares promovidos por organizaciones reconocidas como Cisco, Amazon Web Services (AWS), Red Hat, WatchGuard, Microsoft y Hewlett Packard (HP), además de formación en gestión de proyectos bajo el enfoque PMP y en sistemas de gestión basados en normas ISO.

Su preparación comprende áreas como redes, ciberseguridad, telecomunicaciones, computación en la nube, desarrollo de software, bases de datos, infraestructura tecnológica, sistemas operativos y dirección de proyectos. También cuenta con formación y certificaciones relacionadas con los sistemas de gestión de la calidad ISO 9001 y de seguridad de la información ISO/IEC 27001, así como una certificación como Delegado de Protección de Datos, conocimientos que le han permitido intervenir en iniciativas de calidad, seguridad digital, privacidad, protección de datos personales y cumplimiento normativo. Su perfil académico se complementa con más de 150 cursos de formación continua, cuatro diplomados vinculados con programación, investigación científica e inteligencia artificial y la participación en seminarios internacionales relacionados con sus áreas de especialización.

A lo largo de más de quince años de trayectoria, ha asesorado y acompañado a organizaciones públicas y privadas de los sectores educativo, financiero, bancario, inmobiliario, de la construcción y del desarrollo tecnológico en proyectos de infraestructura, redes, telecomunicaciones, seguridad informática y transformación digital. En el ámbito de la educación superior, ha liderado y participado en la

implementación de soluciones de software a medida para la gestión académica, administrativa y contable, contribuyendo a la automatización de actividades, la integración de información, la generación de indicadores y el fortalecimiento de la toma de decisiones institucionales. Su experiencia también comprende el diseño instruccional de contenidos y entornos virtuales de aprendizaje, la administración de plataformas educativas, el desarrollo de sistemas de información, la mejora y optimización de procesos y la implementación, mantenimiento y evaluación de sistemas de gestión de la calidad. Su ejercicio profesional se caracteriza por integrar la tecnología, la educación, la investigación y la gestión organizacional desde una perspectiva interdisciplinaria orientada a la innovación, la mejora continua, la seguridad de la información y el cumplimiento normativo.

Mantiene un firme compromiso con la ética profesional, la actualización permanente y la generación de soluciones tecnológicas que contribuyan a elevar la calidad educativa, fortalecer la eficiencia de las organizaciones y producir un impacto positivo y sostenible en la sociedad.

Mgtr. Darwin Paul Pozo Pilacuán

Investigador Independiente
paulpozowwe@gmail.com
https://orcid.org/0009-0001-9398-5501
Quito, Pichincha, Ecuador

Semblanza



Darwin Paul Pozo Pilacuán es profesional en Contabilidad y Auditoría, Contador Público Autorizado y Magíster en Gestión Financiera, Inversiones y Activos Digitales. Actualmente se desempeña como Financial Analyst y Docente de Educación superior; y cuenta con experiencia profesional en auditoría financiera, aseguramiento, control interno, análisis y gestión financiera, desarrollada en firmas de auditoría y consultoría nacionales e internacionales. Su trayectoria comprende también la participación en proyectos empresariales de diversos sectores y actividades relacionadas con el análisis de información financiera y el apoyo a la toma de decisiones.

En el ámbito académico, se desempeña como docente de educación superior y capacitador en áreas de contabilidad, auditoría, control interno, administración y análisis financieros. Su formación profesional se complementa con estudios y certificaciones en finanzas, inversiones, activos digitales, normas internacionales de información financiera e inteligencia artificial aplicada a las finanzas. Asimismo, ha fortalecido su preparación mediante programas de formación vinculados con IESE Business School, particularmente en finanzas para directivos, contabilidad financiera, situaciones de negocio y gestión del trabajo.

Como autor, mantiene un especial interés por la incorporación responsable y práctica de la inteligencia artificial en los procesos educativos. Su participación en la presente obra busca aportar conocimientos y reflexiones sobre el uso de herramientas de inteligencia artificial aplicadas a la docencia, promoviendo su aprovechamiento como recurso de apoyo para docentes y estudiantes. A través de esta publicación, procura contribuir al desarrollo de competencias digitales y a la innovación de los procesos de enseñanza y aprendizaje.

MSc. Omar Aníbal Riofrio Olaya

Investigador Independiente
omarriofrioolaya@hotmail.com
https://orcid.org/0009-0001-0457-6459
Milagro, Guayas, Ecuador

Semblanza



Omar Aníbal Riofrio Olaya es un destacado profesional ecuatoriano en el ámbito de la ingeniería civil, reconocido por su compromiso con la excelencia técnica, la innovación sostenible y la formación académica continua. Ingeniero Civil por la Universidad de Guayaquil, ha complementado su preparación con una sólida formación de posgrado orientada al desarrollo integral de la gestión en entornos industriales, empresariales y de infraestructura pública. Posee el Máster Universitario en Prevención de Riesgos Laborales y el Máster Universitario en Sistema Integrado de Gestión: Medio Ambiente, Prevención de Riesgos Laborales, Calidad y Responsabilidad Social Corporativa, ambos otorgados por la Universidad Internacional de La Rioja (España). Actualmente, cursa la Maestría en Ingeniería Civil con especialización en Sanitaria en la Universidad de Guayaquil, reafirmando su constante búsqueda de actualización profesional y su interés por las soluciones sostenibles en el ámbito del saneamiento urbano. Con más de una década de experiencia en el campo de la ingeniería, el Ing. Riofrio ha liderado proyectos de gran relevancia a nivel nacional, destacándose por su capacidad para integrar la planificación técnica con la responsabilidad ambiental y social. Desde el año 2021 se desempeña como Gerente General de la empresa ORO 4& JGC CONSTRUCTORA ORJACONSULTORA S.A.S, donde ha impulsado la implementación de prácticas de gestión eficientes, seguras y sostenibles. Su trayectoria también incluye su labor como consultor independiente en proyectos de infraestructura civil desde 2017, y como perito acreditado del Consejo de la Judicatura del Ecuador desde 2023, donde contribuye con su conocimiento técnico a procesos judiciales que demandan rigor, ética y objetividad profesional. En el ámbito educativo, Omar Riofrio ha ejercido la docencia en distintos niveles y contextos, demostrando su compromiso con la transmisión del conocimiento y la formación de nuevas generaciones. Ha sido docente particular, instructor de ajedrez en la Liga Cantonal de Milagro, docente en la Academia Naval sede Milagro, y docente ocasional en el Instituto Tecnológico Argos, donde ha orientado su enseñanza hacia el desarrollo del pensamiento lógico, la disciplina profesional y la aplicación práctica de la ingeniería. Como conferencista y formador, ha dictado charlas y cursos en instituciones como la Universidad Tecnológica de Babahoyo, el Colegio de Arquitectos del Guayas sede Milagro, la Liga Cantonal de Milagro y la Empresa Pública DASE de

Guayaquil, compartiendo su experiencia sobre temas vinculados a la gestión de riesgos, la ingeniería sostenible y la innovación constructiva. Su participación en estos espacios refleja una visión integradora de la ingeniería como campo de transformación social y desarrollo sostenible. Autor de diversos artículos científicos, su producción académica se orienta hacia la aplicación de sistemas integrados de gestión y la prevención de riesgos en entornos industriales y urbanos, aportando conocimiento técnico a la comunidad científica y profesional. El trabajo del Ing. Omar Riofrio se distingue por su liderazgo ético, su compromiso con la mejora continua y su visión humanista de la ingeniería. Su carrera encarna la convicción de que la construcción y el desarrollo no solo implican levantar estructuras físicas, sino también edificar un futuro basado en la responsabilidad, la innovación y el bienestar colectivo.

El contenido y las ideas expuestas en esta obra se encuentran protegidos por la normativa vigente en materia de propiedad intelectual y constituyen derechos exclusivos de su(s) autor(es).

Derechos de Autor © 2026

Este documento se publica bajo los términos y condiciones de la licencia:



Creative Commons Reconocimiento – No Comercial – Compartir Igual 4.0 Internacional (CC BY-NC-SA 4.0)

Usted es libre de compartir y adaptar el material, siempre que otorgue el crédito correspondiente, no lo utilice con fines comerciales y distribuya sus contribuciones bajo la misma licencia.

Para ver una copia de esta licencia, visite: https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es

Sinopsis

La expansión de la inteligencia artificial generativa está transformando profundamente la forma en que las universidades enseñan, aprenden y evalúan. Este libro analiza los desafíos que surgen cuando herramientas capaces de producir textos, resolver problemas, generar argumentos y apoyar múltiples tareas académicas intervienen en procesos tradicionalmente utilizados para demostrar aprendizaje. Frente a este escenario, se propone superar una visión centrada únicamente en la detección del uso de inteligencia artificial y avanzar hacia modelos evaluativos capaces de hacer visible la comprensión, el razonamiento, la toma de decisiones y la responsabilidad del estudiante.

La obra examina la transformación de la evaluación universitaria, el fortalecimiento de la evaluación auténtica, la incorporación de evidencias de proceso y trazabilidad, así como el valor de las modalidades orales, prácticas y situadas. También aborda los nuevos desafíos relacionados con integridad académica, autoría humana, transparencia y uso responsable de sistemas generativos.

A lo largo de sus capítulos se plantea que la inteligencia artificial no debe entenderse únicamente como una amenaza para la evaluación, sino como un factor que obliga a revisar prácticas tradicionales y fortalecer aquellas que permiten demostrar aprendizaje de manera más completa. El desafío universitario consiste en diseñar evaluaciones donde la tecnología pueda estar presente sin sustituir la participación intelectual, manteniendo visibles la autonomía, el juicio y la capacidad del estudiante para responder por aquello que produce.

Palabras clave: evaluación universitaria; inteligencia artificial generativa; evaluación auténtica; integridad académica; autoría académica.

Synopsis

The expansion of generative artificial intelligence is profoundly transforming the way universities teach, learn, and assess. This book examines the challenges that arise when tools capable of producing texts, solving problems, generating arguments, and supporting multiple academic tasks become involved in processes traditionally used to demonstrate learning. In response to this scenario, the book proposes moving beyond an approach focused solely on detecting the use of artificial intelligence and toward assessment models capable of making students' understanding, reasoning, decision-making, and responsibility visible.

The book examines the transformation of university assessment, the strengthening of authentic assessment, the incorporation of process-based evidence and learning traceability, as well as the value of oral, practical, and situated forms of assessment. It also addresses emerging challenges related to academic integrity, human authorship, transparency, and the responsible use of generative AI systems.

Throughout its chapters, the book argues that artificial intelligence should not be understood solely as a threat to assessment, but rather as a factor that compels universities to reconsider traditional practices and strengthen those that allow learning to be demonstrated more comprehensively. The challenge for higher education is to design assessments in which technology can be present without replacing intellectual engagement, while keeping students' autonomy, judgment, and ability to take responsibility for what they produce clearly visible.

Keywords: university assessment; generative artificial intelligence; authentic assessment; academic integrity; academic authorship.

Índice de Contenido

Sinopsis 12

Synopsis 12

Introducción 17

Capítulo 1. Transformación de la evaluación en la universidad 21

1.1 La evaluación universitaria ante la expansión de la inteligencia artificial generativa 22

1.2 Del control de contenidos a la evaluación de competencias y desempeños ... 25

1.3 Cambios en las prácticas tradicionales de evaluación académica 29

1.4 Nuevas formas de aprender, producir y demostrar conocimiento 32

1.5 El papel del docente en el rediseño de la evaluación universitaria 36

1.6 Alfabetización en inteligencia artificial para docentes y estudiantes 40

1.7 Principios para una evaluación universitaria pertinente en la era de la inteligencia artificial 44

Capítulo 2. Evaluación auténtica frente a contenidos generados por inteligencia artificial 48

2.1 Fundamentos y características de la evaluación auténtica 49

2.2 Limitaciones de las tareas tradicionales frente a la generación automatizada de contenidos 52

2.3 Diseño de actividades vinculadas con problemas y contextos reales 56

2.4 Evaluación de competencias mediante productos y desempeños auténticos . 60

2.5 Resolución de problemas, estudios de caso y aprendizaje basado en proyectos 64

2.6 Uso crítico de la inteligencia artificial dentro de las actividades evaluativas 67

2.7 Criterios y rúbricas para valorar producciones realizadas con apoyo de inteligencia artificial 70

Capítulo 3. Evidencias de proceso y trazabilidad del aprendizaje 74

3.1 Del producto final a la valoración del proceso de aprendizaje 75

3.2 Evidencias progresivas para demostrar la construcción del conocimiento 78

3.3 Portafolios y bitácoras como registros del proceso académico 81

3.4 Borradores, revisiones y versiones sucesivas de los productos evaluados 85

3.5 Registro y declaración del uso de inteligencia artificial durante el aprendizaje 88

3.6 Retroalimentación y metacognición como evidencias de comprensión 91

3.7 Sistemas de trazabilidad para verificar la evolución y autenticidad del trabajo académico 94

Capítulo 4. Evaluación oral, práctica y situada 98

4.1 La recuperación de la oralidad como evidencia de aprendizaje 99

4.2 Defensas orales y sustentaciones académicas frente al uso de inteligencia artificial 102

4.3 Evaluación mediante demostraciones prácticas y resolución de tareas 106

4.4 Simulaciones, escenarios profesionales y situaciones auténticas 108

4.5 Preguntas de profundización para comprobar comprensión y razonamiento 112

4.6 Evaluación situada y transferencia del conocimiento a contextos profesionales 115

4.7 Integración de evidencias escritas, orales y prácticas en la evaluación universitaria 118

Capítulo 5. Integridad académica y nuevos criterios de autoría 123

5.1 Transformación del concepto de integridad académica ante la inteligencia artificial generativa 124

5.2 Del plagio tradicional al uso no declarado de sistemas generativos 127

5.3 Autoría humana y responsabilidad sobre los productos académicos 131

5.4 Transparencia y declaración del uso de inteligencia artificial 133

5.5 Límites éticos entre asistencia tecnológica, colaboración y sustitución del trabajo académico 137

5.6 Políticas institucionales y criterios para el uso responsable de inteligencia artificial 141

5.7 Hacia una cultura de integridad, responsabilidad y autoría académica en la universidad 144

Conclusión.....148

Bibliografía.....151

Introducción

La irrupción de la inteligencia artificial generativa ha modificado de manera profunda las condiciones en las que se produce, comunica y evalúa el conocimiento en la universidad. Herramientas capaces de generar textos, sintetizar información, resolver problemas, formular argumentos y producir diferentes tipos de contenidos intervienen actualmente en actividades que durante años fueron utilizadas como evidencias relativamente directas del aprendizaje. En este escenario, la calidad de un producto final ya no permite asumir automáticamente que el estudiante comprendió, razonó o construyó de manera autónoma aquello que presenta. La evaluación universitaria necesita revisar sus fundamentos y criterios para responder a estas nuevas condiciones de producción académica (Villegas y Carballo, 2026).

Esta transformación no debe entenderse únicamente como una amenaza asociada con posibles usos inadecuados de la tecnología. También constituye una oportunidad para revisar prácticas evaluativas que ya presentaban limitaciones para representar la complejidad del aprendizaje universitario. Wang et al. (2024) señalan que la inteligencia artificial está modificando distintas dimensiones de la educación y las formas de interacción con el conocimiento. Por ello, la evaluación necesita avanzar desde modelos predominantemente centrados en reproducción de contenidos hacia propuestas capaces de valorar comprensión, aplicación, juicio, argumentación, transferencia y toma de decisiones frente a situaciones académicas y profesionales diversas.

Ensayos, informes, cuestionarios, proyectos y otras producciones mantienen valor dentro de la formación superior, pero necesitan interpretarse desde un contexto tecnológico diferente. La disponibilidad de herramientas capaces de producir respuestas estructuradas y formalmente correctas cuestiona especialmente aquellas evaluaciones donde la entrega final constituye la principal evidencia utilizada para inferir aprendizaje. Bearman et al. (2024) destacan la relevancia del juicio evaluativo en un escenario caracterizado por inteligencia artificial generativa. El estudiante necesita demostrar no solo capacidad para obtener respuestas, sino también para reconocer su calidad, detectar limitaciones, comparar alternativas, justificar decisiones y determinar la pertinencia de aquello que incorpora a su trabajo.

El cambio supone prestar mayor atención a lo que el estudiante puede hacer con los conocimientos adquiridos. Las competencias universitarias se hacen visibles cuando los saberes permiten interpretar problemas, establecer relaciones, seleccionar procedimientos y construir respuestas ajustadas a determinadas condiciones. Una evaluación orientada hacia el desempeño requiere situaciones donde las decisiones formen parte de la evidencia y donde la respuesta no pueda reducirse a una reproducción previsible de información. La inteligencia artificial refuerza esta necesidad porque las tareas fácilmente automatizables ofrecen cada vez menos información sobre la participación intelectual que sostuvo un resultado académico.

Dentro de esta perspectiva, la evaluación auténtica adquiere especial importancia. Su valor consiste en relacionar conocimientos con problemas contextualizados, propósitos concretos y actuaciones que demandan interpretación y toma de decisiones. Gedera (2026) propone una concepción de evaluación auténtica habilitada por inteligencia artificial en la que la tecnología puede integrarse sin eliminar la agencia humana. Esta posibilidad permite superar la oposición entre prohibición y aceptación indiscriminada. Una actividad puede admitir recursos generativos y continuar siendo exigente cuando obliga al estudiante a verificar, adaptar, seleccionar y justificar los resultados utilizados de acuerdo con criterios previamente establecidos.

La transformación evaluativa también requiere ampliar la atención desde el producto hacia el proceso de aprendizaje. Una entrega final puede presentar coherencia y calidad sin mostrar cómo fue construida, qué dificultades aparecieron o qué decisiones explican su configuración definitiva. Borradores, versiones sucesivas, portafolios, bitácoras, retroalimentaciones y registros de cambios ofrecen evidencias sobre la trayectoria seguida. Chhetri Thapa y Lewis (2026) destacan la importancia de preservar la autenticidad epistémica mediante evaluaciones orientadas al proceso. Esta perspectiva permite reconocer transformaciones en la comprensión y hacer visible la participación del estudiante más allá de la apariencia final de una producción.

La trazabilidad adquiere entonces un sentido pedagógico diferente de la vigilancia. Su finalidad no consiste en registrar exhaustivamente cada acción, sino en disponer de información suficiente para comprender cómo evolucionó una producción y qué decisiones explican sus modificaciones. En actividades mediadas por inteligencia artificial resulta útil identificar aportes tecnológicos relevantes, alternativas descartadas, contenidos reformulados y respuestas a la retroalimentación. Estas

evidencias permiten distinguir entre obtener un resultado y comprender aquello que se produce. Al mismo tiempo, favorecen la reflexión del estudiante sobre sus estrategias y fortalecen una evaluación interesada en la construcción progresiva del conocimiento.

Las modalidades orales, prácticas y situadas complementan esta perspectiva porque permiten observar el aprendizaje mediante interacción y desempeño directo. Explicar una decisión, responder preguntas, ejecutar un procedimiento, defender una propuesta o actuar frente a un escenario profesional proporciona información que no siempre queda representada en un documento. Stephenson et al. (2025) muestran que la evaluación oral puede aportar evidencias relevantes sobre comprensión y desempeño en educación superior. Estas modalidades no deben concebirse únicamente como mecanismos destinados a comprobar posibles usos de inteligencia artificial, sino como oportunidades para ampliar las formas mediante las cuales el estudiante demuestra razonamiento, apropiación y capacidad de respuesta.

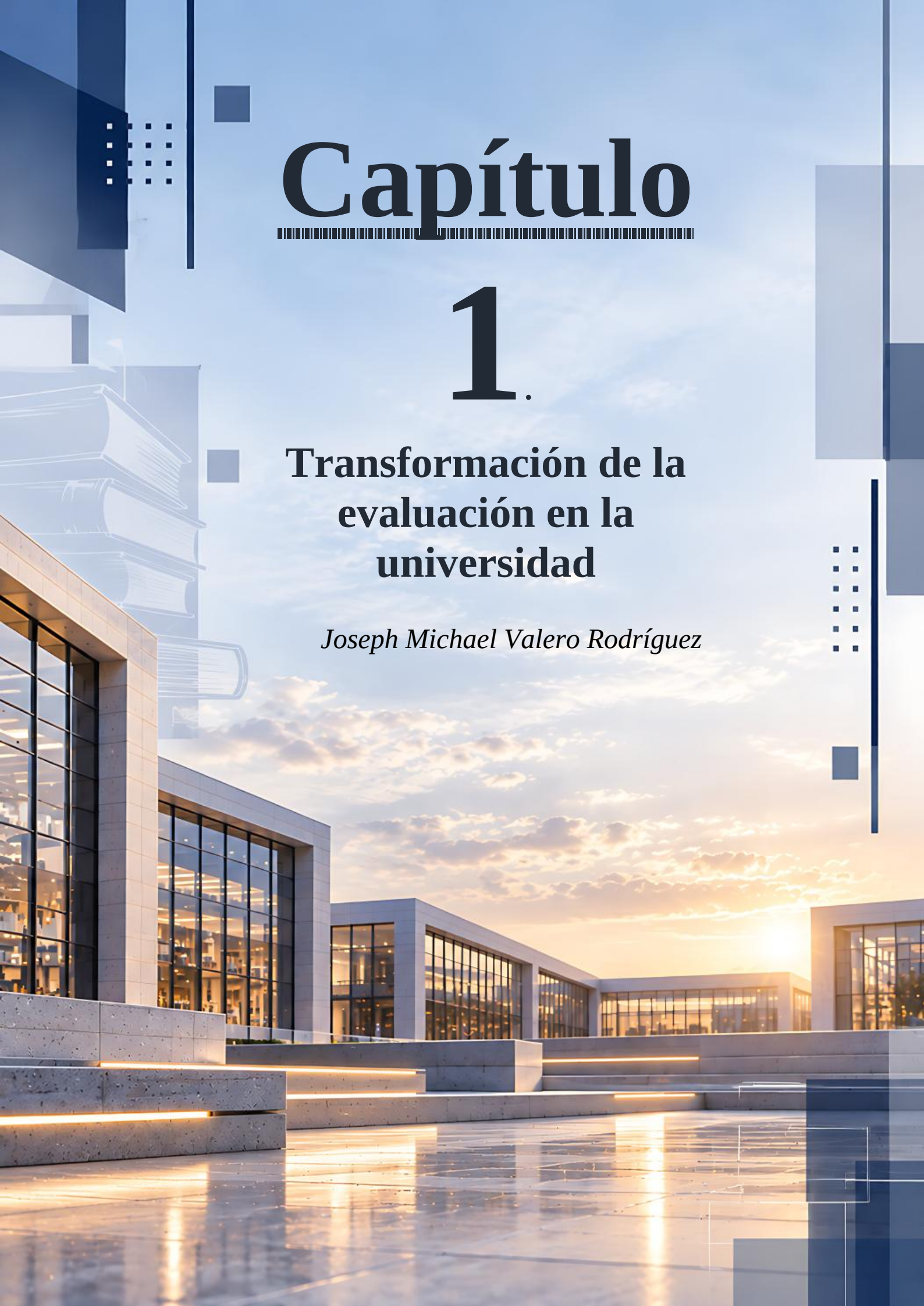
La integración de evidencias escritas, orales y prácticas permite construir juicios más completos sobre el aprendizaje. Cada modalidad aporta información diferente: la escritura puede mostrar organización y argumentación; la oralidad permite observar apropiación y capacidad explicativa; el desempeño práctico evidencia aplicación, transferencia y adaptación. Cuando estas evidencias convergen, fortalecen la interpretación del nivel alcanzado; cuando presentan discrepancias, permiten reconocer dimensiones que requieren atención específica. De esta manera, la diversidad evaluativa reduce la dependencia de un único producto y favorece una comprensión más amplia de competencias cuya complejidad difícilmente puede representarse mediante una sola actividad.

La expansión de los sistemas generativos también modifica la discusión sobre integridad y autoría académica. El problema ya no se limita a identificar plagio o reproducción de contenidos ajenos. Una producción puede ser formalmente original y, sin embargo, reflejar una participación humana insuficiente para demostrar la competencia evaluada. Navarrete-Cazales et al. (2026) señalan que estas tecnologías obligan a reconsiderar las relaciones entre autoría e integridad en educación superior. La responsabilidad académica comienza a relacionarse con el control intelectual sobre las decisiones, la verificación de contenidos, la capacidad para justificar lo presentado y la transparencia respecto de la intervención tecnológica.

La autoría necesita entenderse, en consecuencia, más allá de quién escribió materialmente cada palabra. En entornos mediados por inteligencia artificial, conservar una participación autoral significativa implica definir propósitos, seleccionar información, evaluar alternativas, transformar aportes y asumir responsabilidad sobre el resultado. Esta perspectiva permite diferenciar asistencia, colaboración y sustitución del trabajo académico. No todo uso tecnológico posee las mismas implicaciones: su legitimidad depende del propósito de la evaluación, del tipo de competencia que debe demostrarse y del grado de control que mantiene el estudiante sobre las decisiones fundamentales de su producción.

La transparencia constituye una condición necesaria dentro de este nuevo marco. Declarar la utilización de inteligencia artificial no debería entenderse automáticamente como reconocimiento de una conducta irregular, sino como información que permite interpretar cómo intervino la tecnología durante la actividad. UNESCO (2023) plantea la necesidad de promover una incorporación responsable y centrada en las personas. Para la universidad, ello implica establecer criterios comprensibles sobre usos permitidos, formas de declaración y situaciones donde determinadas intervenciones pueden resultar incompatibles con el aprendizaje esperado. La claridad contribuye a reducir ambigüedades y fortalece la responsabilidad compartida entre estudiantes, docentes e instituciones.

Este libro articula estas transformaciones mediante cinco ejes complementarios: el cambio de la evaluación universitaria, el fortalecimiento de la evaluación auténtica, la incorporación de evidencias de proceso y trazabilidad, la recuperación de modalidades orales, prácticas y situadas, y la redefinición de integridad y autoría académica. La cuestión central ya no consiste en decidir si la inteligencia artificial debe estar presente en la universidad, sino en determinar cómo evaluar cuando producir información resulta cada vez más sencillo. El desafío es construir sistemas donde comprender, argumentar, aplicar, revisar, transferir, justificar y asumir responsabilidad continúen siendo acciones visibles y atribuibles al estudiante. La respuesta universitaria no consiste en eliminar la inteligencia artificial, sino en rediseñar la evaluación para que el aprendizaje humano pueda seguir siendo demostrado.



Capítulo 1.

Transformación de la evaluación en la universidad

Joseph Michael Valero Rodríguez

1.1 La evaluación universitaria ante la expansión de la inteligencia artificial generativa

La expansión de la inteligencia artificial generativa ha introducido una transformación significativa en la manera en que las universidades conciben la evaluación del aprendizaje. La posibilidad de generar textos, explicaciones, síntesis, respuestas estructuradas y propuestas académicas en pocos segundos cuestiona la relación tradicional entre producción escrita y dominio del conocimiento. Villegas y Carballo (2026) advierten que este escenario obliga a revisar los procedimientos utilizados para interpretar el desempeño estudiantil, debido a que la presencia de sistemas generativos modifica las condiciones bajo las cuales se producen las evidencias académicas y exige criterios evaluativos ajustados a esta nueva realidad.

La incorporación de estas tecnologías no representa únicamente la aparición de una herramienta adicional dentro del entorno educativo. También altera los supuestos sobre los que históricamente se han construido muchas actividades de evaluación universitaria. Un producto correctamente redactado ya no constituye, por sí solo, una prueba suficiente de que el estudiante comprendió, analizó o elaboró intelectualmente el contenido presentado. La generación automatizada introduce una separación potencial entre la calidad visible del producto y el aprendizaje efectivamente alcanzado, situación que obliga a reconsiderar qué se evalúa y qué puede inferirse legítimamente a partir de una entrega académica (Wang et al., 2024).

Vera (2023) sostiene que la integración de inteligencia artificial en la educación superior plantea oportunidades importantes, pero también desafíos relacionados con los procesos pedagógicos y evaluativos. Desde esta perspectiva, el problema no consiste simplemente en permitir o prohibir determinadas herramientas, sino en reconocer que estas tecnologías forman parte de un ecosistema académico cada vez más digitalizado. La universidad debe responder con criterios educativos capaces de distinguir entre apoyo tecnológico, construcción intelectual y sustitución del esfuerzo cognitivo, evitando respuestas basadas únicamente en la vigilancia o en la prohibición generalizada del uso de sistemas generativos.

La rapidez con la que los sistemas generativos pueden producir contenidos introduce además una nueva tensión entre eficiencia y aprendizaje. Los estudiantes pueden utilizar estas herramientas para organizar información, explorar alternativas o mejorar

determinados productos, pero también pueden delegar procesos que tradicionalmente requerían comprensión, selección y elaboración personal. Por ello, la discusión evaluativa debe centrarse menos en la existencia de la tecnología y más en las condiciones pedagógicas bajo las cuales se incorpora. La evaluación necesita ofrecer evidencias suficientes para determinar qué parte del resultado corresponde realmente al razonamiento, las decisiones y la comprensión del estudiante (Avalos, 2024).

Uno de los cambios más relevantes se relaciona con la pérdida de certeza respecto del origen de algunas producciones académicas. Durante décadas, el docente pudo asumir razonablemente que un ensayo, una respuesta extensa o una síntesis presentada por un estudiante reflejaba, con diferentes niveles de calidad, un proceso de elaboración personal. Villegas y Carballo (2026) muestran que la inteligencia artificial generativa obliga a revisar esta asociación, porque actualmente una producción coherente puede surgir de interacciones complejas entre estudiante y sistema. La evaluación universitaria necesita, por tanto, considerar no solo el resultado visible, sino las circunstancias en que fue construido.

Esta situación no implica que la escritura académica, los informes o las tareas domiciliarias hayan perdido completamente su valor. El cambio radica en que ya no pueden interpretarse de manera aislada como evidencia automática de aprendizaje. La disponibilidad de herramientas generativas modifica las condiciones de producción y obliga a contextualizar mejor las evidencias obtenidas. La transformación evaluativa consiste, en buena medida, en reconocer que la calidad formal de un documento puede estar parcialmente mediada por tecnologías capaces de corregir, reorganizar o producir contenido, sin que ello permita determinar directamente el grado de comprensión alcanzado por quien presenta el trabajo (Wang et al., 2024).

Vera (2023) plantea que la inteligencia artificial puede convertirse en un recurso educativo valioso cuando su incorporación responde a propósitos pedagógicos definidos. Este planteamiento resulta particularmente importante para la evaluación, porque evita reducir el debate a una oposición entre tecnología y aprendizaje. El desafío universitario consiste en establecer condiciones que permitan aprovechar las capacidades de estas herramientas sin perder de vista la responsabilidad intelectual del estudiante. Evaluar después de la expansión de la IA generativa requiere reconocer que la tecnología puede formar parte del proceso académico, pero no debe sustituir la demostración de conocimientos, razonamientos y decisiones propias.

Otro efecto de esta transformación aparece en la relación entre evaluación y confianza. Cuando docentes e instituciones desconocen cómo se produjo un trabajo, puede surgir una tendencia a sospechar de cualquier resultado especialmente elaborado o técnicamente correcto. Ese escenario puede deteriorar la relación pedagógica si la evaluación se convierte exclusivamente en un mecanismo de detección. La respuesta requiere fortalecer diseños evaluativos que permitan observar mejor el aprendizaje y reducir la dependencia de inferencias basadas únicamente en el producto final. La evaluación debe generar condiciones para reconocer el desempeño del estudiante sin partir sistemáticamente de la presunción de fraude (Avalos, 2024).

La expansión de los sistemas generativos también evidencia que la evaluación universitaria necesita evolucionar al mismo ritmo que las formas de producción del conocimiento. Las prácticas académicas contemporáneas incluyen buscadores, gestores bibliográficos, plataformas virtuales, herramientas de corrección y, cada vez con mayor frecuencia, sistemas capaces de generar contenidos. Villegas y Carballo (2026) sitúan este proceso dentro de una revisión más amplia de la evaluación universitaria, en la que el uso de inteligencia artificial obliga a reconsiderar la validez de determinadas evidencias. La universidad no puede mantener intactos sus procedimientos cuando las condiciones tecnológicas que rodean al aprendizaje han cambiado sustancialmente.

El desafío es especialmente relevante porque la inteligencia artificial generativa no interviene únicamente al final de una tarea. Puede utilizarse desde la búsqueda inicial de ideas hasta la organización de argumentos, la formulación de ejemplos, la revisión lingüística o la construcción de una respuesta completa. Esta diversidad de usos dificulta establecer una frontera única entre producción humana y asistencia tecnológica. Por ello, la evaluación necesita reconocer distintos grados de intervención de la IA y valorar sus efectos sobre el aprendizaje, evitando asumir que todos los usos representan el mismo nivel de dependencia o la misma implicación académica (Wang et al., 2024).

Avalos (2024) destaca que la inteligencia artificial también abre posibilidades para transformar la retroalimentación y los procesos de evaluación, lo cual permite observar que el cambio no debe entenderse únicamente en términos de riesgo. Las mismas tecnologías que cuestionan ciertos formatos tradicionales pueden contribuir a generar nuevas oportunidades de acompañamiento, revisión y mejora. Sin embargo, su valor depende del diseño pedagógico que oriente su utilización. La expansión de la IA

generativa coloca a la universidad frente a la necesidad de redefinir sus expectativas evaluativas y establecer con mayor precisión qué evidencias deben demostrar los estudiantes durante su formación.

La evaluación universitaria entra así en una etapa en la que resulta insuficiente valorar únicamente la corrección formal de una respuesta o la calidad aparente de un producto académico. El escenario actual exige reconocer que el aprendizaje debe ser demostrado mediante evidencias capaces de mantener su valor incluso cuando existen herramientas que pueden producir contenido sofisticado. La transformación no consiste en abandonar las formas conocidas de evaluación, sino en revisar sus propósitos, alcances y condiciones de validez. Este cambio prepara el tránsito hacia modelos donde las competencias, los desempeños y la participación intelectual del estudiante adquieren una importancia creciente (Vera, 2023).

1.2 Del control de contenidos a la evaluación de competencias y desempeños

La transformación de la evaluación universitaria implica desplazar progresivamente el énfasis desde la comprobación de contenidos retenidos hacia la valoración de competencias capaces de movilizar conocimientos en situaciones académicas y profesionales. Ibarra et al. (2026) muestran que la evolución de la evaluación educativa responde a cambios en las concepciones sobre aprendizaje, desempeño y participación del estudiante. En este marco, conocer información continúa siendo necesario, pero deja de constituir el único referente. La evaluación comienza a interesarse también por la manera en que el estudiante interpreta, relaciona, aplica y comunica lo aprendido ante diferentes exigencias.

El control de contenidos suele apoyarse en evidencias centradas en recordar conceptos, reconocer información o reproducir explicaciones previamente estudiadas. Estas tareas pueden cumplir funciones diagnósticas o formativas, pero presentan limitaciones cuando se utilizan como único criterio para establecer el nivel de aprendizaje universitario. La formación superior requiere que el conocimiento pueda emplearse para resolver problemas, argumentar decisiones, analizar situaciones y producir respuestas contextualizadas. Esta orientación amplía el objeto de evaluación, porque incorpora capacidades que no siempre pueden observarse mediante pruebas centradas exclusivamente en respuestas correctas o en la reproducción literal de contenidos (Huauya et al., 2023).

Méndez et al. (2024) destacan la importancia del enfoque basado en competencias para relacionar los aprendizajes con desempeños que integren conocimientos, habilidades y capacidades de aplicación. Esta perspectiva modifica la pregunta central del proceso evaluativo: ya no interesa únicamente determinar cuánto contenido recuerda el estudiante, sino establecer qué puede hacer con aquello que conoce. La competencia se expresa cuando el saber es movilizado de manera pertinente frente a una situación determinada, por lo que su evaluación exige observar decisiones, procedimientos, argumentos y resultados vinculados con propósitos académicos o profesionales claramente definidos.

Este desplazamiento resulta especialmente relevante en un contexto donde la inteligencia artificial generativa puede producir respuestas informativas con rapidez y considerable corrección formal. Si una actividad evalúa solamente la recuperación de información o la elaboración de respuestas previsibles, existe mayor dificultad para determinar qué capacidad intelectual ha demostrado realmente el estudiante. La evaluación de competencias introduce una exigencia distinta: relacionar información, justificar elecciones, adaptar conocimientos y responder ante condiciones específicas. De esta manera, el valor de la evaluación depende menos de la cantidad de contenido reproducido y más de la calidad del desempeño evidenciado (Bearman et al., 2024).

Tabla 1
Diferencias funcionales entre el control de contenidos y la evaluación de competencias y desempeños

Dimensión	Control de contenidos	Evaluación de competencias y desempeños	Implicación para la evaluación universitaria
Objeto principal	Información adquirida	Movilización integrada de saberes	Amplía lo que se considera evidencia de aprendizaje
Evidencia habitual	Respuesta correcta o reproducción conceptual	Decisión, procedimiento, argumentación o actuación	Requiere observar procesos y resultados
Contextualización	General o descontextualizada	Vinculada con situaciones específicas	Favorece tareas con condiciones y propósitos definidos
Papel del estudiante	Responder a lo solicitado	Interpretar, seleccionar, aplicar y justificar	Incrementa su participación intelectual

Dimensión	Control de contenidos	Evaluación de competencias y desempeños	Implicación para la evaluación universitaria
Criterio de calidad	Exactitud del contenido	Pertinencia, aplicación y calidad del desempeño	Exige criterios de valoración más diversos
Relación con la IA generativa	Puede facilitar respuestas directas	Obliga a demostrar apropiación y uso del conocimiento	Reduce la dependencia exclusiva del producto informativo

Nota. La comparación muestra un cambio en el foco de la evaluación y no supone eliminar la valoración de conocimientos conceptuales, sino integrarlos con capacidades de aplicación, razonamiento y desempeño.

La transición hacia competencias tampoco significa que los contenidos pierdan importancia dentro de la formación universitaria. Huauya et al. (2023) permiten comprender que el desarrollo competencial necesita conocimientos que puedan ser activados y articulados durante el aprendizaje. El cambio consiste en dejar de tratarlos como resultado final suficiente. Un estudiante puede recordar definiciones y, sin embargo, presentar dificultades para seleccionar conceptos pertinentes ante un caso, interpretar información contradictoria o fundamentar una decisión. La evaluación debe identificar esa diferencia y valorar hasta qué punto el conocimiento disponible puede convertirse en una actuación intelectualmente consistente.

La noción de desempeño complementa esta perspectiva porque permite observar la competencia mediante acciones concretas. Un desempeño puede manifestarse cuando el estudiante explica un procedimiento, resuelve una situación, construye una propuesta, defiende una elección o utiliza conocimientos disciplinares para responder ante determinadas condiciones. Méndez et al. (2024) relacionan el enfoque competencial con cambios en las experiencias educativas y en las formas de valorar el rendimiento. Desde esta lógica, la evaluación deja de limitarse al resultado cuantificable y se aproxima a evidencias que permiten reconocer cómo se articulan diferentes capacidades durante la realización de una tarea.

La adopción de este enfoque exige también revisar los criterios empleados por el docente. Cuando se evalúan contenidos, es frecuente establecer respuestas esperadas relativamente delimitadas; cuando se evalúan competencias, resulta necesario considerar dimensiones como pertinencia, razonamiento, aplicación, coherencia y capacidad de adaptación. Bearman et al. (2024) destacan la importancia del juicio

evaluativo en escenarios donde los estudiantes deben distinguir la calidad de diferentes producciones y decisiones. Esto implica construir criterios suficientemente claros para orientar el desempeño sin reducir las actividades a respuestas predeterminadas que limiten la posibilidad de demostrar comprensión y autonomía.

El cambio tiene además consecuencias sobre el papel del error dentro de la evaluación. En modelos centrados fundamentalmente en el control de contenidos, una respuesta incorrecta suele interpretarse como ausencia o insuficiencia de conocimiento. En una evaluación orientada al desempeño, el error puede proporcionar información sobre estrategias, interpretaciones, decisiones o dificultades específicas durante la aplicación del saber. Ibarra et al. (2026) permiten situar esta transformación dentro de una evolución hacia funciones evaluativas más amplias. Analizar cómo se produce una respuesta puede resultar tan relevante como determinar si el resultado final coincide con una solución previamente establecida.

Otro aspecto relevante es la contextualización. Las competencias adquieren significado cuando el estudiante debe movilizar aprendizajes frente a condiciones concretas, restricciones o propósitos definidos. Una misma base conceptual puede utilizarse de manera diferente según el problema planteado, por lo que la evaluación debe generar oportunidades para observar esa capacidad de adaptación. El enfoque competencial favorece así una relación más estrecha entre formación universitaria y situaciones que demandan interpretación y actuación, sin que ello implique convertir toda evaluación en una reproducción exacta del ejercicio profesional. Lo importante es comprobar la funcionalidad del conocimiento adquirido (Méndez et al., 2024).

La evaluación de competencias también modifica la interpretación del rendimiento académico. Obtener una calificación elevada en una prueba de contenidos no garantiza necesariamente que el estudiante pueda integrar esos conocimientos en una situación diferente, del mismo modo que una actuación competente requiere una base conceptual suficiente para sostener decisiones fundamentadas. Huauya et al. (2023) muestran la relación entre evaluación formativa y desarrollo de competencias, lo que refuerza la necesidad de utilizar evidencias que permitan reconocer avances y dificultades en dimensiones diversas. La valoración del aprendizaje se vuelve, de este modo, más amplia que la simple acumulación de respuestas correctas.

Este cambio de enfoque adquiere especial importancia frente a tecnologías capaces de recuperar y reorganizar información con una eficacia creciente. El desafío universitario consiste en preservar el conocimiento disciplinar como fundamento, pero situarlo dentro de evaluaciones que demanden comprensión, aplicación, razonamiento y toma de decisiones. Bearman et al. (2024) permiten vincular esta transformación con la necesidad de fortalecer la capacidad de juzgar la calidad de respuestas y producciones. La evaluación universitaria avanza así desde la comprobación de contenidos hacia evidencias de competencia y desempeño, preparando el terreno para revisar las prácticas evaluativas que todavía dependen excesivamente de formatos tradicionales.

1.3 Cambios en las prácticas tradicionales de evaluación académica

Las prácticas tradicionales de evaluación académica han estado asociadas durante décadas con exámenes escritos, cuestionarios, ensayos, informes y tareas desarrolladas bajo criterios relativamente estables de corrección. La expansión de la inteligencia artificial generativa ha alterado este escenario porque algunas de esas actividades pueden ser resueltas, total o parcialmente, mediante sistemas capaces de producir respuestas estructuradas en pocos segundos. Avalos (2024) señala que esta transformación exige revisar el sentido de la evaluación y la calidad de las evidencias obtenidas, especialmente cuando el producto presentado ya no permite identificar con claridad cuánto del trabajo corresponde a elaboración propia.

Uno de los principales cambios afecta a las tareas que solicitan reproducir, resumir o reorganizar información disponible. Estas actividades pueden seguir teniendo utilidad pedagógica, pero ofrecen menor capacidad para discriminar entre comprensión auténtica y generación automatizada cuando se formulan de manera genérica. Naveda (2025) plantea que la incorporación de inteligencia artificial en los procesos evaluativos demanda nuevas formas de seguimiento y retroalimentación. En este contexto, las prácticas tradicionales necesitan evolucionar hacia actividades que permitan observar interpretación, toma de decisiones y justificación, evitando que la evaluación dependa exclusivamente de productos que pueden ser generados con escasa participación intelectual del estudiante.

Tabla 2

Transformaciones necesarias en prácticas tradicionales de evaluación académica

Práctica tradicional	Riesgo en el contexto de IA generativa	Ajuste evaluativo necesario	Evidencia que conviene priorizar
Ensayo domiciliario genérico	Generación automática de argumentos y estructura	Incorporar contexto, decisiones y defensa del contenido	Justificación de elecciones y proceso de elaboración
Cuestionario de respuesta extensa	Producción inmediata de respuestas plausibles	Añadir preguntas de aplicación y contraste	Razonamiento y capacidad de transferencia
Resumen académico	Síntesis automatizada de textos	Solicitar interpretación y relación con casos específicos	Comprensión y selección de información relevante
Informe descriptivo	Elaboración rápida de contenido formalmente correcto	Exigir análisis, evidencia situada y criterios explícitos	Capacidad de análisis y fundamentación
Tarea domiciliaria aislada	Dificultad para conocer el proceso seguido	Incorporar avances, revisión y retroalimentación	Evolución del aprendizaje
Examen exclusivamente memorístico	Menor capacidad para valorar comprensión profunda	Combinar conocimiento con aplicación y argumentación	Dominio conceptual y desempeño

Nota. Los ajustes propuestos no implican eliminar los formatos tradicionales, sino modificar su diseño para incrementar la capacidad de obtener evidencias válidas de aprendizaje en contextos donde existen herramientas generativas.

Los exámenes también enfrentan transformaciones, aunque su vulnerabilidad depende de las condiciones en que se desarrollan. Las pruebas presenciales y supervisadas continúan ofreciendo posibilidades de control sobre el entorno, pero una evaluación centrada únicamente en recordar información puede resultar limitada para valorar capacidades complejas. Heil e Ifenthaler (2023) muestran que la evaluación en entornos digitales amplía las posibilidades de formatos y procedimientos, pero también obliga a reconsiderar aspectos relacionados con validez, autenticidad y diseño. El reto no consiste únicamente en trasladar exámenes tradicionales a plataformas digitales, sino en revisar aquello que realmente se pretende observar mediante ellos.

Los ensayos académicos requieren una consideración particular porque durante mucho tiempo han sido utilizados para valorar comprensión, argumentación y capacidad de

escritura. La inteligencia artificial generativa puede intervenir en todas estas dimensiones, desde la organización inicial hasta la producción de párrafos completos. Pearce y Chiavaroli (2023) sostienen que la aparición de estas herramientas obliga a reconsiderar el diseño de la evaluación, precisamente porque la calidad del producto final puede dejar de representar de manera directa la capacidad del estudiante. El ensayo mantiene valor cuando se inserta en un proceso evaluativo que permita comprobar decisiones, razonamiento y apropiación intelectual.

También se modifica la lógica de las tareas desarrolladas fuera del aula. Tradicionalmente, el trabajo domiciliario ofrecía tiempo para investigar, consultar fuentes y construir respuestas más elaboradas. En la actualidad, ese mismo espacio permite utilizar sistemas generativos de manera intensiva, lo que hace necesario reconsiderar la forma en que se interpreta la evidencia producida. Avalos (2024) permite observar que la transformación evaluativa no depende únicamente de la tecnología disponible, sino de cómo se diseñan las actividades y de los criterios utilizados para valorar sus resultados. El problema central está en la capacidad de la tarea para mostrar aprendizaje, no en su modalidad presencial o remota.

Las prácticas tradicionales también están cambiando en relación con la retroalimentación. El modelo centrado en una calificación final ofrece información limitada cuando se pretende comprender cómo aprendió el estudiante o qué aspectos necesita mejorar. Naveda (2025) destaca el potencial de la inteligencia artificial dentro de la evaluación formativa, lo que refuerza la necesidad de utilizar la evaluación como proceso y no únicamente como mecanismo de certificación. Esta transformación favorece actividades con entregas parciales, revisión, comentarios y oportunidades de mejora, aunque el desarrollo detallado de estas evidencias de proceso corresponde a etapas posteriores del análisis evaluativo.

Otro cambio importante consiste en abandonar la idea de que mayor control equivale necesariamente a mejor evaluación. La supervisión, las restricciones y los mecanismos de autenticación pueden ser necesarios en determinadas situaciones, pero no sustituyen un diseño capaz de producir evidencias significativas. Heil e Ifenthaler (2023) muestran que la evaluación en educación superior puede adoptar configuraciones diversas según los objetivos, recursos y condiciones del aprendizaje. En consecuencia, una respuesta universitaria frente a la IA generativa no debería limitarse a aumentar la vigilancia, sino

combinar condiciones de control con actividades que permitan demostrar comprensión, aplicación y razonamiento.

La revisión de prácticas tradicionales tampoco exige eliminar automáticamente exámenes, informes, ensayos o cuestionarios. El cambio más relevante consiste en modificar la función que cumplen dentro de un sistema evaluativo más amplio. Un examen puede seguir siendo útil para comprobar conocimiento conceptual; un ensayo puede valorar argumentación; y un informe puede evidenciar capacidad de análisis. Sin embargo, estos instrumentos adquieren mayor solidez cuando se complementan con otras evidencias y cuando sus consignas demandan decisiones que no puedan resolverse mediante respuestas genéricas. Pearce y Chiavaroli (2023) permiten situar esta adaptación como parte de una reconsideración más profunda del diseño evaluativo.

La transformación de las prácticas académicas conduce así a una evaluación menos dependiente de formatos rutinarios y más atenta a la calidad de las evidencias que cada actividad produce. El desafío consiste en conservar aquello que sigue siendo pedagógicamente útil y modificar lo que ha perdido capacidad para demostrar aprendizaje en las nuevas condiciones tecnológicas. Naveda (2025) permite comprender esta transición desde una perspectiva formativa, donde evaluar implica acompañar y obtener información relevante sobre el desempeño. Este desplazamiento prepara el análisis de nuevas formas de aprender, producir y demostrar conocimiento dentro de entornos universitarios atravesados por inteligencia artificial generativa.

1.4 Nuevas formas de aprender, producir y demostrar conocimiento

Las nuevas formas de aprender en la universidad están cada vez más vinculadas con procesos de interacción, exploración y revisión apoyados por tecnologías capaces de generar contenido. Chen et al. (2025) muestran que la alfabetización en inteligencia artificial implica no solo utilizar herramientas, sino comprender cómo interactuar con ellas, evaluar sus respuestas y reconocer sus implicaciones. En este escenario, aprender deja de asociarse exclusivamente con recibir y reproducir información, porque el estudiante puede contrastar alternativas, formular nuevas preguntas y reconstruir conocimientos a partir de intercambios sucesivos que requieren criterio, supervisión y participación intelectual activa.

La producción académica también adquiere características más iterativas. Un texto, una propuesta o una solución pueden construirse mediante sucesivas etapas de búsqueda,

generación, selección, corrección y reformulación, en lugar de desarrollarse como un producto lineal elaborado desde el inicio hasta su entrega. Esta dinámica modifica la relación entre estudiante y contenido, porque producir conocimiento implica tomar decisiones sobre aquello que se conserva, transforma o descarta. La intervención tecnológica puede ampliar posibilidades creativas y analíticas, pero su valor formativo depende de que el estudiante mantenga control sobre el propósito, la pertinencia y la calidad de lo producido (Tzirides et al., 2024).

Ossa y Willatt (2023) evidencian que la inteligencia artificial generativa puede participar en procesos de retroalimentación de la escritura académica, ofreciendo oportunidades para revisar formulaciones, identificar alternativas y mejorar determinados aspectos del texto. Esta posibilidad transforma la producción escrita en una actividad potencialmente más dialógica, donde el estudiante puede comparar versiones y reconsiderar decisiones antes de presentar un resultado definitivo. Sin embargo, el aprendizaje no se encuentra en aceptar automáticamente las sugerencias recibidas, sino en analizar su pertinencia, justificar modificaciones y conservar una participación consciente en la construcción de los argumentos y significados desarrollados.

La disponibilidad de sistemas generativos modifica también la manera en que los estudiantes acceden a explicaciones y ejemplos. Una duda puede ser abordada mediante preguntas sucesivas, reformulaciones o solicitudes de ejemplos adaptados a distintos niveles de complejidad. Esto permite personalizar ciertos momentos del aprendizaje y ampliar las oportunidades de exploración fuera del aula. Aun así, la utilidad educativa de estas interacciones depende de la capacidad para identificar errores, simplificaciones o respuestas insuficientemente fundamentadas. La generación inmediata de contenido no garantiza aprendizaje, porque comprender exige interpretar, relacionar y evaluar críticamente la información obtenida (Chen et al., 2025).

Bearman et al. (2024) destacan la importancia del juicio evaluativo en un contexto donde los estudiantes disponen de herramientas capaces de generar producciones plausibles y formalmente elaboradas. Esta capacidad de juzgar adquiere especial relevancia porque aprender ya no consiste únicamente en producir una respuesta, sino también en reconocer qué características hacen que una respuesta sea adecuada, limitada o mejorable. Comparar alternativas, identificar inconsistencias y seleccionar opciones fundamentadas se convierten en actividades cognitivas centrales. La

universidad necesita, por tanto, promover experiencias donde el estudiante participe activamente en la valoración de la calidad del conocimiento que utiliza y produce.

Las nuevas formas de demostrar conocimiento también amplían el repertorio de evidencias disponibles. El aprendizaje puede manifestarse mediante explicaciones razonadas, análisis de alternativas, productos multimodales, propuestas aplicadas o decisiones justificadas, sin quedar restringido a una única forma de presentación. Esta diversidad permite reconocer capacidades que no siempre se observan adecuadamente mediante formatos uniformes. La presencia de inteligencia artificial refuerza esta necesidad porque un producto convencional puede resultar insuficiente para mostrar la participación intelectual del estudiante. Demostrar conocimiento implica evidenciar comprensión, criterio y capacidad de adaptación frente a situaciones concretas (Tzirides et al., 2024).

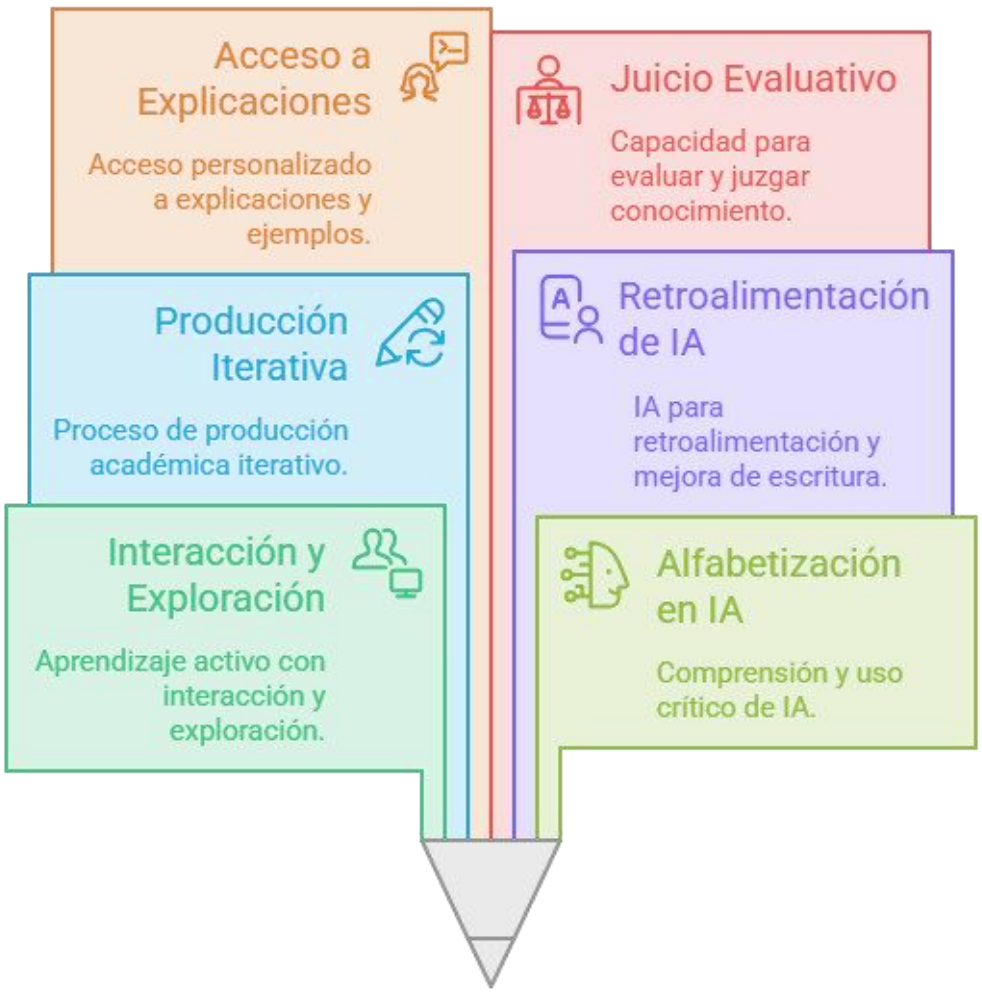
La producción universitaria puede incorporar además procesos de coedición en los que el estudiante utiliza herramientas generativas para ensayar estructuras, revisar claridad o explorar posibles enfoques. Ossa y Willatt (2023) permiten observar que esta interacción puede adquirir valor formativo cuando funciona como apoyo para revisar la escritura y no como sustitución del proceso intelectual. El estudiante necesita decidir qué sugerencias aceptar, cuáles modificar y cuáles rechazar, lo que introduce una dimensión crítica dentro de la producción académica. La calidad del aprendizaje se relaciona entonces con la capacidad de mantener agencia sobre las decisiones que configuran el resultado.

Otra transformación aparece en la relación entre aprender y preguntar. Frente a modelos donde el estudiante recibe información previamente organizada, los sistemas generativos permiten formular consultas, solicitar explicaciones alternativas y profundizar selectivamente en determinados aspectos. Esta interacción puede favorecer procesos exploratorios si las preguntas evolucionan conforme aumenta la comprensión. Sin embargo, una secuencia de respuestas generadas automáticamente puede producir una apariencia de dominio sin aprendizaje profundo cuando el estudiante no examina críticamente el contenido. La alfabetización en inteligencia artificial requiere, por ello, competencias para formular, interpretar y valorar interacciones de manera consciente (Chen et al., 2025).

Tzirides et al. (2024) plantean la alfabetización en inteligencia artificial como una capacidad que combina participación humana y recursos tecnológicos, lo que resulta especialmente pertinente para comprender las nuevas formas de producción del conocimiento. El estudiante no actúa únicamente como receptor de resultados generados, sino como participante que formula propósitos, establece criterios y decide cómo integrar la información obtenida. Esta relación modifica la idea de autoría puramente individual sin eliminar la responsabilidad humana. La producción académica adquiere valor cuando la tecnología amplía posibilidades de análisis y creación, pero las decisiones fundamentales permanecen vinculadas con el juicio del estudiante.

Figura 1

El futuro de la educación y sus principales transformaciones



Nota. Se ilustran las tendencias educativas emergentes que impulsan nuevas formas de enseñanza y aprendizaje mediante la innovación tecnológica y el desarrollo de competencias.

La demostración del aprendizaje también exige mayor atención a la capacidad de explicar las decisiones tomadas durante una producción. Un estudiante puede presentar un resultado correcto, pero su comprensión se hace más visible cuando puede justificar

por qué eligió determinada alternativa, cómo evaluó otras posibilidades o qué criterios utilizó para modificar una propuesta. Bearman et al. (2024) relacionan esta capacidad con el desarrollo del juicio evaluativo, particularmente relevante cuando existen sistemas capaces de generar múltiples soluciones plausibles. La evidencia de aprendizaje se desplaza así desde la mera posesión de una respuesta hacia la capacidad de comprender y defender su calidad.

Estas transformaciones muestran que aprender, producir y demostrar conocimiento son procesos cada vez menos separables dentro de la experiencia universitaria. La interacción con inteligencia artificial puede acompañar la exploración, la escritura, la revisión y la valoración, pero cada una de estas acciones requiere participación intelectual para adquirir significado educativo. La universidad necesita reconocer esta nueva ecología del aprendizaje sin reducirla a dependencia tecnológica ni conservar modelos que ignoren las condiciones actuales de producción académica. La ampliación de posibilidades exige, al mismo tiempo, nuevas responsabilidades docentes para orientar y diseñar experiencias evaluativas coherentes con estos cambios (Ossa y Willatt, 2023).

1.5 El papel del docente en el rediseño de la evaluación universitaria

El rediseño de la evaluación universitaria frente a la inteligencia artificial generativa exige una transformación del papel docente, que ya no puede limitarse a seleccionar instrumentos, asignar calificaciones y comprobar respuestas finales. Sánchez (2023) sostiene que la incorporación de inteligencia artificial en la educación superior amplía las posibilidades pedagógicas, pero también obliga al profesorado a desarrollar criterios para integrarla de manera pertinente. En este escenario, el docente asume una función de diseñador de experiencias evaluativas capaces de obtener evidencias relevantes sobre comprensión, aplicación, razonamiento y participación intelectual del estudiante.

Esta transformación comienza con la definición precisa de aquello que se pretende evaluar. Antes de seleccionar una tarea, el docente necesita distinguir si busca comprobar conocimientos conceptuales, capacidad de argumentación, aplicación profesional, resolución de problemas o integración de diferentes competencias. La existencia de herramientas generativas vuelve especialmente importante esta decisión, porque una actividad mal alineada con los resultados de aprendizaje puede producir evidencias formalmente correctas, pero poco informativas. Rediseñar la evaluación

supone establecer correspondencia entre propósito formativo, actividad solicitada, evidencia esperada y criterios utilizados para valorar el desempeño alcanzado por el estudiante.

Galindo-Domínguez et al. (2024) muestran que la inteligencia artificial puede participar incluso en procesos vinculados con la evaluación, lo que exige conservar el juicio profesional del docente como componente central de las decisiones académicas. Los sistemas automatizados pueden aportar información, sugerencias o valoraciones, pero su intervención no elimina la necesidad de interpretar el contexto, las características de la tarea y el desempeño observado. El profesorado debe valorar críticamente las respuestas producidas por estas tecnologías y determinar hasta qué punto resultan coherentes con los objetivos formativos, evitando delegar decisiones evaluativas que requieren comprensión pedagógica y responsabilidad académica.

El docente también necesita anticipar cómo una herramienta generativa podría intervenir en una actividad antes de aplicarla. Esta previsión permite identificar qué partes pueden resolverse mediante producción automatizada y cuáles exigen decisiones, experiencias o razonamientos que deben provenir del estudiante. No se trata de diseñar tareas imposibles de realizar con inteligencia artificial, sino de establecer condiciones donde su eventual utilización no elimine la necesidad de aprender. Una evaluación sólida puede permitir determinados apoyos tecnológicos y, simultáneamente, exigir explicación, contextualización, verificación o adaptación, de modo que el desempeño siga siendo intelectualmente significativo.

Tabla 3

Funciones docentes en el rediseño de la evaluación universitaria ante la inteligencia artificial generativa

Función docente	Pregunta orientadora	Decisión evaluativa asociada	Resultado esperado
Definir el propósito	¿Qué aprendizaje debe demostrarse?	Precisar conocimientos, competencias o desempeños	Evidencias alineadas con los resultados de aprendizaje
Anticipar el uso de IA	¿Qué partes de la tarea pueden automatizarse?	Ajustar consignas, condiciones y exigencias	Participación intelectual observable
Diseñar la actividad	¿Qué debe hacer el estudiante para demostrar aprendizaje?	Incorporar análisis, aplicación y decisiones	Tareas con mayor valor evaluativo

Función docente	Pregunta orientadora	Decisión evaluativa asociada	Resultado esperado
Establecer criterios	¿Qué características tendrá un desempeño de calidad?	Construir criterios y rúbricas pertinentes	Valoración transparente y consistente
Interpretar evidencias	¿Qué demuestra realmente el producto presentado?	Contrastar resultados, proceso y desempeño	Juicios docentes mejor fundamentados
Orientar el uso tecnológico	¿Cómo puede emplearse la IA sin sustituir el aprendizaje?	Delimitar usos pedagógicamente aceptables	Integración responsable de herramientas generativas

Nota. Las funciones se presentan como componentes articulados del diseño evaluativo; su aplicación depende de los resultados de aprendizaje, las características de la asignatura y las condiciones en las que se desarrolla la evaluación.

La formulación de consignas adquiere especial relevancia dentro de este nuevo papel docente. Una instrucción genérica puede facilitar respuestas estandarizadas, mientras que una consigna construida alrededor de condiciones específicas puede demandar interpretación y toma de decisiones. Sánchez (2023) señala que el profesorado necesita desarrollar capacidades para integrar la inteligencia artificial como recurso educativo, lo que también implica comprender sus alcances y limitaciones. El rediseño requiere formular tareas suficientemente claras para orientar al estudiante, pero abiertas a procesos de razonamiento que permitan diferenciar entre reproducción automática y construcción significativa del conocimiento.

Otra responsabilidad consiste en diseñar criterios de evaluación capaces de valorar algo más que la presentación formal del producto. La corrección lingüística, la estructura o la apariencia profesional pueden mejorarse mediante herramientas automatizadas, por lo que el docente necesita incorporar dimensiones relacionadas con fundamentación, pertinencia, razonamiento, aplicación y capacidad para justificar decisiones. Los criterios deben hacer visible aquello que la asignatura considera aprendizaje valioso y ofrecer referencias comprensibles para estudiantes y evaluadores. Esta claridad favorece una valoración más consistente y reduce la dependencia de impresiones generales sobre productos que pueden haber recibido diferentes niveles de asistencia tecnológica.

Khosravi et al. (2022) destacan la relevancia de comprender y explicar las decisiones producidas por sistemas de inteligencia artificial, aspecto que también resulta pertinente para el trabajo docente. Si una herramienta participa en la retroalimentación,

clasificación o análisis de producciones estudiantiles, el profesor necesita conocer las razones y limitaciones asociadas con sus resultados antes de incorporarlos a una decisión evaluativa. La automatización puede apoyar determinadas tareas, pero la responsabilidad sobre las consecuencias académicas permanece vinculada al juicio humano. El docente actúa, así como mediador entre posibilidades tecnológicas, criterios pedagógicos y necesidades concretas de aprendizaje.

El rediseño también modifica la relación del docente con la retroalimentación. En lugar de concentrarse exclusivamente en señalar errores después de una entrega, puede organizar oportunidades para que el estudiante interprete observaciones, revise decisiones y mejore progresivamente su desempeño. La inteligencia artificial puede ampliar las posibilidades de generar comentarios o apoyos, pero su presencia no sustituye la necesidad de retroalimentación vinculada con objetivos, criterios y particularidades del proceso educativo. La función docente consiste en asegurar que la información recibida contribuya realmente al aprendizaje y no se convierta en una acumulación automática de recomendaciones sin jerarquía ni propósito formativo.

La orientación sobre el uso de inteligencia artificial constituye igualmente una responsabilidad pedagógica. UNESCO (2023) plantea la necesidad de desarrollar capacidades para utilizar estas tecnologías de manera crítica, segura y centrada en propósitos educativos. En la evaluación, ello supone comunicar qué usos son compatibles con la actividad, qué decisiones deben permanecer bajo responsabilidad del estudiante y qué evidencias permiten reconocer su participación. Esta orientación no equivale a establecer una política institucional completa, sino a proporcionar reglas pedagógicas claras dentro de cada experiencia evaluativa para que la tecnología pueda utilizarse sin desdibujar los objetivos de aprendizaje.

El papel docente también requiere revisar continuamente la eficacia de las evaluaciones implementadas. Una actividad puede haber funcionado adecuadamente durante años y perder capacidad para producir evidencias válidas cuando cambian las herramientas disponibles o las formas de trabajo académico. Galindo-Domínguez et al. (2024) muestran que la interacción entre inteligencia artificial y evaluación plantea nuevas posibilidades de análisis, pero también exige examinar críticamente los resultados obtenidos. El profesorado necesita observar patrones de respuesta, dificultades recurrentes y formas emergentes de uso tecnológico para ajustar progresivamente tareas, criterios y procedimientos sin depender de soluciones permanentes.

El rediseño de la evaluación universitaria sitúa, por tanto, al docente como responsable de construir condiciones en las que el aprendizaje pueda hacerse visible de manera confiable y pedagógicamente pertinente. Su función combina planificación, anticipación tecnológica, formulación de consignas, establecimiento de criterios, interpretación de evidencias y orientación del estudiante. Esta transformación no reduce la importancia del profesor frente a sistemas generativos; por el contrario, incrementa la necesidad de decisiones profesionales capaces de otorgar sentido educativo a las herramientas disponibles. Desde esta función renovada se vuelve indispensable fortalecer también la alfabetización en inteligencia artificial de docentes y estudiantes.

1.6 Alfabetización en inteligencia artificial para docentes y estudiantes

La alfabetización en inteligencia artificial se ha convertido en una competencia necesaria para participar de manera informada en entornos universitarios donde las herramientas generativas intervienen cada vez con mayor frecuencia. Hornberger et al. (2023) señalan que comprender la inteligencia artificial implica reconocer sus principios básicos, posibilidades y limitaciones, no solo saber utilizar una aplicación concreta. En el ámbito educativo, esta alfabetización permite que docentes y estudiantes interpreten mejor el funcionamiento de los sistemas, formulen expectativas realistas y eviten asumir que una respuesta técnicamente elaborada representa necesariamente conocimiento verdadero, suficiente o pedagógicamente pertinente.

Para los estudiantes, alfabetizarse en inteligencia artificial significa aprender a interactuar con estas herramientas de manera estratégica y crítica. No basta con redactar instrucciones o solicitar respuestas; también es necesario comprender qué información proporcionar, cómo revisar los resultados y qué criterios utilizar para determinar su calidad. Esta capacidad adquiere especial importancia en contextos académicos, donde una respuesta puede parecer convincente y, sin embargo, contener errores, omisiones o afirmaciones difíciles de verificar. La alfabetización debe permitir al estudiante mantener control intelectual sobre el proceso y no limitarse a aceptar automáticamente la producción generada (Chen et al., 2025).

Jin et al. (2025) plantean la alfabetización en inteligencia artificial como una capacidad susceptible de ser evaluada mediante dimensiones relacionadas con comprensión, uso y valoración de sistemas generativos. Esta perspectiva resulta relevante para la universidad porque permite superar una visión instrumental centrada exclusivamente

en saber “usar” una herramienta. Un estudiante alfabetizado necesita reconocer cuándo una tecnología puede ser útil, qué riesgos introduce y qué acciones de verificación resultan necesarias. De esta manera, la competencia tecnológica se articula con habilidades de análisis, selección y juicio que fortalecen la participación consciente en actividades académicas mediadas por inteligencia artificial.

La alfabetización docente presenta exigencias particulares porque el profesor necesita comprender la tecnología no solo como usuario, sino también como responsable del diseño pedagógico. Debe identificar qué tareas pueden ser transformadas por sistemas generativos, qué limitaciones presentan sus resultados y cómo orientar a los estudiantes en su utilización. Esta preparación influye directamente en la calidad de las decisiones evaluativas, ya que permite distinguir entre usos que favorecen la exploración y aquellos que pueden reducir oportunidades de aprendizaje. El desarrollo de estas capacidades requiere formación continua y reflexión sobre las consecuencias educativas de cada integración tecnológica (Tzirides et al., 2024).

Hornberger et al. (2023) muestran que el conocimiento sobre inteligencia artificial no puede darse por supuesto entre estudiantes universitarios. La familiaridad cotidiana con herramientas digitales no garantiza comprensión de los principios que orientan su funcionamiento ni capacidad para interpretar críticamente sus resultados. Esta diferencia es importante porque el uso frecuente puede generar una sensación de dominio que no necesariamente corresponde con una alfabetización suficiente. La universidad necesita crear oportunidades formativas que permitan comprender conceptos básicos, reconocer límites y desarrollar criterios para evaluar la confiabilidad y pertinencia de las respuestas producidas por sistemas automatizados.

Un componente central de esta alfabetización es la capacidad para evaluar resultados generados. Los modelos pueden producir respuestas coherentes, bien organizadas y lingüísticamente correctas, pero ello no asegura exactitud conceptual, vigencia de la información o adecuación al contexto solicitado. Por esta razón, el estudiante debe aprender a contrastar datos, revisar fuentes, identificar inconsistencias y reconocer cuándo una respuesta requiere comprobación adicional. Esta habilidad transforma el uso de inteligencia artificial en una actividad intelectualmente exigente y evita que la facilidad de acceso a contenido reduzca la participación crítica en el aprendizaje (Chen et al., 2025).

La formulación de instrucciones constituye otra dimensión importante, aunque no debe confundirse con alfabetización completa. Jin et al. (2025) permiten comprender que interactuar eficazmente con inteligencia artificial implica más que elaborar prompts detallados. El estudiante necesita definir un propósito, proporcionar información relevante, interpretar la respuesta y ajustar la interacción cuando los resultados no cumplen con los criterios establecidos. La calidad de la consulta depende de la claridad del objetivo académico y de la capacidad para reconocer qué aspectos requieren precisión adicional, lo que convierte la interacción con sistemas generativos en un proceso de decisión y revisión constante.

Para el profesorado, esta competencia también implica aprender a reconocer diferentes niveles de intervención tecnológica en los trabajos académicos. No todos los usos de inteligencia artificial tienen el mismo efecto sobre el aprendizaje: solicitar una explicación, comparar alternativas, revisar estilo o generar una respuesta completa representan formas distintas de participación. Comprender estas diferencias permite formular orientaciones más precisas y evitar reglas excesivamente generales. Una alfabetización docente sólida facilita decisiones basadas en el propósito de la actividad y en el tipo de aprendizaje esperado, en lugar de depender exclusivamente de prohibiciones o permisos indiferenciados (Tzirides et al., 2024).

Chen et al. (2025) destacan además la dimensión ética asociada con la alfabetización en inteligencia artificial, vinculada con la evaluación responsable de las implicaciones de su uso. En el contexto universitario, esta dimensión debe desarrollarse sin convertir el aprendizaje tecnológico en una discusión exclusivamente normativa. El estudiante necesita comprender que sus decisiones tienen consecuencias sobre la calidad de su aprendizaje, la confiabilidad de sus productos y la responsabilidad que asume sobre lo presentado. La formación debe favorecer una relación consciente con la tecnología, donde utilizar una herramienta implique también saber cuándo limitar su intervención y cuándo recurrir al juicio propio.

La alfabetización también contribuye a reducir desigualdades derivadas de diferencias en acceso, experiencia y conocimiento tecnológico. Cuando el uso de inteligencia artificial se incorpora a actividades académicas sin formación previa, algunos estudiantes pueden beneficiarse de competencias adquiridas fuera de la universidad mientras otros enfrentan mayores dificultades para aprovechar las mismas herramientas. Hornberger et al. (2023) permiten reconocer la necesidad de evaluar y

desarrollar estas capacidades de manera sistemática. La formación universitaria debe procurar condiciones donde el aprendizaje sobre inteligencia artificial no dependa únicamente de experiencias personales o de la familiaridad previa con determinadas plataformas.

Figura 2

Equilibrio de competencias en inteligencia artificial entre estudiantes y docentes



Nota. Se representan las competencias complementarias de estudiantes y docentes para integrar la inteligencia artificial mediante un uso crítico, ético y pedagógicamente responsable.

La construcción conjunta de alfabetización entre docentes y estudiantes puede favorecer una cultura académica más crítica frente a la inteligencia artificial. El profesorado necesita comprender cómo se utilizan estas herramientas y los estudiantes requieren criterios para interpretar sus resultados, lo que convierte la formación en un proceso compartido de actualización. Tzirides et al. (2024) proponen una perspectiva donde capacidades humanas y recursos de inteligencia artificial pueden complementarse cuando existen conocimientos suficientes para orientar la interacción. Esta relación resulta más productiva cuando la tecnología amplía posibilidades de aprendizaje sin desplazar el razonamiento, la responsabilidad ni la capacidad de juicio.

Fortalecer la alfabetización en inteligencia artificial permite establecer una base común para las decisiones evaluativas que la universidad necesita adoptar. Docentes y estudiantes pueden participar de manera más consciente cuando comprenden qué hacen los sistemas generativos, qué limitaciones presentan y qué responsabilidades surgen durante su utilización. La evaluación universitaria requiere esta competencia porque difícilmente puede rediseñarse de manera pertinente si quienes participan en ella desconocen las características fundamentales de las herramientas disponibles. El desarrollo de esta alfabetización prepara así el terreno para formular principios evaluativos más claros, realistas y coherentes con las nuevas condiciones tecnológicas (Jin et al., 2025).

1.7 Principios para una evaluación universitaria pertinente en la era de la inteligencia artificial

Una evaluación universitaria pertinente en la era de la inteligencia artificial debe conservar como eje central la demostración efectiva del aprendizaje, aun cuando las condiciones tecnológicas hayan modificado la forma de producir respuestas y productos académicos. Villegas y Carballo (2026) muestran que la expansión de la inteligencia artificial generativa obliga a revisar los criterios utilizados para valorar el desempeño estudiantil. La pertinencia evaluativa depende, por tanto, de que las actividades mantengan correspondencia con los resultados de aprendizaje y permitan distinguir entre la disponibilidad de una herramienta tecnológica y la apropiación real de conocimientos, capacidades y decisiones.

El primer principio corresponde a la validez de la evidencia obtenida. Una evaluación resulta válida cuando permite inferir de manera razonable aquello que pretende valorar y no únicamente la calidad formal del producto presentado. En contextos donde la inteligencia artificial puede generar respuestas estructuradas, la actividad debe exigir evidencias relacionadas con comprensión, aplicación, juicio o razonamiento. Esta exigencia no implica desconfiar automáticamente de los productos académicos, sino diseñarlos de manera que proporcionen información suficiente sobre el aprendizaje alcanzado y sobre la participación intelectual del estudiante durante su construcción (Wang et al., 2024).

UNESCO (2023) plantea que la incorporación de inteligencia artificial en educación debe mantener un enfoque centrado en las personas, lo que traslada a la evaluación la

necesidad de preservar el propósito formativo y el juicio humano. La tecnología puede apoyar determinadas fases del aprendizaje, pero los objetivos educativos no deberían organizarse alrededor de las capacidades de una plataforma. El principio de centralidad humana supone que el estudiante continúe tomando decisiones relevantes y que el docente conserve responsabilidad sobre los criterios utilizados para interpretar el desempeño y establecer las consecuencias académicas derivadas de una evaluación.

La transparencia constituye otro principio necesario porque docentes y estudiantes requieren conocer las condiciones bajo las cuales se desarrolla cada actividad. Esto supone especificar qué se pretende demostrar, qué criterios se utilizarán, qué recursos están permitidos y qué nivel de apoyo tecnológico resulta compatible con el propósito evaluativo. La claridad reduce interpretaciones ambiguas y permite que el uso de inteligencia artificial se relacione con expectativas pedagógicas explícitas. Una evaluación pertinente evita que las reglas aparezcan únicamente después de detectar problemas y procura que las condiciones sean comprensibles desde el inicio del proceso (Villegas y Carballo, 2026).

Tabla 4
Principios orientadores para una evaluación universitaria pertinente en contextos de inteligencia artificial

Principio	Finalidad evaluativa	Pregunta orientadora para el diseño	Riesgo que ayuda a reducir
Validez	Obtener evidencias acordes con el aprendizaje esperado	¿La actividad demuestra realmente lo que se pretende evaluar?	Confundir calidad del producto con aprendizaje
Centralidad humana	Mantener juicio, responsabilidad y participación intelectual	¿Qué decisiones significativas debe realizar el estudiante?	Sustitución del razonamiento por automatización
Transparencia	Clarificar condiciones y expectativas	¿Qué usos tecnológicos y criterios deben conocerse previamente?	Ambigüedad sobre usos permitidos
Equidad	Evitar ventajas derivadas únicamente del acceso tecnológico	¿Todos cuentan con condiciones razonables para demostrar aprendizaje?	Brechas de acceso y experiencia
Explicabilidad	Comprender criterios y resultados de la evaluación	¿Puede justificarse cómo se llegó a una valoración?	Dependencia de decisiones automatizadas opacas

Principio	Finalidad evaluativa	Pregunta orientadora para el diseño	Riesgo que ayuda a reducir
Adaptabilidad	Ajustar prácticas ante cambios tecnológicos y disciplinares	¿La evaluación sigue produciendo evidencias útiles?	Permanencia de formatos que han perdido pertinencia

Nota. Los principios se presentan como criterios articulados para orientar decisiones de diseño, aplicación e interpretación de evaluaciones universitarias; su importancia relativa puede variar según la disciplina, el propósito formativo y las condiciones institucionales.

La equidad adquiere especial relevancia porque la incorporación de inteligencia artificial puede ampliar diferencias existentes entre estudiantes con distintos niveles de acceso, experiencia tecnológica o capacidad para utilizar herramientas generativas. Una evaluación pertinente no debería asumir que todos poseen las mismas condiciones de interacción con estos sistemas. El diseño debe considerar alternativas razonables, formación suficiente y criterios que valoren el aprendizaje más que la sofisticación tecnológica del producto. Esta precaución evita que el rendimiento académico dependa de ventajas externas al propósito educativo y contribuye a preservar oportunidades comparables para demostrar competencias (UNESCO, 2023).

La explicabilidad también debe orientar las decisiones evaluativas cuando intervienen sistemas automatizados. Khosravi et al. (2022) destacan la importancia de comprender cómo la inteligencia artificial produce determinadas recomendaciones o resultados dentro de contextos educativos. Aplicado a la evaluación, este principio implica que una calificación o retroalimentación relevante no debería depender de procesos que docentes y estudiantes no puedan interpretar razonablemente. Si una herramienta tecnológica participa en la valoración, su resultado necesita ser contrastado con criterios pedagógicos explícitos y permanecer sujeto a revisión humana, especialmente cuando existen consecuencias importantes para la trayectoria académica.

Otro principio es la adaptabilidad de las prácticas evaluativas. Las herramientas de inteligencia artificial evolucionan con rapidez y pueden alterar en poco tiempo las condiciones bajo las cuales una actividad conserva su capacidad para demostrar aprendizaje. Wang et al. (2024) evidencian que la inteligencia artificial está produciendo transformaciones amplias dentro de la educación, lo que obliga a evitar diseños considerados permanentes o inmunes al cambio tecnológico. La evaluación necesita revisarse periódicamente a partir de sus resultados, del comportamiento de los estudiantes y de las posibilidades emergentes de automatización, manteniendo flexibilidad sin perder coherencia curricular.

La pertinencia requiere asimismo equilibrio entre prevención y aprovechamiento educativo. Una evaluación diseñada exclusivamente para impedir el uso de inteligencia artificial corre el riesgo de limitar oportunidades legítimas de aprendizaje, mientras que una integración sin criterios puede debilitar la calidad de las evidencias obtenidas. Villegas y Carballo (2026) permiten situar esta tensión como uno de los principales desafíos de la evaluación universitaria contemporánea. Las decisiones deberían establecer cuándo la tecnología puede contribuir al proceso y cuándo determinadas capacidades deben demostrarse sin su intervención, siempre en función de los objetivos de aprendizaje y no de posiciones generales frente a la herramienta.

Estos principios muestran que la transformación evaluativa requiere algo más que sustituir instrumentos o incorporar nuevas plataformas. Validez, centralidad humana, transparencia, equidad, explicabilidad y adaptabilidad conforman un marco para tomar decisiones coherentes ante escenarios tecnológicos cambiantes. Khosravi et al. (2022) refuerzan la necesidad de mantener comprensibles las decisiones mediadas por inteligencia artificial, mientras que UNESCO (2023) sitúa la acción educativa bajo responsabilidad humana. Sobre esta base, la evaluación puede avanzar hacia diseños más auténticos, capaces de exigir desempeños significativos sin ignorar las posibilidades que ofrecen los sistemas generativos.

Capítulo 2.

Evaluación auténtica frente a contenidos generados por inteligencia artificial

Carlos Fernando Castillo Montenegro



2.1 Fundamentos y características de la evaluación auténtica

La evaluación auténtica se fundamenta en la necesidad de valorar aprendizajes que puedan movilizarse más allá de la reproducción de información. Su propósito consiste en situar al estudiante frente a tareas que exijan aplicar conocimientos, interpretar condiciones, tomar decisiones y construir respuestas vinculadas con situaciones significativas. García (2024) plantea que la evaluación auténtica permite aproximar las actividades académicas a contextos donde el aprendizaje adquiere utilidad práctica. Desde esta perspectiva, evaluar implica observar cómo el estudiante integra saberes y criterios para responder ante desafíos que demandan comprensión, juicio y capacidad de actuación.

Una característica central de este enfoque es la contextualización de las tareas. La actividad evaluativa adquiere mayor autenticidad cuando presenta propósitos, destinatarios, restricciones o condiciones que obligan al estudiante a adaptar el conocimiento disponible. No se trata necesariamente de reproducir de manera exacta una situación profesional, sino de construir escenarios que exijan decisiones semejantes a las que podrían surgir fuera de una prueba convencional. Gedera (2026) vincula esta lógica con diseños de evaluación capaces de conservar autenticidad incluso cuando la inteligencia artificial forma parte del entorno académico y de producción.

La evaluación auténtica también se distingue por valorar desempeños complejos en lugar de respuestas aisladas. Una competencia universitaria puede requerir integrar conocimientos disciplinares, habilidades de comunicación, razonamiento y capacidad para justificar decisiones. Por esta razón, una única respuesta correcta no siempre permite representar adecuadamente el nivel de aprendizaje alcanzado. Kofinas et al. (2025) muestran que las evaluaciones auténticas adquieren especial relevancia en contextos donde la inteligencia artificial generativa puede producir contenido, porque desplazan la atención hacia tareas que demandan interpretación, aplicación y participación intelectual situada del estudiante.

Otro fundamento consiste en la relación entre tarea, criterio y evidencia. Una actividad auténtica necesita definir con claridad qué aprendizaje pretende hacer visible y qué características permitirán reconocer un desempeño de calidad. Gu (2021) destaca la importancia de validar la evaluación a partir de los argumentos y evidencias que permiten interpretar el aprendizaje observado. Desde esta lógica, la autenticidad no

depende únicamente de que la tarea parezca realista, sino de que produzca información válida sobre aquello que se quiere valorar y permita establecer juicios suficientemente fundamentados sobre el desempeño del estudiante.

Tabla 5
Características funcionales de una evaluación auténtica en educación superior

Característica	Función dentro de la evaluación	Evidencia esperada	Valor frente a contenidos generados por IA
Contextualización	Sitúa el aprendizaje en condiciones específicas	Adaptación del conocimiento al escenario	Reduce respuestas genéricas
Complejidad	Integra varias capacidades en una misma tarea	Decisiones, procedimientos y productos articulados	Exige participación más allá de la generación automática
Relevancia	Vincula la actividad con propósitos académicos o profesionales	Respuestas útiles y justificadas	Obliga a interpretar condiciones concretas
Criterios explícitos	Define qué se considera desempeño de calidad	Correspondencia entre producto y criterios	Favorece valoración transparente
Juicio del estudiante	Exige seleccionar, comparar y justificar	Argumentación y toma de decisiones	Permite observar agencia intelectual
Aplicación	Moviliza conocimiento en lugar de reproducirlo	Transferencia a situaciones específicas	Dificulta que el producto final sea la única evidencia

Nota. Las características se complementan entre sí y permiten diseñar evaluaciones donde el valor de la evidencia depende de la aplicación, la contextualización y la participación intelectual del estudiante, no únicamente de la calidad formal del producto.

La autenticidad implica además que el estudiante comprenda para qué realiza una actividad y qué función cumple aquello que produce. Un informe dirigido a un destinatario específico, una propuesta que responda a una necesidad concreta o una recomendación basada en restricciones determinadas pueden generar mayor implicación cognitiva que una tarea formulada únicamente para obtener una calificación. García (2024) relaciona la evaluación auténtica con aprendizajes significativos y con una retroalimentación orientada a mejorar el desempeño. El propósito de la actividad se convierte así en parte de la experiencia evaluativa y orienta las decisiones realizadas.

La presencia de inteligencia artificial generativa hace más visible la diferencia entre una tarea auténtica y una actividad basada en producción genérica. Gedera (2026) propone un marco de evaluación auténtica habilitada por inteligencia artificial donde el diseño debe preservar la agencia humana y la relación entre tarea, contexto y aprendizaje. Una actividad puede permitir el uso de herramientas generativas y seguir siendo exigente si obliga al estudiante a seleccionar, verificar, adaptar y justificar los resultados obtenidos. La autenticidad no depende, por tanto, de excluir la tecnología, sino de evitar que esta sustituya el núcleo intelectual del desempeño.

La complejidad constituye otra característica relevante, porque las situaciones auténticas suelen admitir diferentes caminos de resolución. Frente a ejercicios con respuestas previsibles, una tarea compleja puede exigir priorizar información, identificar restricciones y argumentar por qué una alternativa resulta más pertinente que otra. Kofinas et al. (2025) vinculan la evaluación auténtica con la necesidad de preservar la integridad y el valor del aprendizaje en contextos mediados por inteligencia artificial. Cuando existen múltiples posibilidades razonables, el estudiante necesita demostrar cómo llegó a una decisión y qué criterios utilizó para sostenerla.

La evaluación auténtica también requiere criterios capaces de reconocer calidad sin convertir la actividad en una secuencia rígida de pasos. Gu (2021) señala que la interpretación de resultados necesita apoyarse en argumentos sólidos sobre aquello que la evidencia permite afirmar. Esto implica construir criterios relacionados con pertinencia, fundamentación, aplicación, coherencia y calidad de las decisiones. Los criterios orientan tanto al docente como al estudiante, pero deben permitir diferentes formas válidas de resolver una tarea. La autenticidad perdería valor si una situación abierta terminara evaluándose mediante una única respuesta predeterminada.

Otro rasgo importante es la posibilidad de incorporar retroalimentación vinculada con el desempeño. En una tarea auténtica, los comentarios pueden centrarse en cómo se interpretó el problema, qué decisiones fueron adecuadas, qué aspectos requieren mayor fundamentación y cómo podría mejorarse la respuesta ante condiciones similares. García (2024) destaca el papel de la retroalimentación como parte de una evaluación orientada al aprendizaje. Esta función permite que la actividad no termine con una calificación, sino que aporte información útil para fortalecer capacidades que posteriormente podrán transferirse a otros escenarios académicos o profesionales.

La autenticidad también exige equilibrio entre realismo y viabilidad. Una actividad puede aproximarse a situaciones profesionales sin reproducir todas sus condiciones, siempre que conserve los elementos necesarios para movilizar conocimientos y decisiones relevantes. Gedera (2026) plantea que el diseño auténtico debe considerar tanto las posibilidades tecnológicas como las experiencias que permiten demostrar aprendizaje significativo. Esto evita asumir que una evaluación es auténtica únicamente porque utiliza casos reales, simulaciones o herramientas digitales. La autenticidad depende de la relación entre propósito, contexto, actuación requerida y calidad de la evidencia producida.

En contextos donde la inteligencia artificial puede generar contenidos sofisticados, la evaluación auténtica ofrece una base para desplazar el interés desde la producción de respuestas hacia la demostración de capacidades. Kofinas et al. (2025) permiten comprender que el valor de estas evaluaciones radica en exigir decisiones, aplicación y responsabilidad sobre lo producido. El reto no consiste en crear tareas imposibles de resolver con apoyo tecnológico, sino en diseñar actividades donde el aprendizaje siga siendo necesario para interpretar, adaptar y justificar. Sobre estos fundamentos puede analizarse con mayor precisión por qué determinadas tareas tradicionales presentan limitaciones frente a la generación automatizada.

2.2 Limitaciones de las tareas tradicionales frente a la generación automatizada de contenidos

Las tareas tradicionales presentan limitaciones crecientes cuando se apoyan en consignas previsibles, descontextualizadas o centradas en productos que pueden ser generados automáticamente con escasa intervención intelectual. Thanh et al. (2023) muestran que los sistemas de inteligencia artificial generativa pueden resolver actividades auténticas con niveles de desempeño que obligan a reconsiderar la capacidad de ciertas tareas para discriminar aprendizaje. El problema no radica únicamente en que la tecnología produzca respuestas correctas, sino en que algunas actividades permiten obtener resultados aceptables sin exigir comprensión profunda, juicio propio ni adaptación significativa a condiciones específicas.

Una primera limitación aparece en las tareas formuladas de manera excesivamente general. Solicitudes como explicar un concepto, resumir un tema o redactar una reflexión amplia pueden ser respondidas por sistemas generativos con rapidez y

coherencia formal. En estos casos, el producto final ofrece poca información sobre el proceso cognitivo seguido por el estudiante y dificulta establecer qué conocimientos fueron realmente movilizados. La debilidad se encuentra en la estructura de la consigna, porque permite respuestas intercambiables y poco situadas, reduciendo la capacidad de la evaluación para identificar comprensión individual o razonamiento aplicado (Villegas y Carballo, 2026).

Kofinas et al. (2025) advierten que la integridad de las evaluaciones auténticas puede verse afectada cuando las tareas no exigen suficiente participación personal, contextualización o responsabilidad sobre las decisiones tomadas. Esta observación resulta igualmente válida para actividades tradicionales que dependen de la entrega de un único producto. Cuando el docente recibe solo la versión final de un ensayo, informe o resolución, dispone de escasas evidencias para distinguir entre elaboración propia, asistencia tecnológica legítima y generación sustitutiva. La evaluación pierde capacidad interpretativa cuando el producto se convierte en la única base para inferir aprendizaje.

Otra limitación se relaciona con la previsibilidad de las respuestas esperadas. Las tareas construidas alrededor de patrones conocidos, definiciones estandarizadas o preguntas recurrentes pueden ser abordadas con facilidad por herramientas entrenadas sobre grandes volúmenes de información. Pearce y Chiavaroli (2023) sostienen que la inteligencia artificial generativa obliga a replantear formas de evaluación que dependen demasiado de respuestas convencionales. Cuando el estudiante puede obtener una respuesta plausible sin necesidad de analizar restricciones, comparar alternativas o defender una postura, la actividad ofrece menor resistencia a la automatización y menor capacidad para revelar la calidad del razonamiento.

También resultan vulnerables las actividades cuyo criterio principal de calidad se concentra en la corrección formal. Redacción fluida, estructura lógica, ortografía adecuada y presentación ordenada pueden seguir siendo elementos valiosos, pero ya no permiten inferir por sí mismos dominio conceptual. Los sistemas generativos son capaces de producir textos técnicamente correctos incluso cuando el usuario posee un conocimiento limitado del tema. En consecuencia, una evaluación centrada excesivamente en la forma puede sobrevalorar aspectos fácilmente automatizables y subestimar dimensiones como pertinencia, interpretación, fundamentación y capacidad para responder ante condiciones cambiantes (Thanh et al., 2023).

Villegas y Carballo (2026) señalan que la evaluación universitaria necesita adaptarse a escenarios donde el uso de inteligencia artificial modifica las condiciones de producción académica. Esto afecta especialmente a las tareas domiciliarias desarrolladas sin interacción posterior con el docente. La disponibilidad de tiempo, acceso a herramientas y posibilidad de generar múltiples versiones permiten producir trabajos de alta calidad aparente sin que necesariamente exista un proceso equivalente de aprendizaje. La limitación no se resuelve prohibiendo el trabajo fuera del aula, sino aumentando la riqueza de las evidencias solicitadas y la especificidad de las decisiones requeridas.

Otra debilidad aparece cuando las actividades no incorporan información local, datos concretos o condiciones propias de una situación particular. Una consigna demasiado amplia puede ser respondida mediante contenido genérico, mientras que una tarea situada obliga a interpretar elementos que requieren selección y adaptación. Kofinas et al. (2025) destacan la importancia de preservar el valor auténtico de la evaluación frente a la inteligencia artificial, lo que implica diseñar actividades donde las respuestas deban ajustarse a contextos específicos. La contextualización aumenta la exigencia cognitiva porque reduce la utilidad de soluciones universales producidas sin análisis de las circunstancias planteadas.

Las tareas tradicionales también pueden resultar limitadas cuando no exigen justificar decisiones. Un estudiante puede entregar una solución correcta sin explicar por qué eligió determinado procedimiento, qué alternativas consideró o cómo evaluó la calidad de la respuesta obtenida. Pearce y Chiavaroli (2023) plantean que la evaluación debe adaptarse a un entorno donde las herramientas generativas pueden producir resultados convincentes. En este contexto, la justificación adquiere valor porque permite observar el razonamiento subyacente y no únicamente el resultado. La ausencia de explicación reduce la capacidad de la evaluación para diferenciar entre posesión de una respuesta y comprensión de su fundamento.

La uniformidad de las tareas constituye otra limitación relevante. Cuando todos los estudiantes reciben exactamente la misma consigna, con los mismos datos y sin necesidad de personalización, aumenta la posibilidad de utilizar respuestas generadas o estructuras fácilmente replicables. Esto no significa que toda evaluación deba individualizarse por completo, pero sí que conviene incorporar márgenes de elección, variables contextuales o decisiones que obliguen a adaptar el conocimiento. Las

actividades con diferentes rutas de solución permiten observar mejor la participación intelectual del estudiante y disminuyen la dependencia de respuestas estandarizadas producidas automáticamente (Villegas y Carballo, 2026).

Figura 3

Limitaciones de las tareas tradicionales frente a la generación automatizada



Nota. Se identifican las debilidades de las evaluaciones convencionales ante la inteligencia artificial, especialmente en la autenticidad del aprendizaje, la diversidad de respuestas y la valoración del razonamiento crítico.

Thanh et al. (2023) evidencian que los sistemas generativos pueden alcanzar resultados relevantes incluso ante tareas diseñadas con cierta autenticidad, lo que demuestra que la solución no consiste simplemente en hacer las actividades más complejas. Una tarea difícil puede seguir siendo automatizable si su complejidad es predecible o si toda la

información necesaria está contenida en la consigna. El diseño debe incorporar componentes que demanden interpretación, selección, reflexión y responsabilidad sobre las decisiones. La resistencia a la automatización depende menos de aumentar artificialmente la dificultad y más de exigir formas de razonamiento que hagan visible la apropiación del conocimiento.

Otra limitación de algunos formatos tradicionales es su escasa capacidad para identificar cómo se construyó el producto. Cuando la evaluación se concentra exclusivamente en la entrega final, puede ignorar revisiones, cambios de criterio, errores corregidos y decisiones intermedias que aportan información valiosa sobre el aprendizaje. Kofinas et al. (2025) relacionan la autenticidad con la necesidad de preservar evidencias que permitan sostener la integridad del desempeño. Aunque el análisis detallado del proceso corresponde a apartados posteriores, resulta claro que las tareas tradicionales pierden fuerza cuando no ofrecen oportunidades para observar más de una evidencia del trabajo realizado.

Las limitaciones descritas muestran que el problema no reside en categorías como ensayo, informe, examen o tarea domiciliaria, sino en la forma específica en que se diseñan. Pearce y Chiavaroli (2023) permiten comprender que la inteligencia artificial generativa obliga a reconsiderar qué características hacen que una actividad siga siendo válida como evidencia de aprendizaje. Una tarea tradicional puede conservar utilidad si incorpora contexto, decisiones, justificación y criterios claros; por el contrario, puede volverse débil si depende de respuestas genéricas y productos aislados. Esta distinción abre la necesidad de diseñar actividades más vinculadas con problemas y contextos reales.

2.3 Diseño de actividades vinculadas con problemas y contextos reales

El diseño de actividades vinculadas con problemas y contextos reales constituye una de las estrategias más relevantes para fortalecer la autenticidad de la evaluación universitaria. Gedera (2026) plantea que una tarea auténtica debe situar al estudiante ante condiciones que exijan interpretar información, tomar decisiones y producir respuestas ajustadas a un propósito concreto. Esta lógica desplaza las consignas generales hacia situaciones donde el conocimiento necesita ser movilizado de manera contextualizada. La calidad de la evaluación depende entonces de que el problema

planteado obligue a comprender sus particularidades y no pueda resolverse únicamente mediante una respuesta estandarizada.

Un contexto realista aporta información que condiciona la manera en que debe abordarse una actividad. Datos disponibles, características de una población, restricciones presupuestarias, necesidades institucionales o condiciones profesionales pueden modificar las decisiones que el estudiante necesita tomar. Pineda et al. (2024) muestran que los proyectos integradores favorecen el desarrollo de competencias cuando articulan conocimientos con situaciones que demandan aplicación y colaboración. En evaluación, esta contextualización permite observar cómo el estudiante selecciona recursos, establece prioridades y adapta aprendizajes previos frente a circunstancias específicas, evitando respuestas desvinculadas de las exigencias planteadas.

El problema evaluativo debe poseer suficiente apertura para permitir razonamiento sin convertirse en una consigna ambigua. Una tarea auténtica puede admitir varias soluciones plausibles, siempre que existan criterios que permitan valorar la pertinencia de cada alternativa. Bearman et al. (2024) destacan la importancia del juicio evaluativo para reconocer y justificar la calidad de diferentes producciones. Por ello, el diseño debe ofrecer un escenario con información suficiente para orientar al estudiante, pero también con espacios de decisión donde sea necesario comparar opciones, asumir posiciones y explicar las razones que sustentan la respuesta construida.

La vinculación con contextos reales no implica trasladar literalmente una actividad profesional al aula. El diseño universitario puede simplificar determinadas condiciones siempre que conserve los elementos esenciales que activan el conocimiento que se pretende evaluar. Gedera (2026) sostiene que la autenticidad debe responder a la relación entre tarea, contexto y desempeño, más que a una reproducción superficial de la realidad. Una situación ficticia puede ser evaluativamente auténtica si obliga a interpretar variables relevantes, mientras que un caso real puede perder valor si solo solicita describir información sin exigir decisiones ni aplicación significativa.

Tabla 6

Componentes para diseñar actividades evaluativas vinculadas con problemas y contextos reales

Componente	Función en el diseño	Ejemplo de incorporación	Evidencia que permite observar
Problema central	Define el desafío que debe resolverse	Necesidad, conflicto o decisión profesional	Capacidad de análisis y resolución
Contexto específico	Sitúa la actividad en condiciones determinadas	Institución, comunidad, organización o caso	Adaptación del conocimiento
Restricciones	Obligan a priorizar y justificar decisiones	Tiempo, presupuesto, normativa o recursos	Juicio y toma de decisiones
Destinatario	Otorga propósito comunicativo al producto	Cliente, paciente, directivo, comunidad o equipo	Adecuación de la respuesta
Información disponible	Proporciona evidencias para fundamentar la actuación	Datos, documentos, indicadores o antecedentes	Selección e interpretación de información
Criterios de calidad	Orientan el desempeño esperado	Pertinencia, viabilidad, fundamentación y coherencia	Calidad del producto y de las decisiones

Nota. Los componentes pueden combinarse de acuerdo con la disciplina y el nivel formativo; su finalidad es construir tareas donde la respuesta dependa de interpretar condiciones concretas y justificar decisiones, no solamente de producir contenido formalmente correcto.

Una estrategia especialmente útil consiste en incorporar restricciones que obliguen al estudiante a abandonar respuestas ideales o excesivamente generales. Presupuestos limitados, tiempos definidos, requisitos normativos, información incompleta o necesidades contradictorias introducen condiciones semejantes a las que aparecen en escenarios profesionales. Pineda et al. (2024) permiten observar cómo las experiencias integradoras favorecen la articulación de competencias cuando el estudiante necesita responder a problemas complejos. Las restricciones aumentan el valor evaluativo porque obligan a priorizar, justificar renuncias y seleccionar alternativas viables en lugar de limitarse a proponer soluciones conceptualmente correctas.

El destinatario constituye otro elemento que puede fortalecer la autenticidad de una actividad. No es igual redactar un texto “sobre” un problema que elaborar una recomendación para una autoridad, una propuesta para una organización o una explicación dirigida a una comunidad específica. El destinatario obliga a seleccionar lenguaje, nivel de detalle, argumentos y evidencias de acuerdo con una finalidad

concreta. Esta exigencia amplía la participación intelectual del estudiante y dificulta la utilización mecánica de contenidos generados, porque la respuesta necesita ser adaptada a necesidades particulares y criterios de pertinencia definidos previamente (Gedera, 2026).

Thanh et al. (2023) muestran que los sistemas generativos pueden resolver tareas con apariencia auténtica, lo que evidencia que la contextualización por sí sola no garantiza una evaluación resistente a la automatización. El diseño debe exigir que el estudiante trabaje con información específica, interprete evidencias y responda a condiciones que no puedan resolverse mediante formulaciones genéricas. Incorporar datos propios de la asignatura, documentos seleccionados por el docente o antecedentes construidos para el caso puede aumentar la necesidad de análisis. La autenticidad depende de cómo se utiliza el contexto, no simplemente de mencionarlo en la consigna.

También resulta conveniente incorporar momentos donde el estudiante deba explicar cómo llegó a una propuesta. Bearman et al. (2024) vinculan el aprendizaje con la capacidad de juzgar la calidad de diferentes alternativas, aspecto especialmente importante cuando existen herramientas capaces de generar múltiples respuestas plausibles. Solicitar criterios utilizados, opciones descartadas o razones que llevaron a priorizar una estrategia permite observar dimensiones del desempeño que no aparecen en el producto final. Estas exigencias transforman la tarea en una oportunidad para evaluar no solo qué solución fue presentada, sino también la calidad del razonamiento que la sostiene.

Los problemas reales pueden diseñarse con diferentes grados de complejidad según el nivel formativo. En cursos iniciales, la actividad puede ofrecer información estructurada y alternativas relativamente delimitadas; en niveles avanzados, puede incorporar incertidumbre, datos contradictorios o decisiones con consecuencias múltiples. Pineda et al. (2024) evidencian el valor de los proyectos integradores para articular aprendizajes transversales, lo que permite comprender que la autenticidad debe ajustarse progresivamente a las capacidades esperadas. El desafío evaluativo necesita ser suficiente para movilizar conocimientos sin exceder las posibilidades formativas del estudiante ni convertir la dificultad en un fin por sí mismo.

La incorporación de inteligencia artificial dentro de estas actividades puede ser posible cuando su uso no sustituye la interpretación del contexto ni las decisiones esenciales.

Gedera (2026) propone una evaluación auténtica habilitada por inteligencia artificial donde la tecnología puede participar como recurso, siempre que la agencia del estudiante permanezca visible. Una herramienta puede ayudar a explorar alternativas o estructurar información, pero la actividad debe exigir verificar, adaptar y fundamentar cualquier resultado utilizado. De esta manera, la evaluación no se centra en impedir el acceso tecnológico, sino en preservar aquellas acciones intelectuales que permiten demostrar aprendizaje significativo.

El diseño de actividades vinculadas con problemas y contextos reales fortalece la evaluación cuando transforma el conocimiento en una herramienta para actuar frente a condiciones concretas. Thanh et al. (2023) muestran que la capacidad de la inteligencia artificial para producir respuestas obliga a diseñar tareas donde el valor no dependa únicamente del contenido generado. Contexto, restricciones, destinatarios, evidencias y decisiones permiten construir actividades con mayor capacidad para revelar competencias. Esta lógica conduce naturalmente a examinar cómo los productos y desempeños derivados de estas tareas pueden utilizarse como evidencias auténticas del aprendizaje universitario.

2.4 Evaluación de competencias mediante productos y desempeños auténticos

La evaluación de competencias mediante productos y desempeños auténticos permite observar con mayor precisión cómo el estudiante moviliza conocimientos, habilidades y criterios frente a situaciones que demandan actuación. Huauya et al. (2023) señalan que el desarrollo de competencias requiere procesos evaluativos capaces de integrar conocimientos con capacidades de aplicación. Desde esta perspectiva, la competencia no se demuestra únicamente mediante respuestas correctas, sino a través de evidencias donde el estudiante interpreta, decide, produce y justifica. El producto o desempeño se convierte así en una manifestación concreta del aprendizaje alcanzado y de su capacidad para utilizarlo.

Un producto auténtico puede adoptar distintas formas según la disciplina y los resultados de aprendizaje esperados. Informes técnicos, propuestas de intervención, planes, diagnósticos, prototipos, proyectos, estudios aplicados o productos comunicativos pueden funcionar como evidencias cuando obligan al estudiante a integrar conocimientos y tomar decisiones. Méndez et al. (2024) destacan que el enfoque basado en competencias favorece una relación más estrecha entre formación y

desempeño. Por ello, el valor de un producto no depende únicamente de su presentación final, sino de la calidad de las decisiones y fundamentos que sostienen su elaboración.

La autenticidad del producto aumenta cuando responde a un propósito claramente definido y a condiciones semejantes a aquellas en las que el conocimiento podría utilizarse fuera del aula. Gedera (2026) plantea que la evaluación auténtica debe articular tarea, contexto y agencia del estudiante, especialmente en entornos donde la inteligencia artificial puede intervenir en la producción académica. Un producto se vuelve evaluativamente significativo cuando refleja adaptación a restricciones, consideración de destinatarios y aplicación de criterios disciplinares. Estas características permiten observar desempeño real y disminuyen la dependencia de respuestas genéricas o descontextualizadas.

Los desempeños auténticos, a diferencia de los productos estáticos, permiten valorar acciones observables durante la ejecución de una tarea. Un estudiante puede analizar un caso, defender una propuesta, ejecutar un procedimiento, interpretar datos o resolver un problema bajo determinadas condiciones. García (2024) sostiene que la evaluación auténtica favorece una valoración más cercana a los procesos reales de aprendizaje y aplicación. La observación del desempeño aporta información sobre cómo el estudiante organiza sus conocimientos, responde ante dificultades y ajusta sus decisiones, dimensiones que no siempre resultan visibles mediante una entrega escrita aislada.

La evaluación por competencias requiere criterios que permitan distinguir entre una producción formalmente correcta y un desempeño realmente competente. Huauya et al. (2023) vinculan la evaluación formativa con el desarrollo de competencias, lo que implica valorar no solo el resultado, sino también dimensiones relacionadas con aplicación, integración y mejora. Los criterios pueden incluir pertinencia, fundamentación, viabilidad, coherencia, capacidad de adaptación y calidad de las decisiones. Esta diversidad permite reconocer diferentes niveles de logro y evita que la calificación dependa exclusivamente de aspectos superficiales o fácilmente automatizables mediante herramientas tecnológicas.

En contextos donde la inteligencia artificial generativa puede contribuir a redactar, organizar o mejorar productos académicos, resulta especialmente importante que la evaluación exija evidencias de desempeño intelectual. Gedera (2026) propone diseños donde la IA puede formar parte del entorno sin sustituir la agencia del estudiante. Una

propuesta generada con apoyo tecnológico puede seguir siendo válida si el estudiante demuestra que comprende el problema, selecciona criterios, adapta las respuestas y asume responsabilidad sobre las decisiones incorporadas. La autenticidad se preserva cuando el producto refleja juicio y apropiación, no solo una producción técnicamente elaborada.

Méndez et al. (2024) muestran que el enfoque basado en competencias adquiere mayor sentido cuando las actividades permiten conectar conocimientos con situaciones de aplicación. Esto significa que el producto evaluado debe responder a demandas concretas y no limitarse a exhibir información aprendida. Por ejemplo, una propuesta profesional debe mostrar no solo conocimientos conceptuales, sino también capacidad para priorizar, justificar, anticipar consecuencias y ajustar decisiones. La evaluación gana profundidad cuando el producto actúa como evidencia integrada de diferentes capacidades y no como simple contenedor de contenido académico.

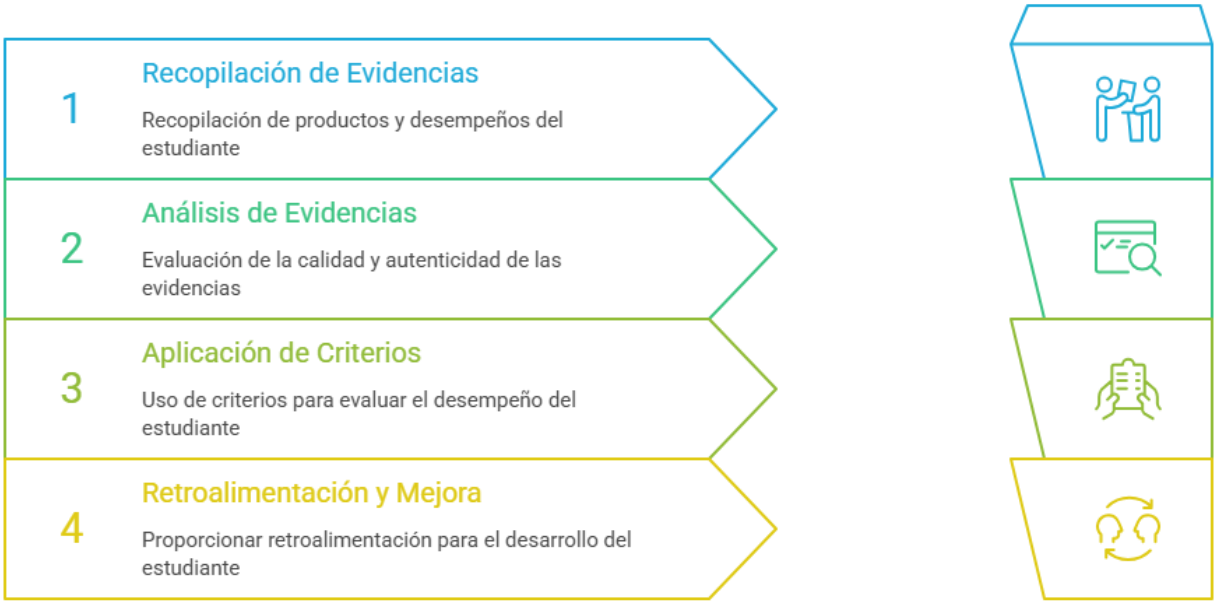
La valoración de desempeños auténticos también permite incorporar variabilidad en las respuestas. Dos estudiantes pueden resolver una misma tarea mediante estrategias diferentes y alcanzar resultados igualmente válidos si sus decisiones están bien fundamentadas. García (2024) resalta que la evaluación auténtica debe valorar aprendizajes significativos y no reducirlos a respuestas únicas. Esta apertura exige criterios suficientemente claros para reconocer calidad sin imponer una sola ruta de solución. El desempeño competente se manifiesta precisamente en la capacidad para seleccionar y defender una alternativa adecuada frente a varias posibilidades razonables.

Otro elemento relevante es la posibilidad de evaluar progresivamente la complejidad del desempeño. En niveles iniciales, los productos pueden requerir aplicación de conocimientos en escenarios relativamente estructurados; en niveles avanzados, pueden incorporar incertidumbre, información contradictoria o restricciones múltiples. Huauya et al. (2023) muestran que el desarrollo de competencias implica una construcción gradual, por lo que las exigencias evaluativas deben aumentar de manera coherente con la formación. Esta progresión permite observar cómo el estudiante pasa de aplicar procedimientos conocidos a enfrentar situaciones donde necesita integrar saberes y tomar decisiones con mayor autonomía.

Los productos auténticos también pueden favorecer una relación más clara entre evaluación y futuro desempeño profesional. Un estudiante que elabora un diagnóstico, diseña una intervención o construye una propuesta enfrenta demandas más cercanas a las que podría encontrar en contextos reales. Méndez et al. (2024) relacionan el enfoque competencial con mejoras en la experiencia formativa cuando las actividades permiten aplicar conocimientos de manera significativa. Esta conexión fortalece la percepción de relevancia de la evaluación y puede aumentar la implicación del estudiante, porque el trabajo deja de sentirse como una tarea exclusivamente académica.

La evaluación de desempeños requiere, además, que el docente observe no solo qué hace el estudiante, sino cómo responde ante cambios o dificultades. Gedera (2026) señala que la autenticidad se fortalece cuando las tareas permiten evidenciar agencia y adaptación. Una respuesta competente puede incluir modificaciones frente a información nueva, reconocimiento de limitaciones o reformulación de decisiones iniciales. Estas acciones aportan evidencias de aprendizaje más ricas que un producto fijo y permiten valorar la flexibilidad con la que el estudiante aplica conocimientos en condiciones variables.

Figura 4
Proceso de evaluación de competencias auténticas



Nota. Se presenta una secuencia de cuatro etapas orientadas a verificar el desempeño estudiantil mediante evidencias, criterios de valoración y retroalimentación para la mejora continua del aprendizaje.

Evaluar competencias mediante productos y desempeños auténticos transforma la evidencia académica en una demostración más completa del aprendizaje. García (2024)

permite entender que la autenticidad se construye cuando la evaluación conecta saberes, aplicación y retroalimentación significativa. En un contexto atravesado por inteligencia artificial generativa, esta lógica adquiere mayor relevancia porque desplaza el énfasis desde la apariencia final del producto hacia la capacidad para actuar, decidir y justificar. Esta perspectiva prepara el terreno para analizar metodologías como resolución de problemas, estudios de caso y aprendizaje basado en proyectos.

2.5 Resolución de problemas, estudios de caso y aprendizaje basado en proyectos

La resolución de problemas, los estudios de caso y el aprendizaje basado en proyectos constituyen estrategias especialmente adecuadas para una evaluación auténtica, porque obligan al estudiante a movilizar conocimientos ante situaciones que demandan análisis y toma de decisiones. Pineda et al. (2024) muestran que los proyectos integradores pueden favorecer el desarrollo de competencias cuando articulan diferentes saberes alrededor de una situación concreta. En estos enfoques, el aprendizaje se evidencia mediante la capacidad para interpretar información, formular alternativas y construir respuestas pertinentes, en lugar de limitarse a reproducir contenidos previamente aprendidos.

La resolución de problemas adquiere valor evaluativo cuando el estudiante debe enfrentarse a situaciones cuya solución no resulta inmediata ni completamente predecible. El problema debe exigir identificar información relevante, establecer relaciones, seleccionar estrategias y justificar el procedimiento utilizado. Huauya et al. (2023) permiten comprender que el desarrollo de competencias requiere oportunidades para aplicar conocimientos y no únicamente recordarlos. Desde esta perspectiva, la calidad del desempeño se observa tanto en la respuesta alcanzada como en la lógica empleada para construirla, especialmente cuando existen diferentes caminos posibles para resolver una misma situación.

Los estudios de caso permiten trasladar esta lógica a escenarios contextualizados donde convergen hechos, actores, restricciones y decisiones. García (2024) señala que la evaluación auténtica debe aproximarse a situaciones que permitan al estudiante aplicar conocimientos de manera significativa. Un caso bien diseñado puede incluir información suficiente para interpretar una problemática, pero también incertidumbres que obliguen a priorizar y argumentar. Su valor no reside únicamente en presentar una

situación realista, sino en exigir que el estudiante relacione conceptos con evidencias concretas y adopte una posición razonada frente al escenario planteado.

El aprendizaje basado en proyectos amplía esta dinámica porque organiza el trabajo alrededor de un producto o solución que se construye durante un periodo determinado. Pineda et al. (2024) destacan que los proyectos integradores pueden articular competencias transversales y favorecer procesos de evaluación entre pares. En este tipo de experiencia, la evaluación puede considerar planificación, toma de decisiones, integración de conocimientos y calidad del producto desarrollado. El proyecto permite observar capacidades que difícilmente aparecen en una actividad breve, como organización, adaptación, coordinación y respuesta ante dificultades surgidas durante el proceso.

Tabla 7

Características evaluativas de la resolución de problemas, los estudios de caso y el aprendizaje basado en proyectos

Estrategia	Núcleo de la actividad	Evidencia principal	Capacidad que permite observar	Riesgo que debe evitarse
Resolución de problemas	Buscar una solución ante una situación desafiante	Procedimiento y respuesta fundamentada	Análisis, selección de estrategias y razonamiento	Reducir el problema a una fórmula repetitiva
Estudio de caso	Interpretar una situación contextualizada	Diagnóstico, decisión o recomendación	Integración de información y juicio contextual	Convertir el caso en una pregunta descriptiva
Aprendizaje basado en proyectos	Construir progresivamente un producto o solución	Proyecto desarrollado y decisiones asociadas	Integración de competencias, planificación y adaptación	Valorar solo el producto final
Combinación de estrategias	Resolver un problema mediante análisis y producción	Solución aplicada y argumentada	Transferencia e integración de saberes	Sobrecargar la actividad sin claridad evaluativa

Nota. Las tres estrategias pueden complementarse dentro de una misma experiencia, siempre que cada componente responda a resultados de aprendizaje definidos y genere evidencias diferenciadas del desempeño estudiantil.

Una característica común de estas metodologías es que permiten introducir información incompleta o cambiante para observar cómo responde el estudiante. En un problema o caso, pueden incorporarse nuevos datos que obliguen a revisar una decisión inicial; en un proyecto, pueden modificarse restricciones o condiciones de implementación. Edgar

et al. (2022) muestran que las experiencias de simulación pueden generar aprendizajes valiosos cuando los estudiantes interactúan con situaciones que exigen respuesta y adaptación. Estas variaciones aportan autenticidad porque permiten evaluar flexibilidad y capacidad para reconsiderar estrategias, en lugar de premiar únicamente la ejecución de un plan previamente fijado.

La inteligencia artificial generativa puede formar parte de estas experiencias sin anular necesariamente su valor evaluativo. Un estudiante podría utilizarla para explorar alternativas, formular posibles soluciones o identificar información preliminar, pero tendría que seleccionar, verificar y adaptar aquello que resulte pertinente. La actividad conserva autenticidad cuando las decisiones esenciales permanecen bajo responsabilidad del estudiante y cuando el producto exige conexión con las condiciones específicas del problema. En este sentido, la evaluación debe centrarse en cómo se utiliza la información obtenida y no únicamente en si existió o no apoyo tecnológico (García, 2024).

Huauya et al. (2023) permiten situar estas estrategias dentro de una lógica formativa en la que las competencias se desarrollan mediante aplicación, retroalimentación y mejora. Un problema puede resolverse inicialmente de manera parcial, un caso puede reinterpretarse después de recibir nueva información y un proyecto puede modificarse a partir de observaciones del docente o de los pares. Estas oportunidades permiten que la evaluación genere aprendizaje mientras produce evidencias sobre el desempeño. La corrección deja de concentrarse exclusivamente en identificar errores y pasa a examinar cómo el estudiante utiliza la retroalimentación para fortalecer sus decisiones.

El diseño de estas metodologías también exige evitar una complejidad artificial que termine dificultando la interpretación de los resultados. Un proyecto excesivamente amplio puede hacer difícil reconocer qué competencia se está evaluando, mientras que un caso sobrecargado de información puede medir capacidad de procesamiento más que comprensión disciplinar. Pineda et al. (2024) muestran que la integración de competencias requiere una organización clara de objetivos y actividades. La autenticidad debe construirse mediante desafíos relevantes y no mediante acumulación indiscriminada de tareas, de modo que cada evidencia obtenida mantenga relación directa con los aprendizajes esperados.

La resolución de problemas, los estudios de caso y el aprendizaje basado en proyectos permiten ampliar la evaluación universitaria hacia escenarios donde el estudiante debe actuar con mayor autonomía y responsabilidad. Edgar et al. (2022) evidencian el valor de experiencias que aproximan el aprendizaje a situaciones de desempeño, mientras que García (2024) destaca la importancia de tareas auténticas para fortalecer la relación entre evaluación y aprendizaje. Estas estrategias ofrecen condiciones adecuadas para integrar posteriormente la inteligencia artificial de manera crítica, no como sustituto del razonamiento, sino como recurso que puede ser examinado, utilizado y cuestionado dentro de actividades evaluativas.

2.6 Uso crítico de la inteligencia artificial dentro de las actividades evaluativas

El uso crítico de la inteligencia artificial dentro de las actividades evaluativas exige superar dos posiciones extremas: la prohibición absoluta y la incorporación indiscriminada. Bearman et al. (2024) sostienen que la presencia de sistemas generativos obliga a fortalecer el juicio evaluativo de los estudiantes, especialmente cuando deben distinguir entre producciones de diferente calidad. Desde esta perspectiva, permitir inteligencia artificial puede ser pedagógicamente válido si la actividad exige analizar, verificar, adaptar y justificar sus resultados. La tecnología se integra entonces como recurso de apoyo, mientras el aprendizaje continúa dependiendo de decisiones intelectuales atribuibles al estudiante.

La incorporación crítica comienza por definir con precisión para qué puede utilizarse la inteligencia artificial dentro de una tarea. Una actividad puede autorizarla para generar alternativas, explorar explicaciones, revisar la claridad de un texto o contrastar posibles soluciones, pero establecer límites cuando se pretende valorar capacidades que deben ser demostradas directamente. UNESCO (2023) plantea que el uso educativo de inteligencia artificial debe mantener propósitos pedagógicos claros y preservar la participación humana. Por ello, la decisión de permitir una herramienta debe responder al aprendizaje esperado y no únicamente a su disponibilidad tecnológica o popularidad entre los estudiantes.

Otro criterio importante consiste en exigir que los resultados generados sean sometidos a evaluación crítica. Utilizar inteligencia artificial no debería significar transferir automáticamente sus respuestas al producto académico. Khosravi et al. (2022) destacan la importancia de la explicabilidad en sistemas de inteligencia artificial aplicados a

educación, lo que refuerza la necesidad de comprender y cuestionar los resultados producidos. En una actividad evaluativa, el estudiante puede ser requerido para identificar supuestos, detectar errores, contrastar información o explicar por qué una determinada recomendación generada resulta pertinente frente a otras posibles alternativas.

La inteligencia artificial también puede integrarse mediante actividades donde el propio resultado generado se convierta en objeto de análisis. El docente puede proporcionar una respuesta producida por un sistema y solicitar al estudiante que valore su precisión, profundidad, coherencia o adecuación a criterios disciplinares. Bearman et al. (2024) relacionan este tipo de trabajo con el desarrollo del juicio sobre la calidad, una capacidad especialmente relevante cuando las herramientas pueden producir contenidos convincentes. En estos casos, la IA no reemplaza la tarea, sino que crea un material sobre el cual el estudiante debe ejercer análisis y valoración.

Tabla 8

Formas de integrar críticamente la inteligencia artificial en actividades evaluativas

Modalidad de uso	Acción del estudiante	Exigencia evaluativa	Evidencia de aprendizaje
Generación de alternativas	Solicita varias propuestas o soluciones	Comparar y seleccionar la más pertinente	Capacidad de juicio y selección
Revisión de respuestas	Analiza contenido generado por IA	Detectar errores, sesgos o insuficiencias	Evaluación crítica
Apoyo a la escritura	Utiliza sugerencias para mejorar un texto	Justificar cambios y conservar coherencia argumentativa	Control intelectual sobre la producción
Exploración inicial	Consulta conceptos, ejemplos o posibilidades	Contrastar y profundizar la información obtenida	Capacidad de investigación y verificación
Simulación de perspectivas	Solicita respuestas desde diferentes enfoques	Comparar argumentos y reconocer limitaciones	Análisis de alternativas
Producción asistida	Integra IA en una parte de la tarea	Adaptar el resultado al contexto y criterios establecidos	Aplicación y responsabilidad sobre el producto

Nota. Las modalidades propuestas requieren que el estudiante conserve responsabilidad sobre las decisiones incorporadas al producto y que el uso de inteligencia artificial responda a objetivos evaluativos previamente definidos.

Galindo-Domínguez et al. (2024) muestran que las valoraciones proporcionadas por inteligencia artificial pueden presentar coincidencias y diferencias respecto de las

realizadas por personas, lo que convierte su uso en una oportunidad para desarrollar análisis comparativo. Una actividad puede solicitar al estudiante contrastar una evaluación generada por IA con criterios previamente definidos, identificar discrepancias y argumentar cuál valoración resulta más adecuada. Esta estrategia permite utilizar la tecnología como instrumento de contraste, no como autoridad definitiva. El aprendizaje se evidencia en la capacidad para interpretar criterios y defender una decisión evaluativa razonada.

El uso crítico también exige reconocer que diferentes niveles de asistencia tecnológica producen efectos distintos sobre la tarea. Solicitar una explicación complementaria no equivale a pedir la elaboración completa de un informe, del mismo modo que recibir una sugerencia de estilo no implica delegar la construcción del argumento. UNESCO (2023) enfatiza la necesidad de orientar el uso responsable de sistemas generativos, lo que en evaluación se traduce en diferenciar funciones permitidas según el propósito de aprendizaje. La claridad sobre estos niveles permite evitar reglas demasiado generales y favorece decisiones más coherentes con cada actividad.

La verificación de información constituye otra competencia fundamental cuando la inteligencia artificial forma parte de la evaluación. Los sistemas pueden producir respuestas plausibles pero incorrectas, incompletas o insuficientemente contextualizadas, por lo que el estudiante necesita comprobar aquello que pretende utilizar. Khosravi et al. (2022) permiten comprender la importancia de hacer comprensibles y examinables los resultados producidos por sistemas automatizados. En una tarea universitaria, verificar no significa únicamente confirmar datos, sino también revisar supuestos, evaluar la pertinencia de las fuentes y determinar si la respuesta se ajusta realmente al problema planteado.

El docente puede además diseñar actividades donde utilizar inteligencia artificial sea parte explícita de la consigna. Por ejemplo, los estudiantes pueden generar una propuesta inicial, identificar sus debilidades y posteriormente reconstruirla mediante criterios disciplinares. Bearman et al. (2024) muestran que desarrollar juicio evaluativo requiere oportunidades para comparar calidad y tomar decisiones fundamentadas. Esta lógica convierte el uso de IA en una experiencia de aprendizaje observable, porque el valor de la actividad se encuentra en aquello que el estudiante modifica, rechaza o mejora y en las razones que utiliza para justificar tales decisiones.

La incorporación crítica de sistemas generativos también requiere que los criterios de evaluación se ajusten a la naturaleza de la actividad. Si la IA está permitida, valorar únicamente corrección formal o riqueza expositiva puede resultar insuficiente, ya que esas dimensiones pueden haber sido ampliamente asistidas. Galindo-Domínguez et al. (2024) evidencian la necesidad de analizar cómo intervienen estas herramientas en los procesos evaluativos. Resulta más pertinente incorporar criterios relacionados con verificación, contextualización, calidad del razonamiento, capacidad de adaptación y justificación de decisiones, de modo que la calificación refleje participación intelectual y no solamente acabado técnico.

El uso crítico de inteligencia artificial dentro de la evaluación se sostiene, por tanto, en una relación de subordinación pedagógica: la herramienta debe responder al propósito de aprendizaje y no determinarlo. UNESCO (2023) sitúa la acción humana como elemento central en la incorporación educativa de sistemas generativos, mientras que Khosravi et al. (2022) refuerzan la importancia de mantener comprensibles sus aportes. Permitir IA puede enriquecer una actividad cuando obliga a analizar, contrastar y decidir; su valor disminuye cuando sustituye precisamente la capacidad que se pretendía evaluar. Desde esta lógica resulta necesario revisar también los criterios con los que se valoran producciones asistidas.

2.7 Criterios y rúbricas para valorar producciones realizadas con apoyo de inteligencia artificial

La valoración de producciones realizadas con apoyo de inteligencia artificial exige criterios capaces de distinguir entre la calidad formal del producto y la calidad del aprendizaje demostrado. Una respuesta puede presentar buena estructura, redacción fluida y apariencia académica sin reflejar necesariamente comprensión suficiente por parte del estudiante. Romero et al. (2025) destacan la utilidad de las rúbricas para organizar dimensiones de desempeño y hacer explícitas las expectativas evaluativas. En este contexto, los criterios deben orientarse hacia aspectos como fundamentación, pertinencia, adaptación, juicio y responsabilidad sobre las decisiones incorporadas al trabajo.

Una rúbrica diseñada para producciones asistidas por inteligencia artificial no debería limitarse a valorar ortografía, organización o presentación. Estas dimensiones continúan siendo relevantes, pero pueden ser ampliamente favorecidas por herramientas

generativas. García (2024) plantea que la evaluación auténtica requiere criterios vinculados con aprendizajes significativos y con la capacidad para aplicar conocimientos. Por ello, resulta necesario incorporar dimensiones que permitan observar la relación entre producto y comprensión, como la calidad de los argumentos, la adecuación al contexto, la selección de evidencias y la coherencia entre decisiones y propósitos académicos.

Galindo-Domínguez et al. (2024) muestran que las valoraciones producidas por inteligencia artificial pueden aproximarse en ciertos aspectos a las realizadas por docentes, aunque no necesariamente coinciden de manera completa. Esta situación evidencia la importancia de establecer criterios explícitos antes de evaluar. Una rúbrica bien construida ayuda a reducir interpretaciones improvisadas y permite centrar la valoración en dimensiones previamente definidas. Cuando la IA interviene en el producto, la claridad de los criterios resulta aún más necesaria porque evita premiar automáticamente atributos formales que pueden haber sido generados sin una participación intelectual equivalente.

Uno de los criterios más importantes es la capacidad de adaptación. El estudiante debe demostrar que el contenido utilizado responde a las condiciones específicas de la tarea y no constituye una reproducción genérica de información generada. Esto puede observarse en la incorporación de datos concretos, la adecuación a destinatarios, la respuesta a restricciones y la conexión con los objetivos planteados. Bearman et al. (2024) destacan la importancia del juicio evaluativo para reconocer la calidad de una producción, lo que implica valorar no solo si una respuesta parece correcta, sino si resulta pertinente frente a una situación determinada.

Otro criterio necesario es la fundamentación. Una producción asistida por inteligencia artificial debe mostrar razones que permitan comprender por qué se seleccionaron determinadas ideas, estrategias o soluciones. No basta con presentar un contenido convincente si el estudiante no puede sostenerlo mediante conceptos disciplinares, evidencias o argumentos consistentes. García (2024) vincula la evaluación auténtica con procesos donde el estudiante demuestra comprensión mediante aplicación significativa. La fundamentación permite diferenciar una respuesta aceptada de manera pasiva de una producción que ha sido examinada, reconstruida y asumida intelectualmente por quien la presenta.

La verificación también merece un lugar explícito dentro de las rúbricas. Gombert et al. (2024) muestran que los sistemas automatizados pueden apoyar procesos de evaluación y retroalimentación, pero la calidad de los resultados depende de cómo se interpretan y utilizan. En una producción asistida por IA, el estudiante debería demostrar que revisó datos, conceptos, referencias o afirmaciones relevantes antes de incorporarlos. Valorar esta capacidad permite reconocer una práctica académica más rigurosa y evita que la corrección aparente de una respuesta sea confundida con confiabilidad. La verificación convierte el uso tecnológico en una actividad crítica.

Las rúbricas también pueden incluir un criterio relacionado con la transformación del material generado. Un estudiante que utiliza inteligencia artificial de manera activa debería ser capaz de modificar, reorganizar, complementar o rechazar partes de las respuestas obtenidas según las necesidades de la tarea. Bearman et al. (2024) permiten vincular esta capacidad con el desarrollo del juicio sobre calidad. La transformación muestra que la tecnología fue utilizada como recurso y no como sustituto completo de la elaboración. Su valoración debe centrarse en la calidad de las decisiones realizadas y no simplemente en la cantidad de cambios introducidos.

Romero et al. (2025) señalan que las rúbricas pueden favorecer la evaluación formativa al ofrecer criterios comprensibles sobre el desempeño esperado. Esta función resulta especialmente útil cuando las actividades permiten el uso de inteligencia artificial, porque los estudiantes necesitan saber desde el inicio qué aspectos serán valorados. La transparencia de los criterios permite orientar el trabajo hacia dimensiones como análisis, argumentación, contextualización y verificación. Además, reduce la posibilidad de que el estudiante interprete que el uso de una herramienta garantiza automáticamente un mejor resultado, ya que la rúbrica deja claro qué evidencia de aprendizaje debe demostrar.

La construcción de niveles de desempeño requiere evitar descriptores vagos como “excelente”, “bueno” o “insuficiente” sin especificar qué caracteriza cada nivel. Gombert et al. (2024) evidencian el valor de una retroalimentación informativa y vinculada con aspectos concretos del desempeño. En una rúbrica, los niveles deberían describir diferencias observables, por ejemplo, entre aceptar contenido generado sin revisión, adaptarlo parcialmente o integrarlo críticamente mediante verificación y fundamentación. Estos descriptores permiten que la calificación refleje progresiones

reales en la calidad del trabajo y facilitan una retroalimentación más útil para el estudiante.

La valoración debe mantener también equilibrio entre producto y proceso, aunque la rúbrica se aplique principalmente a una entrega final. Galindo-Domínguez et al. (2024) muestran que la inteligencia artificial puede participar en distintos momentos de la evaluación, lo que hace necesario interpretar los resultados con cautela. El docente puede considerar evidencias complementarias cuando necesite comprender cómo se construyó una producción, pero sin convertir cada rúbrica en un instrumento de vigilancia. El objetivo es valorar la calidad académica y la participación intelectual observable, no intentar reconstruir exhaustivamente cada interacción tecnológica realizada durante el trabajo.

El diseño de criterios y rúbricas para producciones asistidas por inteligencia artificial debe conservar una lógica pedagógica antes que tecnológica. Romero et al. (2025) permiten comprender que una rúbrica adquiere valor cuando hace explícitos los aprendizajes esperados y orienta la mejora del desempeño. En este escenario, los criterios centrales deben concentrarse en aquello que la universidad desea preservar: comprensión, juicio, fundamentación, contextualización, verificación y responsabilidad sobre lo producido. Esta perspectiva permite evaluar trabajos apoyados por IA sin reducir la valoración a detectar su uso, y cierra el capítulo con una mirada centrada en la calidad del aprendizaje demostrado.



Capítulo 3.

Evidencias de proceso y trazabilidad del aprendizaje

David Israel Báez Herrera

3.1 Del producto final a la valoración del proceso de aprendizaje

La evaluación universitaria centrada exclusivamente en el producto final resulta cada vez más insuficiente para interpretar con precisión cómo se construyó el aprendizaje. Chhetri Thapa y Lewis (2026) plantean que, en contextos donde la inteligencia artificial generativa puede participar en la producción académica, adquiere mayor importancia valorar los procesos mediante los cuales el estudiante desarrolla, revisa y transforma su trabajo. Esta perspectiva desplaza la atención desde la simple entrega de una respuesta terminada hacia las decisiones, estrategias y razonamientos que permiten comprender cómo se alcanzó el resultado presentado.

El producto final continúa siendo una evidencia relevante, pero ya no debería funcionar como única base para establecer el nivel de aprendizaje alcanzado. Un informe, ensayo o proyecto puede mostrar calidad formal y coherencia sin revelar necesariamente las operaciones cognitivas que intervinieron en su elaboración. La valoración del proceso permite ampliar esta mirada al considerar cómo el estudiante organizó información, ajustó decisiones y respondió ante dificultades. Esta ampliación resulta especialmente necesaria cuando existen herramientas capaces de intervenir en diferentes etapas de la producción académica y mejorar considerablemente la apariencia del resultado final (Sharma, 2026).

Banihashem et al. (2022) destacan que los procesos de retroalimentación adquieren mayor valor cuando permiten al estudiante utilizar información para mejorar su desempeño. Esta idea refuerza la importancia de observar el aprendizaje como una trayectoria y no únicamente como un resultado puntual. Cuando el estudiante recibe comentarios, modifica una propuesta o corrige una interpretación, esas acciones ofrecen información sobre su capacidad para comprender y responder a orientaciones. La evaluación del proceso permite reconocer estas transformaciones y evita que la calificación dependa exclusivamente de aquello que aparece en la última versión presentada.

La valoración del proceso también modifica la interpretación del error. En una evaluación centrada únicamente en el producto, los errores suelen identificarse como deficiencias del resultado; en una perspectiva procesual, pueden convertirse en evidencias sobre cómo evoluciona la comprensión. Un planteamiento inicial equivocado puede ser significativo si posteriormente el estudiante reconoce su

limitación, revisa su estrategia y mejora la solución. Zhang y Tur (2024) muestran que los portafolios permiten documentar experiencias de aprendizaje y evolución, lo que evidencia el potencial de conservar diferentes momentos para interpretar mejor el desarrollo académico.

Tabla 9

Diferencias entre la valoración centrada en el producto final y la valoración del proceso de aprendizaje

Dimensión	Enfoque centrado en el producto final	Enfoque centrado en el proceso	Aporte evaluativo
Unidad principal de análisis	Resultado entregado	Trayectoria de construcción	Permite interpretar cómo se alcanzó el resultado
Evidencias consideradas	Producto terminado	Decisiones, ajustes, avances y resultado	Amplía la base para emitir juicios
Tratamiento del error	Deficiencia del producto	Información sobre evolución y aprendizaje	Facilita comprender cambios en el desempeño
Retroalimentación	Predomina después de la entrega	Puede incorporarse durante el desarrollo	Favorece mejora y autorregulación
Papel del estudiante	Presentar una respuesta	Construir, revisar y justificar	Hace visible la participación intelectual
Relación con IA generativa	Dificultad para conocer su intervención	Mayor posibilidad de interpretar el proceso	Reduce dependencia exclusiva de la apariencia final

Nota. Ambos enfoques pueden complementarse; la valoración del proceso no elimina la importancia del producto, sino que incorpora evidencias adicionales para interpretar con mayor precisión el aprendizaje demostrado.

La perspectiva procesual resulta especialmente pertinente ante la inteligencia artificial generativa porque permite observar elementos que no siempre quedan representados en una entrega final. Chhetri Thapa y Lewis (2026) señalan que preservar la autenticidad epistémica requiere prestar atención a cómo se construye el conocimiento y no solamente a su expresión final. Una evaluación que considera el proceso puede examinar decisiones, modificaciones y respuestas frente a nueva información, generando evidencias que ayudan a reconocer la participación intelectual del estudiante incluso cuando determinadas herramientas tecnológicas forman parte del trabajo académico.

Otro aporte consiste en permitir una valoración más justa de trayectorias de aprendizaje diferentes. Dos estudiantes pueden alcanzar resultados similares mediante recorridos

distintos, y esas diferencias pueden ofrecer información importante sobre autonomía, capacidad de revisión o necesidad de apoyo. Banihashem et al. (2022) muestran que la retroalimentación puede contribuir a mejorar el aprendizaje cuando se integra de manera significativa durante el proceso. Observar cómo cada estudiante responde a esa información permite valorar no solo lo que logra al final, sino también la manera en que utiliza oportunidades de mejora para fortalecer su desempeño.

La valoración del proceso no significa registrar cada acción realizada por el estudiante ni convertir la evaluación en un mecanismo de vigilancia permanente. Sharma (2026) sostiene que, en escenarios de evaluación aumentados por inteligencia artificial, es importante hacer visible el juicio y las decisiones que sustentan el trabajo académico. El objetivo consiste en seleccionar evidencias relevantes que permitan comprender la evolución del aprendizaje sin exigir una documentación excesiva. La calidad de las evidencias resulta más importante que su cantidad, especialmente cuando se busca mantener equilibrio entre transparencia, viabilidad y autonomía estudiantil.

Zhang y Tur (2024) muestran que los e-portafolios pueden integrar evidencias diversas y facilitar una visión más amplia del aprendizaje desarrollado a lo largo del tiempo. Esta lógica resulta útil para comprender la diferencia entre almacenar productos y documentar procesos. Una colección de trabajos no necesariamente muestra evolución; para que exista valor procesual, las evidencias deben permitir identificar cambios, conexiones y progresos. La evaluación universitaria necesita, por tanto, organizar los registros de manera que puedan interpretarse como una trayectoria y no como una simple acumulación de documentos producidos durante el curso.

La observación del proceso también puede fortalecer la capacidad del docente para realizar juicios evaluativos más fundamentados. Cuando la calificación depende exclusivamente de un producto final, determinadas decisiones pueden quedar ocultas y resultar difíciles de interpretar. En cambio, disponer de evidencias seleccionadas del desarrollo permite comprender mejor qué dificultades surgieron, cómo fueron resueltas y qué apoyos resultaron necesarios. Banihashem et al. (2022) destacan la relevancia de utilizar información del aprendizaje para mejorar la retroalimentación, lo que refuerza el valor de integrar datos procesuales en las decisiones docentes.

Desde la perspectiva del estudiante, valorar el proceso modifica también la manera en que se concibe la evaluación. El trabajo deja de orientarse únicamente hacia la entrega

y comienza a reconocer la importancia de revisar, reconsiderar y mejorar. Chhetri Thapa y Lewis (2026) relacionan esta orientación con la necesidad de preservar una participación epistémica auténtica en contextos mediados por inteligencia artificial. Cuando el estudiante sabe que sus decisiones y evolución pueden formar parte de la evidencia evaluada, aumenta la relevancia de asumir responsabilidad sobre el camino seguido y no solo sobre el producto final.

Esta perspectiva también permite evitar una oposición simplista entre trabajos "humanos" y trabajos "generados por inteligencia artificial". Sharma (2026) plantea que resulta más productivo hacer visible el juicio ejercido durante la evaluación que concentrarse únicamente en determinar si una herramienta intervino. El aprendizaje puede involucrar diferentes recursos, pero la valoración necesita identificar qué decisiones, interpretaciones y responsabilidades permanecen vinculadas al estudiante. Observar el proceso contribuye a reconocer esas dimensiones y permite construir evaluaciones más informativas que aquellas orientadas exclusivamente a clasificar el origen tecnológico de un producto.

La transición desde el producto final hacia la valoración del proceso amplía, por tanto, la capacidad de la evaluación para interpretar el aprendizaje universitario. Zhang y Tur (2024) muestran el valor de registrar experiencias y producciones a lo largo del tiempo, mientras que Banihashem et al. (2022) evidencian la importancia de integrar retroalimentación dentro de la trayectoria formativa. Esta perspectiva no elimina el resultado final, sino que lo sitúa dentro de una secuencia más amplia de decisiones y transformaciones. A partir de esta lógica resulta necesario precisar qué evidencias progresivas pueden demostrar la construcción del conocimiento.

3.2 Evidencias progresivas para demostrar la construcción del conocimiento

Las evidencias progresivas permiten demostrar que el conocimiento no aparece de manera inmediata en un producto final, sino que se construye mediante decisiones, ajustes y reorganizaciones sucesivas. Chhetri Thapa y Lewis (2026) sostienen que una evaluación orientada al proceso adquiere mayor valor cuando hace visible cómo el estudiante participa en la construcción de sus propias respuestas. Desde esta perspectiva, las evidencias intermedias permiten observar momentos significativos del aprendizaje y reconocer transformaciones que podrían desaparecer si solo se considerara la entrega definitiva.

Una evidencia progresiva puede adoptar formas diversas según la naturaleza de la actividad. Esquemas preliminares, mapas de ideas, formulaciones iniciales de problemas, selecciones de fuentes, hipótesis de trabajo o justificaciones parciales pueden aportar información sobre cómo evoluciona la comprensión. Alfredo et al. (2024) muestran que los datos generados durante el aprendizaje pueden utilizarse para interpretar procesos y apoyar decisiones educativas. Lo relevante no es acumular registros, sino seleccionar aquellos que permitan identificar cambios en el razonamiento, conexiones entre conceptos y decisiones que evidencien construcción gradual del conocimiento.

Cobo Dorado (2021) destaca el valor de la bitácora como recurso para registrar experiencias, decisiones y aprendizajes a lo largo de un proceso académico. Este tipo de evidencia resulta especialmente útil porque permite conservar huellas que de otro modo desaparecerían cuando el estudiante presenta únicamente un producto terminado. Una anotación breve puede mostrar por qué se modificó una estrategia, qué dificultad se identificó o cómo cambió una interpretación inicial. Así, la evaluación dispone de elementos para comprender no solo qué sabe el estudiante, sino cómo llegó a saberlo.

Las evidencias progresivas también ayudan a diferenciar entre acumulación de información y construcción real de conocimiento. Un estudiante puede reunir numerosos datos sin integrarlos de manera significativa, mientras que otro puede mostrar una secuencia clara de selección, relación y reorganización conceptual. Zhang y Tur (2024) muestran que los portafolios permiten documentar trayectorias de aprendizaje cuando las evidencias se organizan con sentido y no como simples archivos. La progresión se vuelve visible cuando cada registro mantiene relación con el anterior y permite reconocer una evolución en la calidad de la comprensión.

Otro aporte de estas evidencias es que permiten observar cambios en el nivel de autonomía. Al inicio de una actividad, el estudiante puede depender de orientaciones detalladas, mientras que posteriormente puede tomar decisiones con mayor independencia. Chhetri Thapa y Lewis (2026) vinculan esta evolución con la preservación de una participación auténtica en el aprendizaje. Registrar ciertos momentos del proceso permite identificar cómo se transforma la capacidad para seleccionar información, justificar alternativas y responder ante nuevas exigencias. La autonomía deja entonces de ser una impresión general y comienza a manifestarse mediante evidencias concretas.

La selección de evidencias progresivas debe responder a criterios de utilidad pedagógica. No toda acción necesita ser registrada ni toda fase requiere producir un documento adicional. Alfredo et al. (2024) permiten comprender que el valor de la información educativa depende de su capacidad para apoyar interpretaciones significativas sobre el aprendizaje. Una evidencia es pertinente cuando ayuda a responder preguntas concretas sobre evolución, comprensión o toma de decisiones. Este criterio evita transformar la evaluación del proceso en una carga administrativa y mantiene la atención sobre momentos realmente informativos del desarrollo académico.

Figura 5

Evidencias progresivas para demostrar la construcción del conocimiento



Nota. Se presentan diferentes evidencias que permiten reconocer la evolución del pensamiento, documentar las decisiones y valorar el desarrollo gradual del aprendizaje estudiantil.

La secuencia de evidencias también puede favorecer la retroalimentación oportuna. Cuando el docente dispone de información antes de la entrega final, puede identificar dificultades que todavía pueden ser corregidas y orientar al estudiante hacia decisiones más consistentes. Cobo Dorado (2021) muestra que el registro sistemático del proceso puede contribuir a reflexionar sobre aquello que se está haciendo y sobre los cambios

necesarios. La evidencia progresiva no cumple únicamente una función de comprobación posterior, sino que puede convertirse en un recurso para intervenir pedagógicamente mientras el aprendizaje aún está en construcción.

La inteligencia artificial generativa refuerza la importancia de estas evidencias porque puede intervenir en distintas fases de una actividad académica. Una respuesta final puede contener aportes tecnológicos difíciles de distinguir a simple vista, mientras que una secuencia de decisiones y revisiones permite comprender mejor cómo participó el estudiante. Chhetri Thapa y Lewis (2026) plantean que la autenticidad del aprendizaje se fortalece cuando se hacen visibles los procesos de construcción. Las evidencias progresivas no buscan demostrar ausencia de tecnología, sino mostrar continuidad, criterio y responsabilidad intelectual a lo largo del trabajo realizado.

Zhang y Tur (2024) permiten observar que el valor de una colección de evidencias aumenta cuando existe organización temporal y reflexión sobre su significado. Esto implica que la progresión no debe limitarse a presentar documentos en orden cronológico, sino mostrar relaciones entre ellos. Un esquema inicial, una reformulación posterior y una decisión final adquieren mayor valor cuando permiten explicar qué cambió y por qué. La evaluación puede entonces reconocer evolución conceptual, capacidad de revisión y coherencia entre decisiones sucesivas, en lugar de limitarse a verificar la existencia de múltiples entregas.

Las evidencias progresivas amplían la capacidad de la evaluación para representar el aprendizaje como un proceso dinámico y no como un resultado instantáneo. Alfredo et al. (2024) muestran el potencial de utilizar información generada durante la experiencia educativa para comprender mejor cómo evoluciona el estudiante. Seleccionadas con criterio, estas evidencias permiten observar comprensión, autonomía, cambios de estrategia y respuesta a la retroalimentación. Su organización sistemática conduce naturalmente hacia instrumentos capaces de conservar y estructurar esas huellas, como los portafolios y las bitácoras académicas.

3.3 Portafolios y bitácoras como registros del proceso académico

Los portafolios y las bitácoras constituyen recursos especialmente valiosos para documentar el proceso académico porque permiten organizar evidencias que muestran cómo evoluciona el aprendizaje a lo largo del tiempo. Zhang y Tur (2024) destacan que los portafolios pueden integrar producciones, reflexiones y registros diversos dentro de

una estructura que favorece la comprensión de la trayectoria seguida por el estudiante. Su valor no reside únicamente en reunir documentos, sino en mostrar relaciones entre ellos, identificar cambios y hacer visible la manera en que las decisiones académicas se transforman durante el desarrollo de una actividad.

El portafolio adquiere sentido evaluativo cuando existe una selección intencional de evidencias vinculadas con los resultados de aprendizaje. No se trata de conservar todo aquello que el estudiante produce, sino de elegir materiales capaces de mostrar avances, dificultades, revisiones y logros relevantes. Alfredo et al. (2024) señalan que la información generada durante el aprendizaje puede aportar valor cuando se organiza con criterios que permitan interpretarla. Desde esta perspectiva, el portafolio funciona como una estructura de evidencia que permite al docente observar evolución y al estudiante reconocer con mayor claridad su propio recorrido formativo.

Cobo Dorado (2021) plantea que la bitácora puede utilizarse como estrategia para registrar experiencias, decisiones y reflexiones derivadas del proceso educativo. A diferencia de un portafolio, que suele reunir productos seleccionados, la bitácora se concentra con mayor frecuencia en documentar lo que ocurre durante el desarrollo de una actividad. Puede incluir observaciones sobre dificultades, razones para modificar una estrategia, aprendizajes inesperados o decisiones que marcaron un cambio en el trabajo. Este carácter narrativo permite conservar información que normalmente desaparecería cuando solo se presenta una versión final.

La principal diferencia entre ambos instrumentos se encuentra en la forma en que representan el proceso. El portafolio organiza evidencias significativas y permite observar una trayectoria mediante productos seleccionados, mientras que la bitácora registra de manera más continua las experiencias y decisiones que acompañan esa trayectoria. Zhang y Tur (2024) muestran que el valor del portafolio aumenta cuando existe reflexión sobre las evidencias incorporadas. Por ello, ambos recursos pueden complementarse: uno permite observar qué produjo el estudiante en distintos momentos y el otro ayuda a comprender por qué tomó determinadas decisiones durante ese recorrido.

Los portafolios también favorecen la integración de evidencias heterogéneas. Una misma experiencia formativa puede incluir informes, esquemas, análisis de casos, presentaciones, productos audiovisuales, registros de retroalimentación y reflexiones

personales. Alfredo et al. (2024) permiten comprender que la diversidad de datos educativos puede ofrecer una visión más rica del aprendizaje cuando existe una estructura que facilite su interpretación. El portafolio permite reunir estas evidencias sin reducir la evaluación a un único formato, haciendo posible observar distintas dimensiones del desempeño y relacionarlas con los objetivos planteados en una asignatura o proyecto.

La bitácora, por su parte, permite capturar momentos de incertidumbre y cambio que suelen quedar fuera de los productos formales. Cobo Dorado (2021) destaca su utilidad para promover reflexión sobre la experiencia y para registrar el desarrollo de procesos de aprendizaje. Un estudiante puede documentar por qué descartó una alternativa, qué dificultad encontró al aplicar un concepto o qué cambio realizó después de recibir retroalimentación. Estas anotaciones proporcionan información sobre la forma en que interpreta su propio trabajo y sobre la capacidad para reconocer decisiones que influyen en el resultado alcanzado.

Banihashem et al. (2022) muestran que la retroalimentación adquiere mayor efectividad cuando se integra dentro del proceso y puede ser utilizada para mejorar el desempeño. Los portafolios y bitácoras ofrecen condiciones favorables para documentar esta relación porque permiten conservar tanto la observación recibida como la respuesta del estudiante frente a ella. Una corrección aislada evidencia que existió retroalimentación, pero la comparación entre una versión inicial y otra posterior permite observar cómo fue interpretada. Esta continuidad convierte los registros en evidencia de aprendizaje y no solamente en archivos administrativos del curso.

En contextos donde la inteligencia artificial generativa puede intervenir durante la producción académica, estos instrumentos adquieren una función adicional. Un portafolio puede conservar versiones, decisiones y productos seleccionados, mientras que una bitácora puede registrar momentos en que se utilizó una herramienta, razones para recurrir a ella y modificaciones realizadas posteriormente. Alfredo et al. (2024) muestran la importancia de utilizar información del proceso para comprender mejor el aprendizaje. Estos registros no buscan vigilar cada interacción tecnológica, sino proporcionar contexto suficiente para interpretar la participación intelectual del estudiante dentro de un trabajo mediado por diferentes recursos.

El diseño de estos instrumentos debe evitar que la documentación se convierta en una carga excesiva. Si se exige registrar cada acción, el estudiante puede terminar produciendo evidencias mecánicas que aporten poca información y aumenten innecesariamente el trabajo. Zhang y Tur (2024) permiten comprender que el valor del portafolio depende de una selección significativa y de la reflexión asociada con las evidencias. De manera similar, una bitácora resulta más útil cuando registra momentos relevantes del proceso que cuando se limita a completar entradas rutinarias sin relación clara con el aprendizaje o las decisiones realizadas.

La calidad de los registros también depende de las preguntas que orientan la reflexión. Cobo Dorado (2021) muestra que la bitácora puede estimular procesos de interpretación sobre aquello que se está aprendiendo y haciendo. Preguntas como qué cambió, qué dificultad fue más relevante, qué decisión resultó necesaria o qué evidencia modificó una idea inicial pueden producir información más significativa que solicitudes generales de descripción. Estas orientaciones ayudan al estudiante a identificar aspectos importantes de su trayectoria y facilitan al docente interpretar el sentido de las evidencias registradas dentro del proceso académico.

Los portafolios pueden además favorecer una evaluación más longitudinal, especialmente cuando se utilizan durante periodos amplios de formación. Banihashem et al. (2022) destacan la importancia de disponer de información que permita orientar procesos de mejora, y una colección organizada de evidencias puede mostrar patrones que no serían visibles en actividades aisladas. La evolución de la argumentación, la capacidad para integrar retroalimentación o la progresiva autonomía en la toma de decisiones pueden observarse con mayor claridad cuando existen registros comparables. Esta continuidad amplía la capacidad de valorar desarrollo y no solo resultados momentáneos.

Portafolios y bitácoras fortalecen la evaluación del proceso cuando se utilizan como instrumentos para organizar, interpretar y reflexionar sobre evidencias significativas. Zhang y Tur (2024) muestran el potencial de los portafolios para documentar aprendizajes a lo largo del tiempo, mientras que Cobo Dorado (2021) destaca la capacidad de la bitácora para registrar experiencias y decisiones. Su valor conjunto reside en hacer visible una trayectoria que el producto final no puede representar completamente. Esta lógica prepara el análisis de borradores, revisiones y versiones sucesivas como evidencias específicas de transformación académica.

3.4 Borradores, revisiones y versiones sucesivas de los productos evaluados

Los borradores, las revisiones y las versiones sucesivas permiten observar de manera directa cómo evoluciona un producto académico antes de alcanzar su forma final. Ossa y Willatt (2023) muestran que la retroalimentación apoyada por inteligencia artificial puede participar en procesos de escritura académica y favorecer instancias de revisión. Desde una perspectiva evaluativa, conservar diferentes versiones permite identificar cambios en argumentos, estructura, precisión conceptual y uso de evidencias. La atención deja de centrarse únicamente en el documento terminado y se amplía hacia las decisiones que explican cómo el estudiante transformó progresivamente su trabajo.

Un borrador inicial no debe interpretarse como una versión defectuosa que simplemente necesita corrección, sino como una evidencia del estado de comprensión en un momento determinado. Puede mostrar ideas todavía poco articuladas, decisiones provisionales o dificultades que posteriormente serán reconsideradas. Chhetri Thapa y Lewis (2026) sostienen que la evaluación orientada al proceso permite preservar la autenticidad del aprendizaje al hacer visibles las operaciones intelectuales que conducen al resultado. Comparar el borrador con versiones posteriores permite reconocer cambios relevantes y comprender qué aspectos del pensamiento académico fueron reorganizados durante el desarrollo de la tarea.

La revisión adquiere verdadero valor formativo cuando implica algo más que corregir errores superficiales. Modificar una palabra, mejorar la presentación o ajustar aspectos ortográficos puede perfeccionar el producto, pero no necesariamente demuestra transformación conceptual. Gombert et al. (2024) muestran que la retroalimentación informativa puede orientar mejoras específicas en producciones académicas cuando ofrece información útil sobre el desempeño. Desde esta lógica, una revisión significativa debería evidenciar reorganización de argumentos, incorporación crítica de nueva información, reformulación de interpretaciones o cambios en decisiones que permitan identificar un aprendizaje más profundo.

Las versiones sucesivas también permiten analizar cómo el estudiante utiliza la retroalimentación recibida. Una observación docente, una evaluación entre pares o una sugerencia generada por inteligencia artificial adquieren valor educativo cuando el estudiante puede interpretarlas y decidir cómo incorporarlas. Ossa y Willatt (2023) evidencian que la IA generativa puede funcionar como recurso de retroalimentación en

procesos de escritura, pero su utilidad depende de la capacidad para valorar las sugerencias. La evaluación puede examinar no solo qué cambió entre versiones, sino también si las modificaciones muestran comprensión, criterio y apropiación de las orientaciones recibidas.

Tabla 10

Lectura evaluativa de los cambios entre borradores y versiones sucesivas

Tipo de cambio observado	Ejemplo de modificación	Evidencia que puede aportar	Interpretación evaluativa
Conceptual	Reformulación de una explicación o categoría	Mayor precisión y comprensión	Evolución del dominio disciplinar
Argumentativo	Incorporación o reorganización de razones	Mejor coherencia entre afirmaciones y evidencias	Desarrollo del razonamiento
Metodológico	Cambio de procedimiento o estrategia	Reconocimiento de limitaciones previas	Capacidad de revisión de decisiones
Contextual	Adaptación del contenido a nuevas condiciones	Mayor pertinencia de la respuesta	Transferencia y ajuste del conocimiento
Crítico	Rechazo o modificación de una sugerencia recibida	Juicio sobre la calidad de la retroalimentación	Autonomía intelectual
Formal	Mejora de redacción, estructura o presentación	Mayor claridad comunicativa	Perfeccionamiento del producto, sin implicar por sí solo aprendizaje conceptual

Nota. Los cambios entre versiones deben interpretarse según su profundidad y relación con los resultados de aprendizaje; una mayor cantidad de modificaciones no implica necesariamente una mejora académica más significativa.

Chhetri Thapa y Lewis (2026) destacan que la visibilidad del proceso resulta especialmente importante cuando los sistemas generativos pueden contribuir a producir textos de elevada calidad formal. Una secuencia de versiones puede mostrar qué ideas existían inicialmente, cuáles fueron incorporadas posteriormente y cómo se transformó el razonamiento. Sin embargo, el objetivo no debe ser reconstruir exhaustivamente cada interacción tecnológica. Lo importante es disponer de suficientes evidencias para interpretar la continuidad del trabajo y reconocer si las modificaciones muestran apropiación intelectual, adaptación y capacidad para responder de manera razonada a nueva información.

La comparación entre versiones puede convertirse además en una oportunidad para desarrollar juicio evaluativo. Bearman et al. (2024) señalan que los estudiantes necesitan aprender a reconocer la calidad de diferentes producciones y tomar decisiones fundamentadas sobre cómo mejorarlas. Solicitar que expliquen cuáles fueron los cambios más relevantes, por qué los realizaron y qué versión consideran superior permite transformar la revisión en una actividad de evaluación del propio trabajo. Esta práctica ayuda a que el estudiante no perciba la edición como una corrección mecánica, sino como un proceso de valoración crítica de sus decisiones académicas.

La incorporación de inteligencia artificial exige prestar especial atención a revisiones que pueden realizarse de manera casi automática. Una herramienta puede reescribir párrafos, reorganizar estructuras o mejorar el estilo sin que el estudiante comprenda plenamente las razones del cambio. Gombert et al. (2024) evidencian el potencial de los sistemas automatizados para ofrecer retroalimentación detallada, pero también la necesidad de interpretar adecuadamente sus aportes. Por ello, la evaluación debería valorar la capacidad para seleccionar, rechazar o transformar sugerencias, en lugar de asumir que toda mejora visible entre versiones representa necesariamente aprendizaje.

También conviene establecer un número razonable de versiones relevantes. Solicitar múltiples entregas sin una función pedagógica clara puede convertir el proceso en una rutina documental y aumentar la carga de trabajo sin producir mejores evidencias. Ossa y Willatt (2023) permiten comprender que la retroalimentación resulta útil cuando interviene en momentos donde todavía es posible modificar el producto. Una secuencia breve, como versión inicial, revisión intermedia y entrega final, puede ser suficiente si cada etapa responde a propósitos definidos y permite observar transformaciones vinculadas con los criterios de evaluación establecidos.

Las versiones sucesivas adquieren mayor valor cuando el estudiante puede explicar la relación entre ellas. No basta con conservar documentos diferentes; es necesario identificar qué cambió, qué motivó la modificación y qué aprendizaje se deriva de esa transformación. Bearman et al. (2024) vinculan el juicio evaluativo con la capacidad para reconocer calidad y tomar decisiones sobre el propio desempeño. Esta explicación permite distinguir entre modificaciones superficiales y cambios que evidencian comprensión más profunda, además de ofrecer al docente una base más sólida para interpretar el progreso observado durante la actividad.

Borradores, revisiones y versiones sucesivas convierten la evolución del producto académico en una fuente de evidencia evaluativa. Chhetri Thapa y Lewis (2026) permiten situar esta práctica dentro de una evaluación orientada a preservar la autenticidad del proceso, mientras que Gombert et al. (2024) muestran el valor de una retroalimentación capaz de producir mejoras informadas. El interés principal no reside en controlar cada modificación, sino en reconocer cómo cambia la comprensión y qué decisiones explican esa evolución. Esta lógica conduce hacia la necesidad de documentar específicamente cuándo y cómo interviene la inteligencia artificial durante el aprendizaje.

3.5 Registro y declaración del uso de inteligencia artificial durante el aprendizaje

El registro del uso de inteligencia artificial durante el aprendizaje permite documentar de manera más precisa cómo intervienen estas herramientas en la construcción de un producto académico. Sharma (2026) sostiene que, en contextos evaluativos aumentados por inteligencia artificial, resulta necesario hacer visible el juicio ejercido por el estudiante. Registrar el uso tecnológico no significa controlar cada interacción, sino identificar aquellas acciones que influyen de manera significativa en el desarrollo del trabajo. Esta documentación aporta contexto para interpretar mejor las decisiones, modificaciones y aportes intelectuales que acompañan la producción académica.

Una declaración útil debe diferenciar entre tipos de intervención. No es equivalente utilizar inteligencia artificial para generar ideas iniciales, revisar estilo, sintetizar información, proponer alternativas o redactar partes sustantivas de una entrega. UNESCO (2023) plantea la necesidad de promover un uso responsable y transparente de los sistemas generativos en educación. Desde esta perspectiva, el registro debería permitir reconocer para qué se utilizó la herramienta, en qué momento intervino y qué función cumplió dentro del proceso. Esta diferenciación evita que toda utilización de IA sea interpretada como una práctica homogénea.

Chhetri Thapa y Lewis (2026) destacan la importancia de preservar la autenticidad del proceso cuando la inteligencia artificial participa en actividades académicas. Una forma de hacerlo consiste en solicitar al estudiante que describa brevemente qué aportes tecnológicos incorporó y qué decisiones tomó después de recibirlos. Por ejemplo, puede indicar si aceptó, rechazó o modificó una sugerencia generada. Este tipo de información ayuda a mostrar continuidad entre el uso de la herramienta y la elaboración personal,

sin exigir transcripciones exhaustivas de cada interacción mantenida durante el desarrollo del trabajo.

El registro también puede incluir evidencias seleccionadas de interacción cuando estas resultan relevantes para comprender una decisión. Capturas, fragmentos de prompts, versiones comparadas o descripciones breves pueden servir para mostrar cómo se utilizó la herramienta. Sharma (2026) permite situar estas evidencias como parte de una evaluación donde el juicio necesita hacerse visible. Sin embargo, la finalidad no debería ser acumular pruebas de uso, sino identificar momentos significativos que permitan interpretar mejor el razonamiento del estudiante y su capacidad para intervenir críticamente sobre los resultados generados.

Tan y Maravilla (2024) señalan que la inteligencia artificial generativa no necesariamente debilita la educación si su utilización se integra dentro de marcos que mantengan responsabilidad y participación humana. Esta idea resulta central para comprender la función del registro. Documentar el uso de IA no debe concebirse únicamente como mecanismo de sospecha, sino como una práctica que ayuda a reconocer diferentes formas de apoyo tecnológico. Una declaración clara puede mostrar que la herramienta fue utilizada como recurso complementario y que las decisiones académicas relevantes permanecieron bajo responsabilidad del estudiante.

La calidad del registro depende también de su nivel de detalle. Una declaración excesivamente general, como indicar simplemente que "se utilizó inteligencia artificial", aporta poca información para interpretar el proceso. UNESCO (2023) permite comprender la importancia de establecer orientaciones claras sobre el uso de sistemas generativos. En consecuencia, resulta más útil especificar la herramienta utilizada, el propósito de su empleo, las partes del proceso en las que intervino y el tratamiento dado a sus resultados. Esta información ofrece una base más precisa para comprender la participación tecnológica sin transformar la documentación en una carga desproporcionada.

El momento en que se solicita la declaración también influye en su utilidad. Si el estudiante intenta reconstruir todo el proceso únicamente al finalizar la actividad, puede omitir decisiones relevantes o reducir el registro a una descripción superficial. Chhetri Thapa y Lewis (2026) muestran el valor de una evaluación orientada al proceso, donde las evidencias se generan a medida que el aprendizaje avanza. Por ello, pueden utilizarse

registros breves durante etapas específicas del trabajo, permitiendo que la documentación conserve mayor proximidad con las decisiones realmente tomadas en cada momento.

Otra función del registro consiste en favorecer la reflexión sobre el propio uso de inteligencia artificial. Sharma (2026) plantea que hacer visible el juicio permite valorar con mayor claridad la participación del estudiante. Solicitar que explique por qué utilizó una herramienta, qué obtuvo de ella y qué modificó posteriormente puede convertir la declaración en una actividad metacognitiva. El estudiante no solo informa que empleó IA, sino que analiza cómo esa intervención afectó su trabajo. Esta reflexión ayuda a diferenciar entre dependencia automática y utilización consciente como recurso dentro del proceso académico.

Figura 6

Componentes de una declaración de uso de inteligencia artificial



Nota. Se identifican los elementos necesarios para transparentar el uso de herramientas de inteligencia artificial, documentar su intervención y promover la responsabilidad académica.

Tan y Maravilla (2024) permiten comprender que la responsabilidad académica no desaparece cuando una herramienta participa en la producción. El estudiante continúa siendo responsable de revisar la pertinencia, exactitud y adecuación de aquello que incorpora. Por esta razón, el registro puede incluir una breve explicación sobre los mecanismos de verificación utilizados. Señalar que una respuesta fue contrastada, modificada o descartada permite mostrar que el estudiante no actuó como receptor

pasivo. La documentación adquiere así valor cuando evidencia control intelectual y no cuando se limita a enumerar herramientas utilizadas.

El diseño institucional o docente del registro debe mantener proporcionalidad. No todas las actividades requieren el mismo nivel de documentación, y una tarea breve no necesita un sistema complejo de trazabilidad. UNESCO (2023) plantea la importancia de integrar la inteligencia artificial de acuerdo con el propósito educativo, principio que también puede aplicarse a la declaración de uso. El nivel de detalle debería depender de la relevancia que tenga la herramienta dentro de la actividad, del tipo de producto solicitado y del aprendizaje que se pretende demostrar, evitando procedimientos innecesariamente burocráticos.

El registro y la declaración del uso de inteligencia artificial fortalecen la evaluación cuando permiten comprender cómo se relacionan tecnología, decisiones y aprendizaje. Chhetri Thapa y Lewis (2026) sitúan esta necesidad dentro de una lógica de autenticidad procesual, mientras que Sharma (2026) enfatiza la importancia de hacer visible el juicio del estudiante. Documentar de manera selectiva, clara y proporcional puede aportar información relevante sobre la participación intelectual sin convertir la evaluación en vigilancia permanente. Esta práctica prepara el análisis de la retroalimentación y la metacognición como evidencias adicionales de comprensión.

3.6 Retroalimentación y metacognición como evidencias de comprensión

La retroalimentación adquiere valor evaluativo cuando permite observar cómo el estudiante interpreta información sobre su desempeño y la utiliza para modificar decisiones, estrategias o productos. Banihashem et al. (2022) destacan que las prácticas de retroalimentación en educación superior resultan más efectivas cuando se integran dentro del proceso de aprendizaje y favorecen acciones posteriores. Desde esta perspectiva, la evidencia no se encuentra únicamente en el comentario recibido, sino en la respuesta del estudiante frente a él. Revisar, justificar cambios o mantener una decisión con argumentos puede mostrar comprensión y capacidad para autorregular el propio aprendizaje.

La metacognición complementa esta perspectiva porque permite al estudiante analizar qué sabe, cómo está aprendiendo y qué dificultades necesita resolver. Ossa y Willatt (2023) muestran que la retroalimentación puede apoyar procesos de escritura académica cuando promueve revisión y reflexión sobre las decisiones tomadas. Una respuesta

metacognitiva puede hacer visible por qué se modificó un argumento, qué criterio permitió reconocer un error o qué estrategia resultó más efectiva. Estas explicaciones ofrecen información que no aparece necesariamente en el producto final y ayudan a interpretar con mayor precisión la comprensión alcanzada durante el proceso.

de Schipper et al. (2021) evidencian que la retroalimentación automatizada puede personalizarse y proporcionar información específica sobre el desempeño. Sin embargo, recibir recomendaciones no garantiza que exista aprendizaje. El valor evaluativo aparece cuando el estudiante interpreta la información, determina su pertinencia y decide cómo utilizarla. Esta distinción resulta especialmente importante en contextos donde sistemas automatizados pueden generar comentarios con rapidez. La comprensión puede demostrarse cuando el estudiante explica por qué acepta una sugerencia, por qué la modifica o por qué considera que no resulta adecuada para los objetivos de la actividad.

La metacognición también permite reconocer diferencias entre corregir un producto y comprender la razón de la corrección. Gombert et al. (2024) muestran que la retroalimentación informativa puede orientar mejoras cuando ofrece indicaciones relacionadas con aspectos concretos del desempeño. Un estudiante puede realizar una modificación superficial siguiendo una recomendación sin comprender plenamente su fundamento, mientras que otro puede utilizar la misma observación para reorganizar su razonamiento. La evaluación necesita distinguir estas situaciones mediante evidencias que permitan explicar qué se aprendió, qué cambió en la comprensión y cómo esa nueva interpretación influyó en la versión posterior.

Tabla 11

Indicadores de comprensión derivados de la retroalimentación y la metacognición

Evidencia observada	Acción del estudiante	Indicador de comprensión	Utilidad evaluativa
Interpretación de comentarios	Explica el significado de la retroalimentación recibida	Comprensión del criterio evaluado	Permite valorar apropiación de la observación
Revisión fundamentada	Modifica el producto y justifica el cambio	Capacidad para aplicar retroalimentación	Hace visible aprendizaje entre versiones
Rechazo argumentado	Mantiene una decisión y explica sus razones	Juicio crítico y autonomía	Evita asumir que toda sugerencia debe aceptarse
Identificación de dificultades	Reconoce errores o vacíos propios	Conciencia sobre el desempeño	Facilita autorregulación

Evidencia observada	Acción del estudiante	Indicador de comprensión	Utilidad evaluativa
Selección de estrategias	Describe cómo enfrentó una dificultad	Control sobre el proceso de aprendizaje	Permite observar planificación y ajuste
Proyección de mejoras	Formula acciones para tareas futuras	Transferencia de lo aprendido	Vincula reflexión con desempeño posterior

Nota. Las evidencias metacognitivas adquieren valor cuando permiten relacionar la reflexión del estudiante con decisiones observables y cambios vinculados con los criterios de aprendizaje.

García (2024) destaca que la retroalimentación constituye un componente esencial de una evaluación orientada a mejorar los aprendizajes. Su función no debería limitarse a indicar qué estuvo bien o mal, sino proporcionar información que permita comprender la distancia entre el desempeño actual y el esperado. Cuando el estudiante puede explicar cómo utilizó esa información, la retroalimentación se convierte además en evidencia evaluativa. La respuesta frente al comentario permite observar si reconoce el criterio involucrado, si comprende la dificultad identificada y si puede traducir esa comprensión en una acción de mejora.

La metacognición puede documentarse mediante reflexiones breves vinculadas con momentos relevantes del proceso. Preguntas como qué cambió después de recibir retroalimentación, qué dificultad fue más importante o qué decisión sería diferente en una próxima actividad pueden generar evidencias significativas. Banihashem et al. (2022) muestran que la retroalimentación adquiere mayor potencial cuando promueve participación activa del estudiante. El objetivo no consiste en exigir extensos textos introspectivos, sino en obtener información suficiente para comprender cómo interpreta su desempeño y qué capacidad posee para ajustar estrategias a partir de los resultados obtenidos.

La inteligencia artificial generativa introduce nuevas posibilidades y desafíos dentro de estos procesos. Ossa y Willatt (2023) muestran que puede utilizarse para proporcionar retroalimentación sobre escritura académica, pero el aprendizaje depende de cómo el estudiante analiza y utiliza esas recomendaciones. Una herramienta puede sugerir reorganizar un argumento, ampliar una explicación o cambiar una formulación, pero la comprensión se hace visible cuando el estudiante evalúa la pertinencia de esa propuesta. La metacognición permite distinguir entre aplicar automáticamente una sugerencia y tomar una decisión consciente sobre su incorporación al trabajo.

Gombert et al. (2024) permiten comprender que la calidad de la retroalimentación depende de su capacidad para proporcionar información útil y accionable. En evaluación, esto implica seleccionar comentarios que ayuden al estudiante a identificar aspectos concretos que puede mejorar y evitar observaciones demasiado generales. La evidencia metacognitiva puede mostrar si el estudiante comprendió la información recibida y si fue capaz de transformarla en una estrategia. Esta relación entre comentario, interpretación y acción ofrece una secuencia especialmente valiosa para observar comprensión, porque permite vincular el pensamiento reflexivo con cambios efectivamente realizados.

La retroalimentación y la metacognición fortalecen la evaluación del proceso porque hacen visible cómo el estudiante interpreta su propio desempeño y responde ante oportunidades de mejora. de Schipper et al. (2021) muestran el potencial de sistemas personalizados de retroalimentación, mientras que García (2024) subraya su función dentro de una evaluación orientada al aprendizaje. Cuando las respuestas del estudiante se documentan de manera selectiva, pueden utilizarse como evidencias de comprensión, autorregulación y juicio. Estas evidencias contribuyen a construir una trayectoria más completa y preparan el análisis de sistemas de trazabilidad capaces de integrar la evolución del trabajo académico.

3.7 Sistemas de trazabilidad para verificar la evolución y autenticidad del trabajo académico

Los sistemas de trazabilidad permiten reconstruir de manera organizada la evolución de un trabajo académico y relacionar sus distintas evidencias dentro de una secuencia comprensible. Alfredo et al. (2024) muestran que la información generada durante el aprendizaje puede utilizarse para interpretar procesos, decisiones y cambios significativos. En evaluación, la trazabilidad no consiste en almacenar datos indiscriminadamente, sino en conectar momentos relevantes del desarrollo de una actividad. Esta organización facilita reconocer cómo un producto fue construido, qué transformaciones experimentó y qué evidencias respaldan la participación sostenida del estudiante.

La trazabilidad puede integrar registros procedentes de distintas fuentes, como versiones de documentos, entregas parciales, retroalimentaciones, reflexiones, portafolios o evidencias de interacción con herramientas digitales. Laamanen et al.

(2021) muestran que los mecanismos de autenticación en educación superior pueden contribuir a fortalecer la confiabilidad de los procesos cuando se implementan de manera compatible con las necesidades de los estudiantes. Esta perspectiva permite comprender que verificar autenticidad no depende de un único mecanismo, sino de la combinación razonable de evidencias capaces de aportar continuidad e identificación dentro del proceso académico.

Sharma (2026) sostiene que, frente a evaluaciones mediadas por inteligencia artificial generativa, resulta necesario hacer visible el juicio del estudiante y no concentrarse únicamente en determinar el origen de un producto. Un sistema de trazabilidad puede contribuir a este objetivo si registra decisiones relevantes, modificaciones y momentos de intervención tecnológica. La autenticidad se interpreta entonces a partir de la coherencia entre evidencias y no mediante una búsqueda aislada de señales de generación automática. Este enfoque desplaza la atención desde la detección hacia la reconstrucción razonada de la trayectoria académica.

La utilidad de estos sistemas depende de la capacidad para establecer relaciones entre las evidencias registradas. Chhetri Thapa y Lewis (2026) plantean que la evaluación orientada al proceso permite preservar la autenticidad epistémica cuando hace visible cómo se construye el conocimiento. Una secuencia de documentos adquiere valor trazable si permite identificar continuidad entre una idea inicial, una revisión posterior y una decisión final. La mera existencia de archivos cronológicos no garantiza esta función; es necesario que los registros permitan interpretar cambios, justificar transformaciones y reconocer la evolución intelectual del estudiante.

La trazabilidad también puede apoyarse en herramientas digitales que registran versiones, fechas, interacciones o modificaciones realizadas durante una actividad. Alfredo et al. (2024) destacan el potencial de los sistemas de analítica para utilizar datos educativos en la comprensión del aprendizaje. Sin embargo, estos registros deben interpretarse pedagógicamente y no asumirse como indicadores automáticos de calidad. Un historial extenso de modificaciones no demuestra por sí mismo mayor aprendizaje, del mismo modo que una secuencia breve puede contener decisiones conceptualmente significativas. La información necesita relacionarse con criterios y propósitos evaluativos previamente establecidos.

Khosravi et al. (2022) subrayan la importancia de la explicabilidad cuando la inteligencia artificial interviene en contextos educativos. Este principio resulta igualmente relevante para los sistemas de trazabilidad, porque los estudiantes deberían comprender qué información se registra, para qué será utilizada y cómo puede influir en la evaluación. Un sistema opaco puede generar desconfianza y transformar la documentación del aprendizaje en una práctica de vigilancia. La trazabilidad adquiere legitimidad cuando sus propósitos son claros, los registros son proporcionales y el estudiante puede comprender cómo las evidencias contribuyen a interpretar su evolución académica.

La autenticación constituye una dimensión complementaria de la trazabilidad, pero no debe confundirse con ella. Laamanen et al. (2021) muestran que los mecanismos de e-autenticación pueden contribuir a verificar identidad y participación en entornos de educación superior. Sin embargo, confirmar quién accede a una plataforma no explica necesariamente cómo se construyó el aprendizaje. La trazabilidad añade una dimensión temporal y procesual al conectar acciones, versiones y decisiones. De esta manera, autenticación y seguimiento pueden complementarse sin asumir que la identidad digital es suficiente para demostrar apropiación intelectual o comprensión disciplinar.

Otro criterio importante es la proporcionalidad del sistema. Sharma (2026) plantea que hacer visible el juicio no requiere convertir cada acción académica en una evidencia evaluable. Registrar excesivamente puede generar carga administrativa, afectar la autonomía y producir datos difíciles de interpretar. Por ello, la trazabilidad debería concentrarse en hitos relevantes: formulaciones iniciales, decisiones críticas, revisiones significativas, uso de retroalimentación y versiones finales. Esta selección permite reconstruir el recorrido de manera suficiente sin exigir una vigilancia continua de la actividad ni confundir cantidad de registros con calidad del aprendizaje.

Los sistemas de trazabilidad también pueden favorecer la retroalimentación cuando permiten comparar etapas y reconocer patrones de evolución. Chhetri Thapa y Lewis (2026) destacan que la orientación procesual fortalece la posibilidad de interpretar cómo se transforma el conocimiento durante una actividad. Si el docente puede observar dónde surgió una dificultad, cuándo se modificó una estrategia y qué evidencia provocó el cambio, dispone de mejores elementos para orientar futuras decisiones pedagógicas. La trazabilidad adquiere así una función formativa, porque no solo verifica continuidad, sino que ayuda a comprender qué factores contribuyeron a mejorar el desempeño.

La trazabilidad del aprendizaje fortalece la evaluación cuando integra evidencias relevantes, autenticación razonable y criterios comprensibles dentro de una secuencia coherente. Alfredo et al. (2024) muestran el potencial de los datos educativos para interpretar procesos, mientras que Khosravi et al. (2022) recuerdan la necesidad de mantener explicables las decisiones apoyadas por sistemas tecnológicos. Un diseño equilibrado permite verificar evolución y autenticidad sin reducir la evaluación a vigilancia digital. Esta perspectiva consolida el paso desde productos aislados hacia trayectorias documentadas capaces de representar con mayor precisión cómo se construye el aprendizaje universitario.

Capítulo 4.

Evaluación oral, práctica y situada

Darwin Paul Pozo Pilacuán



4.1 La recuperación de la oralidad como evidencia de aprendizaje

La recuperación de la oralidad como evidencia de aprendizaje responde a la necesidad de observar dimensiones del conocimiento que no siempre quedan representadas adecuadamente en productos escritos. Stephenson et al. (2025) muestran que las evaluaciones orales pueden ofrecer información relevante sobre comprensión, capacidad de respuesta y dominio del contenido. En contextos donde la inteligencia artificial generativa puede intervenir en la producción escrita, la oralidad adquiere renovado valor porque permite observar cómo el estudiante explica, conecta ideas, responde a preguntas y sostiene argumentos mediante una interacción que exige participación inmediata.

La oralidad no debe entenderse únicamente como exposición frente a un docente o grupo. Puede adoptar formas diversas, como entrevistas académicas, conversaciones evaluativas, presentaciones breves, explicaciones de procedimientos o intercambios estructurados de preguntas y respuestas. Brogan et al. (2024) destacan el potencial de la viva voce como modalidad capaz de complementar otras formas de evaluación. Su fortaleza reside en permitir que el estudiante demuestre apropiación del contenido mediante respuestas que requieren comprensión, organización del pensamiento y capacidad para adaptar el discurso a las preguntas que surgen durante la interacción.

Una característica importante de la evaluación oral es su capacidad para mostrar comprensión más allá de la memoria. El estudiante necesita explicar conceptos con sus propias palabras, establecer relaciones y responder ante variaciones de una pregunta inicial. Pearce y Chiavaroli (2023) sostienen que la aparición de inteligencia artificial generativa obliga a reconsiderar las formas tradicionales de evaluación. Desde esta perspectiva, la oralidad permite ampliar la evidencia disponible y comprobar si el estudiante puede sostener intelectualmente aquello que aparece en otros productos académicos, sin depender exclusivamente de la calidad formal de una entrega escrita.

La recuperación de la oralidad también favorece la observación de procesos de razonamiento. Una respuesta oral puede mostrar vacilaciones, reformulaciones, conexiones espontáneas o cambios de criterio que resultan útiles para comprender cómo piensa el estudiante. Stephenson et al. (2025) evidencian que el desempeño oral puede estar influido por diversos factores, lo que exige diseños cuidadosos y criterios claros. La evaluación debe centrarse en la calidad de la comprensión y de los argumentos,

evitando confundir fluidez verbal, seguridad escénica o rapidez de respuesta con dominio académico suficiente.

Tabla 12

Funciones de la oralidad como evidencia de aprendizaje en la evaluación universitaria

Función evaluativa	Manifestación observable	Evidencia de aprendizaje	Precaución necesaria
Explicación	Desarrollo oral de conceptos o decisiones	Comprensión y apropiación del contenido	No valorar solo fluidez verbal
Argumentación	Defensa de una postura con razones	Capacidad para fundamentar	Diferenciar opinión de argumento académico
Profundización	Respuesta a preguntas adicionales	Dominio más allá de respuestas preparadas	Ajustar dificultad al nivel formativo
Adaptación	Reformulación ante nuevas condiciones	Flexibilidad del razonamiento	Evitar preguntas ambiguas
Integración	Relación entre conceptos, datos y experiencias	Organización del conocimiento	Mantener criterios claros de valoración
Comunicación	Presentación de ideas con claridad y coherencia	Capacidad de expresar conocimiento	No sobrevalorar habilidades escénicas

Nota. La oralidad aporta evidencias especialmente útiles cuando se valora mediante criterios previamente definidos y se combina con otras modalidades de evaluación que permitan representar diferentes dimensiones del aprendizaje.

La oralidad puede resultar especialmente útil cuando el docente necesita comprobar la consistencia entre una producción presentada y el conocimiento que el estudiante demuestra sobre ella. Heil e Ifenthaler (2023) señalan que la evaluación en educación superior requiere diversificar formatos y considerar cuidadosamente la validez de las evidencias obtenidas. Una interacción oral breve puede permitir aclarar decisiones, profundizar argumentos o explorar aspectos que no quedaron suficientemente desarrollados en un documento. Esta función complementaria fortalece la interpretación del aprendizaje sin convertir la evaluación oral en un mecanismo de sospecha permanente.

Otra ventaja consiste en que la oralidad permite introducir variaciones durante la evaluación. El docente puede modificar una condición, solicitar un ejemplo distinto o pedir que el estudiante explique una idea desde otra perspectiva. Brogan et al. (2024) muestran que la viva voce puede ofrecer oportunidades para explorar el conocimiento

con mayor profundidad que una respuesta fija. Esta flexibilidad permite observar la capacidad para transferir y adaptar lo aprendido, siempre que las preguntas mantengan relación con los resultados de aprendizaje y no dependan de improvisaciones arbitrarias del evaluador.

La recuperación de la oralidad también exige reconocer posibles desigualdades en el desempeño. Algunos estudiantes pueden experimentar ansiedad, dificultades comunicativas o barreras específicas que influyen en su capacidad para responder oralmente. Stephenson et al. (2025) identifican factores que pueden facilitar o dificultar el rendimiento en evaluaciones orales. Por ello, su incorporación debe contemplar instrucciones claras, criterios transparentes, tiempos razonables y condiciones que reduzcan interferencias ajenas al aprendizaje que se pretende valorar. La oralidad aporta valor cuando se diseña con equidad y no cuando se utiliza como prueba improvisada de espontaneidad.

La planificación de criterios resulta esencial para evitar que la evaluación oral dependa excesivamente de impresiones subjetivas. Pearce y Chiavaroli (2023) permiten situar la necesidad de revisar las prácticas evaluativas frente a tecnologías generativas, pero cualquier modalidad alternativa debe conservar consistencia y transparencia. Los criterios pueden considerar precisión conceptual, capacidad de argumentación, uso pertinente de evidencias, coherencia y respuesta ante preguntas de profundización. Definir estas dimensiones antes de la actividad facilita decisiones más fundamentadas y permite que el estudiante comprenda qué aspectos de su desempeño serán observados.

La oralidad no necesita sustituir las evaluaciones escritas para aportar valor. Heil e Ifenthaler (2023) muestran que la diversidad de formatos puede ampliar las posibilidades de evaluación en educación superior. Una explicación oral puede complementar un informe, una presentación puede acompañar un proyecto y una conversación evaluativa puede utilizarse para aclarar decisiones presentes en un producto. Esta combinación fortalece la calidad de las evidencias porque permite observar distintas formas de demostrar conocimiento. La oralidad funciona mejor cuando se integra estratégicamente y no cuando se convierte en la única vía para valorar todo aprendizaje.

La recuperación de la oralidad adquiere así especial relevancia en un escenario donde las producciones escritas pueden estar ampliamente mediadas por inteligencia artificial

generativa. Brogan et al. (2024) muestran que las evaluaciones orales permiten explorar el conocimiento mediante interacción directa, mientras que Stephenson et al. (2025) destacan la necesidad de diseñarlas con condiciones que favorezcan un desempeño válido. Su principal aporte consiste en hacer visible la comprensión mediante explicación, argumentación y adaptación. Esta base permite avanzar hacia formas más estructuradas, como las defensas orales y sustentaciones académicas frente al uso de inteligencia artificial.

4.2 Defensas orales y sustentaciones académicas frente al uso de inteligencia artificial

Las defensas orales y las sustentaciones académicas adquieren una relevancia renovada en contextos donde la inteligencia artificial generativa puede participar en la elaboración de textos, informes, proyectos y otros productos universitarios. Brogan et al. (2024) destacan el valor de la viva voce como modalidad que permite explorar directamente el conocimiento del estudiante mediante interacción y argumentación. Su utilidad radica en que obliga a explicar decisiones, justificar contenidos y responder preguntas relacionadas con el trabajo presentado, generando evidencias que complementan aquello que puede observarse únicamente en una entrega escrita.

Una defensa oral permite valorar la apropiación intelectual del producto académico al exigir que el estudiante explique cómo se construyó, qué decisiones fueron relevantes y qué fundamentos sostienen sus conclusiones. En escenarios mediados por inteligencia artificial, esta exigencia resulta especialmente importante porque un trabajo puede presentar elevada calidad formal sin revelar con claridad el nivel de comprensión de quien lo entrega. Pearce y Chiavaroli (2023) sostienen que la aparición de sistemas generativos obliga a reconsiderar las formas de evaluación, favoreciendo modalidades capaces de comprobar comprensión y participación real en el aprendizaje.

La sustentación también permite profundizar en aspectos que no quedan completamente desarrollados en el producto escrito. El evaluador puede solicitar aclaraciones sobre una elección metodológica, una interpretación conceptual o una recomendación específica, obligando al estudiante a movilizar conocimientos más allá del contenido previamente preparado. Stephenson et al. (2025) muestran que las evaluaciones orales pueden ofrecer oportunidades valiosas para observar razonamiento y dominio del contenido. Esta capacidad de profundización resulta especialmente útil cuando se busca distinguir

entre una respuesta superficialmente correcta y una comprensión suficientemente sólida para sostenerla ante nuevas preguntas.

Brogan et al. (2024) señalan que la viva voce puede recuperar una forma de evaluación basada en el diálogo académico, donde el conocimiento se demuestra mediante intercambio y argumentación. Esta interacción permite explorar no solo qué sabe el estudiante, sino también cómo responde cuando se cuestionan sus decisiones o se introducen escenarios alternativos. La defensa oral adquiere así una función de comprobación dinámica, porque el estudiante debe reorganizar sus ideas en tiempo real y demostrar que puede sostener intelectualmente aquello que aparece en el producto presentado.

El uso de inteligencia artificial generativa no invalida necesariamente una sustentación, pero sí modifica las preguntas que conviene formular. Kofinas et al. (2025) muestran que las evaluaciones auténticas pueden verse afectadas por el uso de IA cuando no existen mecanismos que hagan visible la participación del estudiante. Una defensa puede incluir preguntas sobre qué partes del trabajo fueron desarrolladas con apoyo tecnológico, qué resultados fueron aceptados o rechazados y qué criterios se utilizaron para realizar esas decisiones. De esta manera, la sustentación permite valorar apropiación, juicio y responsabilidad sobre el producto.

La calidad de una defensa depende también de la preparación del evaluador. No basta con formular preguntas improvisadas o centradas únicamente en detalles memorísticos. Stephenson et al. (2025) evidencian que el desempeño oral puede verse condicionado por características del diseño evaluativo, por lo que resulta necesario establecer criterios y preguntas alineadas con los resultados de aprendizaje. Las preguntas deben buscar comprensión, razonamiento y capacidad de transferencia, evitando convertir la sustentación en una prueba de rapidez verbal. El objetivo es obtener evidencia válida sobre el conocimiento y no generar dificultad artificial.

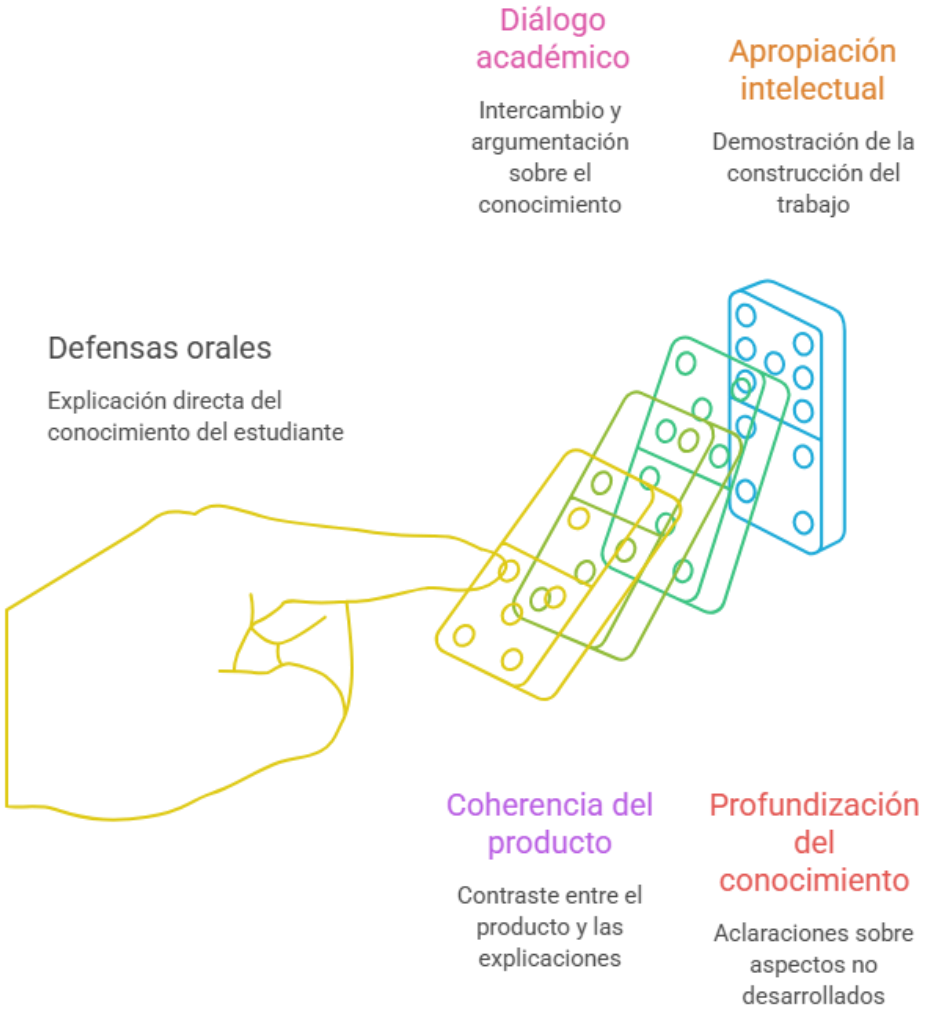
Otra función relevante de la defensa oral consiste en contrastar coherencia entre el producto presentado y las explicaciones del estudiante. Holden et al. (2021) señalan que la integridad académica en contextos de evaluación requiere considerar mecanismos que permitan fortalecer la confiabilidad de las evidencias obtenidas. Cuando el estudiante puede explicar con claridad conceptos, decisiones y procedimientos presentes en su trabajo, aumenta la consistencia entre producto y desempeño observado.

Este contraste no debe utilizarse como mecanismo automático de sospecha, sino como oportunidad para ampliar la base sobre la cual se interpreta el aprendizaje.

Las sustentaciones pueden organizarse en distintos formatos según la complejidad de la actividad. Pueden consistir en exposiciones breves seguidas de preguntas, entrevistas individuales, defensas ante un panel o conversaciones estructuradas sobre un producto previamente entregado. Brogan et al. (2024) muestran que la viva voce puede adaptarse a diferentes contextos y propósitos académicos. La elección del formato debe responder al tipo de evidencia que se busca obtener, evitando procedimientos excesivamente extensos cuando una interacción breve puede ser suficiente para comprobar comprensión y capacidad de argumentación.

Figura 7

Las defensas orales como estrategia de evaluación académica en la era de la inteligencia artificial



Nota. Se destacan los aportes de la defensa oral para verificar la comprensión, fortalecer la argumentación y comprobar la autoría intelectual del trabajo académico.

Pearce y Chiavaroli (2023) permiten comprender que la evaluación universitaria necesita diversificarse frente a la capacidad de la inteligencia artificial generativa para producir respuestas plausibles. Una defensa oral aporta valor porque introduce preguntas imprevisibles que obligan al estudiante a reaccionar ante variaciones del problema o a explicar con sus propias palabras aquello que presentó por escrito. Esta imprevisibilidad no debe confundirse con arbitrariedad; las preguntas deben mantenerse dentro del campo de aprendizaje esperado y utilizarse para observar flexibilidad, profundidad y comprensión, no para sorprender al estudiante.

La equidad también debe considerarse en el diseño de las sustentaciones. Stephenson et al. (2025) muestran que factores como ansiedad, familiaridad con el formato o condiciones comunicativas pueden afectar el desempeño oral. Por ello, las defensas requieren instrucciones previas, criterios transparentes y condiciones razonables que permitan demostrar conocimiento sin favorecer únicamente a quienes poseen mayor fluidez verbal. En determinados casos pueden utilizarse tiempos de preparación, preguntas estructuradas o ajustes de formato, siempre que se preserve la capacidad de valorar comprensión y razonamiento de manera comparable entre estudiantes.

Las defensas orales pueden además complementar otras evidencias sin necesidad de reemplazarlas. Kofinas et al. (2025) destacan la importancia de preservar la autenticidad de las evaluaciones en escenarios donde la IA puede intervenir en la producción académica. Una sustentación posterior a un proyecto, informe o estudio de caso permite ampliar la interpretación del trabajo y comprobar la relación entre producto y conocimiento demostrado. Esta combinación fortalece la evaluación porque integra evidencia escrita con desempeño oral, reduciendo la dependencia de una sola fuente para emitir juicios sobre el aprendizaje alcanzado.

Las defensas y sustentaciones académicas ofrecen, por tanto, una respuesta evaluativa especialmente útil frente a la expansión de la inteligencia artificial generativa. Holden et al. (2021) permiten situar esta práctica dentro de estrategias que fortalecen la confiabilidad de las evidencias, mientras que Brogan et al. (2024) muestran el potencial de la interacción oral para explorar comprensión y argumentación. Su valor no reside en detectar automáticamente el uso de IA, sino en hacer visible la apropiación intelectual del estudiante. Esta lógica prepara el tránsito hacia evaluaciones basadas en demostraciones prácticas y resolución directa de tareas.

4.3 Evaluación mediante demostraciones prácticas y resolución de tareas

La evaluación mediante demostraciones prácticas permite observar directamente cómo el estudiante aplica conocimientos, ejecuta procedimientos y responde ante tareas que exigen actuación. Edgar et al. (2022) muestran que las experiencias basadas en simulación pueden favorecer aprendizajes significativos cuando sitúan al estudiante frente a escenarios que requieren decisiones concretas. Desde esta perspectiva, el desempeño adquiere valor evaluativo porque permite identificar no solo si el estudiante conoce un procedimiento, sino también si puede ejecutarlo con precisión, adaptarlo a determinadas condiciones y explicar las razones que sustentan sus acciones.

La resolución directa de tareas también ofrece evidencias que complementan los productos escritos. Un estudiante puede describir correctamente cómo debería actuar ante una situación y, sin embargo, presentar dificultades cuando necesita ejecutar la misma acción. Gedera (2026) plantea que la evaluación auténtica debe permitir que el estudiante demuestre agencia y capacidad de actuación frente a condiciones significativas. La evaluación práctica hace visible esta diferencia al exigir una respuesta observable, donde conocimientos, habilidades y decisiones deben integrarse dentro de un desempeño que puede analizarse mediante criterios previamente definidos.

Pineda et al. (2024) muestran que las experiencias integradoras permiten articular competencias cuando el estudiante debe resolver situaciones mediante aplicación de conocimientos. En una demostración práctica, esta integración puede observarse en acciones como seleccionar un procedimiento, organizar recursos, interpretar datos o modificar una estrategia frente a una dificultad. El valor de la tarea aumenta cuando el desempeño no depende únicamente de seguir instrucciones mecánicas, sino de comprender qué se está haciendo y por qué. La ejecución se convierte así en una evidencia concreta de la capacidad para transformar conocimientos en acciones pertinentes.

La evaluación práctica puede adoptar formas muy diversas según la disciplina. En ciencias de la salud puede consistir en ejecutar un procedimiento o interpretar una situación clínica; en ingeniería, en operar, diseñar o diagnosticar un sistema; en educación, en planificar y desarrollar una intervención; y en áreas administrativas, en resolver escenarios de gestión. Huauya et al. (2023) señalan que el desarrollo de competencias requiere oportunidades para aplicar conocimientos dentro de actividades

significativas. Esta diversidad confirma que la demostración práctica no pertenece exclusivamente a carreras técnicas, sino a cualquier formación que exija desempeño observable.

Tabla 13
Modalidades de demostración práctica y evidencias de aprendizaje asociadas

Modalidad	Acción principal del estudiante	Evidencia observable	Criterios posibles
Ejecución de procedimiento	Aplica una secuencia de acciones	Precisión, orden y dominio técnico	Exactitud, seguridad y pertinencia
Resolución de tarea	Responde a un desafío concreto	Estrategia seleccionada y resultado	Eficacia, razonamiento y adaptación
Diagnóstico práctico	Analiza información y determina una situación	Interpretación y decisión	Coherencia, fundamentación y precisión
Producción aplicada	Construye un producto funcional	Calidad del resultado y proceso	Viabilidad, funcionalidad y adecuación
Demostración guiada	Ejecuta y explica simultáneamente	Acción acompañada de razonamiento	Comprensión y control del procedimiento
Tarea con variaciones	Ajusta su actuación ante cambios	Capacidad de respuesta	Flexibilidad y transferencia

Nota. Las modalidades deben seleccionarse de acuerdo con los resultados de aprendizaje y pueden combinar ejecución, explicación y adaptación para obtener evidencias más completas del desempeño.

Una característica importante de estas evaluaciones es la posibilidad de incorporar variaciones durante la tarea. Edgar et al. (2022) muestran que las experiencias simuladas pueden enfrentar al estudiante con condiciones que demandan respuesta y adaptación. El docente puede modificar datos, introducir una dificultad o cambiar una restricción para observar si el estudiante comprende el procedimiento o si simplemente reproduce una secuencia previamente memorizada. Esta capacidad de ajuste permite evaluar flexibilidad y transferencia, especialmente cuando el desempeño esperado requiere actuar ante situaciones que no siempre se presentan de manera idéntica.

La inteligencia artificial generativa puede apoyar algunas fases de una tarea práctica sin reemplazar necesariamente el desempeño que debe demostrar el estudiante. Gedera (2026) sostiene que una evaluación auténtica puede incorporar IA siempre que preserve la agencia humana. Por ejemplo, una herramienta puede utilizarse para explorar alternativas antes de una demostración, pero la ejecución, la adaptación y la explicación de decisiones deben permanecer vinculadas al estudiante cuando constituyen el núcleo

de la competencia evaluada. Esta distinción permite integrar tecnología sin desdibujar aquello que se pretende observar directamente durante la actividad.

La construcción de criterios claros resulta indispensable para evitar valoraciones basadas únicamente en impresiones generales. Huauya et al. (2023) muestran que la evaluación de competencias necesita considerar dimensiones relacionadas con desempeño y aplicación. En una demostración práctica, los criterios pueden incluir precisión, seguridad, secuencia, pertinencia de decisiones, capacidad de adaptación y calidad de la explicación. No todas las dimensiones necesitan tener el mismo peso, ya que su relevancia depende de la competencia evaluada. Lo importante es que el estudiante conozca previamente qué aspectos del desempeño serán considerados y por qué.

La resolución de tareas también puede incluir momentos de explicación durante o después de la ejecución. Pineda et al. (2024) permiten comprender que la integración de competencias requiere no solo actuar, sino también articular conocimientos y decisiones. Pedir al estudiante que explique por qué eligió un procedimiento, qué alternativa descartó o cómo respondería ante una variación permite interpretar mejor el desempeño observado. Esta combinación evita que una ejecución técnicamente correcta sea asumida automáticamente como evidencia de comprensión, especialmente cuando determinadas acciones pueden haber sido practicadas de forma rutinaria.

Las demostraciones prácticas y la resolución de tareas fortalecen la evaluación porque permiten observar capacidades que difícilmente pueden inferirse mediante un producto escrito aislado. Edgar et al. (2022) evidencian el valor de experiencias donde el estudiante debe responder activamente ante situaciones de desempeño, mientras que Gedera (2026) destaca la importancia de preservar agencia dentro de evaluaciones auténticas mediadas por tecnología. La combinación de ejecución, adaptación y explicación permite obtener evidencias más completas del aprendizaje y prepara el tránsito hacia escenarios profesionales simulados, donde las condiciones de actuación pueden adquirir mayor complejidad y realismo.

4.4 Simulaciones, escenarios profesionales y situaciones auténticas

Las simulaciones, los escenarios profesionales y las situaciones auténticas permiten aproximar la evaluación universitaria a contextos donde el estudiante debe actuar frente a condiciones similares a las que encontrará fuera del aula. Edgar et al. (2022) muestran

que las experiencias basadas en simulación pueden favorecer aprendizajes significativos cuando exigen aplicar conocimientos en contextos dinámicos. Su valor evaluativo reside en observar cómo el estudiante interpreta información, toma decisiones y ajusta su actuación ante situaciones que demandan una respuesta concreta, superando la lógica de ejercicios descontextualizados o exclusivamente teóricos.

Una simulación académica no necesita reproducir de manera exacta la realidad profesional para resultar útil. Lo importante es que conserve los elementos esenciales de la situación que hacen necesario movilizar conocimientos, habilidades y criterios. Gedera (2026) plantea que la autenticidad evaluativa depende de la relación entre contexto, tarea y agencia del estudiante. Por ello, un escenario puede ser simplificado si mantiene decisiones relevantes, restricciones y consecuencias que permitan observar desempeño. Esta flexibilidad facilita adaptar las simulaciones al nivel formativo sin perder la conexión con situaciones significativas.

Los escenarios profesionales pueden construirse alrededor de problemas, roles o responsabilidades específicas. Un estudiante puede asumir la función de docente, asesor, profesional sanitario, gestor, investigador o responsable de una decisión técnica, según la disciplina. Kofinas et al. (2025) destacan la importancia de preservar la autenticidad de la evaluación frente a la capacidad de la inteligencia artificial generativa para producir contenidos. Los roles profesionales fortalecen esta autenticidad cuando obligan al estudiante a responder a expectativas, destinatarios y condiciones concretas, reduciendo la utilidad de respuestas genéricas o excesivamente abstractas.

La situación auténtica también puede incorporar incertidumbre. En los contextos profesionales, la información disponible no siempre es completa ni permite una única solución evidente. García (2024) sostiene que la evaluación auténtica adquiere valor cuando aproxima el aprendizaje a situaciones donde es necesario aplicar conocimientos de manera significativa. Introducir datos incompletos, restricciones o intereses contrapuestos puede permitir observar cómo el estudiante prioriza información y fundamenta decisiones. La incertidumbre debe estar cuidadosamente diseñada para estimular razonamiento sin convertir la actividad en una experiencia confusa o arbitrariamente difícil.

Tabla 14

Elementos de diseño para simulaciones y escenarios profesionales auténticos

Elemento	Función evaluativa	Ejemplo de aplicación	Evidencia esperada
Rol profesional	Sitúa al estudiante en una responsabilidad concreta	Docente, analista, asesor, gestor o profesional sanitario	Capacidad para actuar según el contexto
Problema central	Define la situación que requiere intervención	Conflicto, diagnóstico, decisión o necesidad	Análisis y resolución
Restricciones	Introducen límites semejantes a condiciones reales	Tiempo, recursos, normativa o información parcial	Priorización y adaptación
Interacción	Exige responder a otros actores o cambios	Usuario, paciente, cliente, equipo o autoridad	Comunicación y capacidad de respuesta
Consecuencias	Relacionan decisiones con posibles resultados	Impactos, riesgos o efectos de una elección	Juicio y anticipación
Reflexión posterior	Permite revisar lo ocurrido	Explicación de decisiones y alternativas	Comprensión y aprendizaje derivado de la experiencia

Nota. Los componentes pueden combinarse según la disciplina y el nivel formativo; su finalidad es producir situaciones donde el desempeño requiera interpretar condiciones, decidir y justificar acciones en un contexto evaluativo significativo.

Edgar et al. (2022) muestran que las experiencias simuladas adquieren mayor valor cuando los estudiantes perciben que pueden aplicar aprendizajes dentro de situaciones relevantes. Esta percepción puede fortalecer la implicación y favorecer la transferencia de conocimientos hacia escenarios posteriores. En evaluación, la simulación permite observar capacidades que permanecen ocultas en una prueba convencional, como respuesta ante presión, adaptación a información nueva o coordinación de diferentes conocimientos. Sin embargo, el realismo debe servir al propósito educativo y no convertirse en un objetivo independiente del aprendizaje que se pretende valorar.

La inteligencia artificial generativa puede formar parte de algunos escenarios simulados. Gedera (2026) plantea que la IA puede integrarse dentro de evaluaciones auténticas si mantiene visible la agencia del estudiante. Por ejemplo, un sistema puede proporcionar información, generar alternativas o representar determinados actores dentro de una situación simulada. El estudiante continúa siendo responsable de interpretar esas respuestas y decidir cómo actuar. Esta integración permite aprovechar

capacidades tecnológicas sin convertir la simulación en una actividad automatizada donde las decisiones centrales dejan de depender del participante evaluado.

Los escenarios profesionales también permiten evaluar capacidades transversales además de conocimientos disciplinares. Comunicación, priorización, razonamiento, adaptación y toma de decisiones pueden observarse simultáneamente cuando forman parte natural de la situación planteada. Kofinas et al. (2025) vinculan las evaluaciones auténticas con la necesidad de preservar evidencias significativas de aprendizaje ante la expansión de sistemas generativos. Esta integración debe mantenerse alineada con los resultados formativos para evitar que una simulación termine valorando demasiadas dimensiones sin distinguir cuáles son realmente centrales para la competencia que se pretende examinar.

La preparación del estudiante influye de manera importante en la validez de estas evaluaciones. Una simulación no debería medir principalmente la familiaridad con el formato, sino las capacidades vinculadas con el aprendizaje esperado. García (2024) destaca que la evaluación auténtica requiere criterios claros y oportunidades para comprender las expectativas de desempeño. Proporcionar información sobre el propósito, las condiciones y los criterios permite que la dificultad provenga del problema evaluado y no de instrucciones ambiguas. La autenticidad no se fortalece mediante sorpresa, sino mediante exigencias realistas y suficientemente comprensibles.

La retroalimentación posterior puede convertir la simulación en una experiencia evaluativa y formativa al mismo tiempo. Edgar et al. (2022) muestran que las experiencias de simulación pueden generar oportunidades de reflexión sobre las decisiones realizadas y sus efectos. Después del escenario, el estudiante puede analizar qué hizo, por qué actuó de determinada manera y qué alternativa habría sido más pertinente. Esta revisión permite conectar desempeño y comprensión, evitando que la actividad se reduzca a una actuación puntual sin análisis posterior de las estrategias, dificultades y aprendizajes derivados de la experiencia.

Las simulaciones, escenarios profesionales y situaciones auténticas amplían la capacidad de la evaluación para observar conocimiento en acción. Gedera (2026) destaca que la autenticidad depende de preservar la agencia y la relación entre tarea y contexto, mientras que Kofinas et al. (2025) subrayan la necesidad de mantener evaluaciones capaces de producir evidencias significativas ante la inteligencia artificial

generativa. Cuando roles, restricciones, interacción y decisiones se integran con criterios claros, estas experiencias permiten comprobar no solo qué sabe el estudiante, sino cómo actúa cuando debe utilizar ese conocimiento frente a situaciones complejas.

4.5 Preguntas de profundización para comprobar comprensión y razonamiento

Las preguntas de profundización permiten comprobar si el estudiante comprende realmente los contenidos y puede sostener sus respuestas más allá de formulaciones previamente preparadas. Stephenson et al. (2025) muestran que las evaluaciones orales ofrecen oportunidades para explorar comprensión, razonamiento y capacidad de respuesta ante nuevas exigencias. Una pregunta bien formulada obliga a explicar relaciones, justificar decisiones o reconsiderar una postura, generando evidencias que complementan el producto inicial. Su valor no reside en sorprender al estudiante, sino en abrir espacios donde el conocimiento deba reorganizarse y expresarse de manera razonada.

Una de sus principales funciones consiste en distinguir entre reconocimiento superficial y comprensión profunda. El estudiante puede repetir una definición correctamente y, sin embargo, presentar dificultades para aplicarla en una situación diferente o explicar sus implicaciones. Brogan et al. (2024) destacan que la viva voce permite ampliar el análisis mediante preguntas que siguen el razonamiento del estudiante. Solicitar ejemplos, consecuencias o fundamentos convierte una respuesta inicial en punto de partida para examinar cómo se organiza el conocimiento y hasta qué nivel puede ser movilizado frente a exigencias que no estaban completamente anticipadas.

Las preguntas de profundización pueden orientarse hacia la justificación de decisiones. Expresiones como “¿por qué eligió esta alternativa?”, “¿qué criterio utilizó?” o “¿qué evidencia sostiene esa conclusión?” obligan al estudiante a hacer explícito el razonamiento que normalmente permanece oculto en el producto presentado. Bearman et al. (2024) relacionan el juicio evaluativo con la capacidad para reconocer y justificar la calidad de diferentes opciones. Esta clase de preguntas permite observar si una decisión fue tomada de manera consciente o si simplemente fue aceptada porque parecía plausible, especialmente cuando existió apoyo de inteligencia artificial.

Otro tipo de preguntas introduce variaciones sobre la situación original. Modificar una condición, cambiar un dato o agregar una restricción permite observar si el estudiante puede adaptar su razonamiento sin depender exclusivamente de una respuesta

memorizada. Pearce y Chiavaroli (2023) sostienen que la inteligencia artificial generativa obliga a reconsiderar las formas de comprobar aprendizaje, precisamente porque puede producir respuestas correctas para escenarios previsibles. Las variaciones permiten examinar transferencia y flexibilidad, siempre que mantengan relación directa con los conocimientos y competencias que se pretende valorar dentro de la actividad.

Las preguntas comparativas también aportan información relevante. Pedir al estudiante que contraste dos alternativas, identifique ventajas y limitaciones o determine cuál opción sería más adecuada bajo determinadas condiciones permite observar la calidad de su juicio. Stephenson et al. (2025) muestran que el desempeño oral puede revelar dimensiones del aprendizaje que no siempre aparecen en evaluaciones estáticas. La comparación exige organizar criterios, reconocer diferencias y establecer prioridades, por lo que ofrece una evidencia más rica que una pregunta centrada únicamente en recuperar información previamente estudiada.

Brogan et al. (2024) permiten comprender que la interacción oral puede adaptarse de acuerdo con las respuestas que van surgiendo durante la evaluación. Esta flexibilidad favorece preguntas encadenadas, donde una primera respuesta conduce a una segunda cuestión más específica. El evaluador puede solicitar aclaración, pedir un contraargumento o explorar una consecuencia no mencionada inicialmente. Sin embargo, esta adaptabilidad necesita criterios claros para evitar diferencias injustificadas entre estudiantes. La profundidad puede variar, pero las preguntas deben mantenerse dentro de un nivel comparable de exigencia y vincularse con los mismos resultados de aprendizaje.

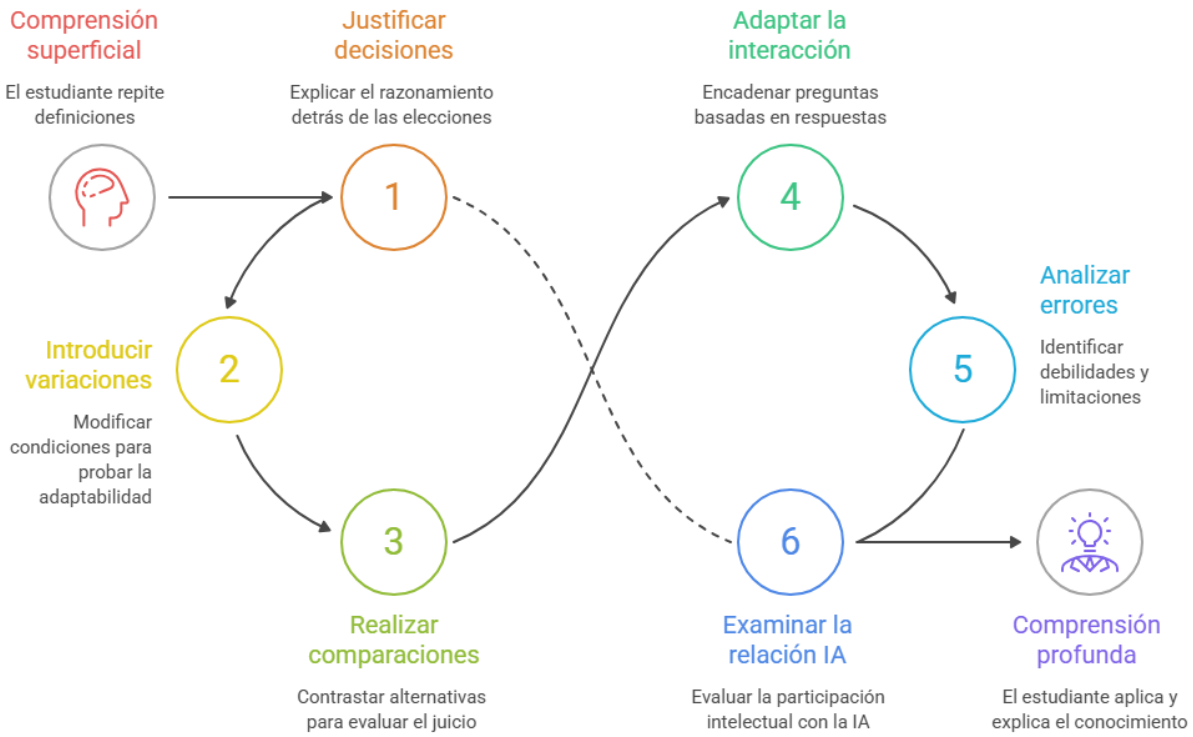
Las preguntas sobre errores o limitaciones también resultan útiles para comprobar comprensión. Solicitar al estudiante que identifique una debilidad en su propia propuesta o que explique qué podría salir mal obliga a adoptar una perspectiva crítica sobre el trabajo realizado. Bearman et al. (2024) destacan la importancia de desarrollar capacidad para juzgar la calidad de las producciones. Reconocer límites demuestra que el estudiante no interpreta su respuesta como una solución incuestionable, sino que puede evaluar sus alcances, condiciones de validez y posibles ajustes frente a información o escenarios diferentes.

En contextos donde se ha utilizado inteligencia artificial generativa, las preguntas de profundización pueden centrarse en la relación entre el estudiante y el contenido

presentado. Preguntar qué parte requirió mayor revisión, qué sugerencia fue descartada o cómo se verificó determinada afirmación permite examinar participación intelectual sin convertir la interacción en interrogatorio sobre la tecnología. Pearce y Chiavaroli (2023) permiten situar estas prácticas dentro de una evaluación que necesita obtener evidencias adicionales de comprensión. El foco debe permanecer en el conocimiento demostrado y en las decisiones que sostienen el producto académico.

Figura 8

Progresión de la comprensión mediante preguntas de profundización



Nota. Se ilustra una secuencia de preguntas que permite superar respuestas superficiales, fortalecer el razonamiento crítico y verificar la comprensión del estudiante frente al uso de inteligencia artificial.

La formulación de preguntas debe evitar ambigüedades innecesarias. Una pregunta demasiado amplia puede dificultar que el estudiante identifique qué tipo de razonamiento se espera, mientras que una excesivamente cerrada puede reducir la oportunidad de profundizar. Stephenson et al. (2025) muestran que las características de la evaluación oral influyen en el desempeño, lo que exige equilibrio entre apertura y claridad. Las mejores preguntas delimitan suficientemente el problema, pero dejan espacio para que el estudiante organice una respuesta propia, seleccione argumentos y demuestre comprensión mediante conexiones que no fueron proporcionadas directamente por el evaluador.

La secuencia de preguntas también puede utilizarse para diferenciar niveles de comprensión. Una primera cuestión puede solicitar explicación, una segunda aplicación y una tercera valoración crítica o adaptación a nuevas condiciones. Brogan et al. (2024) muestran que la evaluación oral permite explorar progresivamente el conocimiento mediante interacción. Esta organización evita comenzar con preguntas excesivamente complejas y permite observar hasta dónde puede avanzar el estudiante en su razonamiento. El objetivo no es aumentar dificultad de manera artificial, sino identificar la profundidad con la que puede utilizar y transformar aquello que ha aprendido.

Las preguntas de profundización fortalecen la evaluación porque convierten la respuesta inicial en una oportunidad para explorar comprensión, razonamiento y capacidad de adaptación. Bearman et al. (2024) permiten relacionar esta práctica con el desarrollo del juicio evaluativo, mientras que Stephenson et al. (2025) muestran el valor de las interacciones orales para obtener evidencias más completas del aprendizaje. Cuando se formulan con claridad y propósito, estas preguntas permiten comprobar no solo qué sabe el estudiante, sino cómo utiliza ese conocimiento frente a nuevas exigencias y condiciones.

4.6 Evaluación situada y transferencia del conocimiento a contextos profesionales

La evaluación situada permite valorar el aprendizaje en condiciones que guardan relación directa con los contextos donde el conocimiento deberá utilizarse posteriormente. Edgar et al. (2022) muestran que las experiencias de simulación pueden fortalecer la aplicación de aprendizajes cuando sitúan al estudiante frente a situaciones que requieren decisiones y respuestas concretas. Desde esta perspectiva, evaluar no consiste únicamente en comprobar si el estudiante recuerda información, sino en observar si puede movilizarla de manera pertinente frente a demandas específicas, restricciones y consecuencias semejantes a aquellas que pueden aparecer en escenarios profesionales.

La transferencia del conocimiento constituye uno de los principales indicadores de que el aprendizaje ha superado la reproducción académica. Méndez et al. (2024) destacan que el enfoque basado en competencias adquiere mayor sentido cuando los estudiantes pueden aplicar lo aprendido en situaciones diferentes de aquellas utilizadas durante la enseñanza. Transferir implica reconocer qué conocimientos son pertinentes, adaptarlos a nuevas condiciones y utilizarlos para resolver problemas. Por ello, una evaluación

situada debe evitar reproducir exactamente los ejercicios practicados y generar oportunidades para comprobar si el estudiante puede responder ante variaciones relevantes.

Pineda et al. (2024) muestran que los proyectos integradores permiten articular competencias y acercar el aprendizaje a problemas vinculados con contextos reales. Esta relación resulta especialmente útil para evaluar transferencia porque obliga al estudiante a conectar conocimientos procedentes de distintas áreas y utilizarlos de manera coordinada. Un desempeño profesional rara vez depende de un contenido aislado; suele exigir interpretar información, establecer prioridades y considerar consecuencias. La evaluación situada puede observar esta integración mediante tareas donde el estudiante deba construir una respuesta coherente frente a una situación que reúna diversas exigencias.

La transferencia también exige distinguir entre aplicar mecánicamente un procedimiento y comprender cuándo resulta adecuado utilizarlo. Kofinas et al. (2025) señalan que las evaluaciones auténticas deben preservar evidencias significativas de aprendizaje incluso en contextos donde la inteligencia artificial puede apoyar la producción académica. Una tarea situada adquiere mayor valor cuando obliga a seleccionar entre diferentes procedimientos, adaptar una estrategia o justificar por qué determinada alternativa resulta pertinente. Esta capacidad de elección permite observar un dominio más profundo que la simple ejecución de una secuencia conocida previamente.

Tabla 15

Indicadores para valorar la transferencia del conocimiento en contextos profesionales

Indicador	Manifestación en el desempeño	Ejemplo de evidencia	Criterio de valoración
Selección pertinente	Elige conocimientos adecuados para la situación	Procedimiento, estrategia o marco conceptual seleccionado	Correspondencia con el problema
Adaptación	Ajusta lo aprendido a condiciones nuevas	Modificación de una estrategia ante restricciones	Flexibilidad y pertinencia
Integración	Articula saberes de distintas áreas	Respuesta que combina conceptos, datos y procedimientos	Coherencia del desempeño

Indicador	Manifestación en el desempeño	Ejemplo de evidencia	Criterio de valoración
Justificación	Explica por qué adopta determinada decisión	Argumentación vinculada con criterios profesionales	Fundamentación
Anticipación	Considera consecuencias de sus decisiones	Identificación de riesgos o efectos posibles	Juicio profesional
Transferencia	Aplica aprendizajes en un escenario distinto	Resolución de una situación no practicada previamente	Capacidad de generalización contextualizada

Nota. Los indicadores permiten valorar la capacidad de movilizar aprendizajes en situaciones diferentes, considerando no solo la aplicación del conocimiento, sino también su adaptación, integración y justificación.

Edgar et al. (2022) evidencian que las experiencias simuladas pueden contribuir a que los estudiantes perciban mayor conexión entre aprendizaje y práctica. Esta relación favorece evaluaciones donde el conocimiento se utiliza dentro de escenarios con propósito claro. Una situación situada puede requerir responder a un usuario, interpretar un caso, resolver una dificultad técnica o tomar una decisión organizacional. Lo importante es que las condiciones obliguen al estudiante a transformar el conocimiento en actuación y permitan observar cómo responde cuando la tarea se aproxima a las demandas propias de una práctica profesional.

La transferencia puede evaluarse mediante cambios deliberados en el contexto. Una misma competencia puede plantearse inicialmente en una situación conocida y posteriormente trasladarse a otra con actores, datos o restricciones diferentes. Méndez et al. (2024) permiten comprender que el enfoque competencial requiere demostrar capacidad de aplicación más allá de una circunstancia específica. Si el estudiante solo puede actuar cuando reconoce exactamente el formato practicado, la evidencia de transferencia es limitada. La evaluación situada debe observar si identifica principios relevantes y los reorganiza de manera adecuada frente a nuevas condiciones.

La inteligencia artificial generativa introduce un elemento adicional dentro de los contextos profesionales contemporáneos. Kofinas et al. (2025) muestran que su presencia obliga a preservar la autenticidad de las evidencias evaluativas. En determinadas actividades situadas, la IA puede formar parte del escenario como recurso de consulta, apoyo o generación de alternativas, siempre que el estudiante conserve responsabilidad sobre la decisión final. La transferencia se hace visible cuando puede

utilizar críticamente la tecnología, contextualizar sus resultados y reconocer cuándo una respuesta generada no se ajusta a las exigencias específicas del problema profesional.

Pineda et al. (2024) muestran que los proyectos integradores pueden favorecer competencias transversales cuando los estudiantes deben responder a situaciones que demandan colaboración, planificación y aplicación. Este tipo de experiencias permite evaluar transferencia desde una perspectiva más amplia, porque el desempeño profesional suele requerir combinar conocimientos con comunicación, organización y toma de decisiones. Sin embargo, la evaluación debe delimitar con claridad qué competencias pretende observar para evitar que una tarea compleja produzca evidencias difíciles de interpretar. La autenticidad necesita acompañarse de criterios específicos y vinculados con resultados de aprendizaje.

La evaluación situada adquiere mayor valor cuando permite observar no solo si el estudiante resuelve una situación, sino cómo interpreta el contexto y adapta sus conocimientos. Edgar et al. (2022) evidencian que las experiencias prácticas y simuladas pueden favorecer aprendizajes aplicados, mientras que Méndez et al. (2024) destacan la importancia de evaluar competencias mediante situaciones de desempeño. La transferencia se convierte así en una evidencia central de aprendizaje universitario, porque muestra que el conocimiento puede utilizarse más allá del aula y prepara la integración de evidencias escritas, orales y prácticas dentro de una valoración más completa.

4.7 Integración de evidencias escritas, orales y prácticas en la evaluación universitaria

La integración de evidencias escritas, orales y prácticas permite construir una valoración más completa del aprendizaje universitario al reconocer que una sola modalidad difícilmente representa todas las capacidades desarrolladas por el estudiante. Heil e Ifenthaler (2023) muestran que la evaluación en educación superior puede adoptar múltiples formatos y que la calidad de las decisiones depende de la adecuación entre evidencia y propósito. Combinar modalidades permite observar conocimientos conceptuales, capacidad argumentativa, desempeño y aplicación, reduciendo la dependencia de un único producto para interpretar el nivel de aprendizaje alcanzado.

Las evidencias escritas aportan información especialmente útil sobre organización conceptual, argumentación, selección de información y capacidad para estructurar

respuestas complejas. Sin embargo, su valor aumenta cuando pueden contrastarse con otras formas de desempeño. Stephenson et al. (2025) muestran que la evaluación oral permite explorar dimensiones de comprensión que no siempre quedan visibles en documentos previamente elaborados. Una producción escrita puede demostrar calidad académica, mientras una explicación posterior permite observar si el estudiante comprende sus fundamentos, puede defender sus decisiones y responde coherentemente ante preguntas vinculadas con el contenido presentado.

Las evidencias prácticas añaden otra dimensión porque permiten observar cómo el estudiante transforma el conocimiento en actuación. Gedera (2026) plantea que la evaluación auténtica requiere tareas donde la agencia y el desempeño puedan hacerse visibles en contextos significativos. Un estudiante puede explicar correctamente un procedimiento y redactar una propuesta bien fundamentada, pero la ejecución práctica permite comprobar si puede aplicar esos conocimientos bajo condiciones concretas. Integrar las tres modalidades favorece una interpretación más rica al relacionar lo que el estudiante escribe, lo que puede explicar y aquello que logra realizar.

La combinación de evidencias no significa repetir la misma actividad en tres formatos. Kofinas et al. (2025) destacan la importancia de preservar la autenticidad de la evaluación mediante tareas que produzcan evidencias significativas. Cada modalidad debería cumplir una función específica: la escrita puede documentar análisis y fundamentación; la oral puede explorar comprensión y razonamiento; la práctica puede mostrar aplicación y capacidad de adaptación. La integración resulta útil cuando las evidencias se complementan y permiten examinar diferentes dimensiones de una misma competencia sin generar actividades redundantes o una carga innecesaria.

Bearman et al. (2024) destacan la relevancia del juicio evaluativo para interpretar la calidad de producciones y decisiones. Esta capacidad resulta especialmente importante cuando se combinan evidencias heterogéneas, porque el docente necesita establecer cómo se relacionan entre sí y qué peso tendrá cada una. Una respuesta oral débil no debería anular automáticamente un producto escrito sólido, del mismo modo que una ejecución práctica satisfactoria no necesariamente compensa una fundamentación insuficiente. La integración requiere criterios que permitan interpretar el conjunto y reconocer fortalezas, inconsistencias y patrones del desempeño del estudiante.

La triangulación de evidencias puede fortalecer la confiabilidad de la evaluación cuando diferentes modalidades apuntan hacia conclusiones similares sobre el aprendizaje. Heil e Ifenthaler (2023) permiten comprender que la selección de formatos debe responder a los objetivos formativos y a las condiciones de aplicación. Si un estudiante explica con claridad un concepto, lo utiliza adecuadamente en una producción escrita y lo aplica correctamente en una situación práctica, existe una base más amplia para interpretar su competencia. Esta convergencia reduce la dependencia de circunstancias particulares asociadas con una sola modalidad de evaluación.

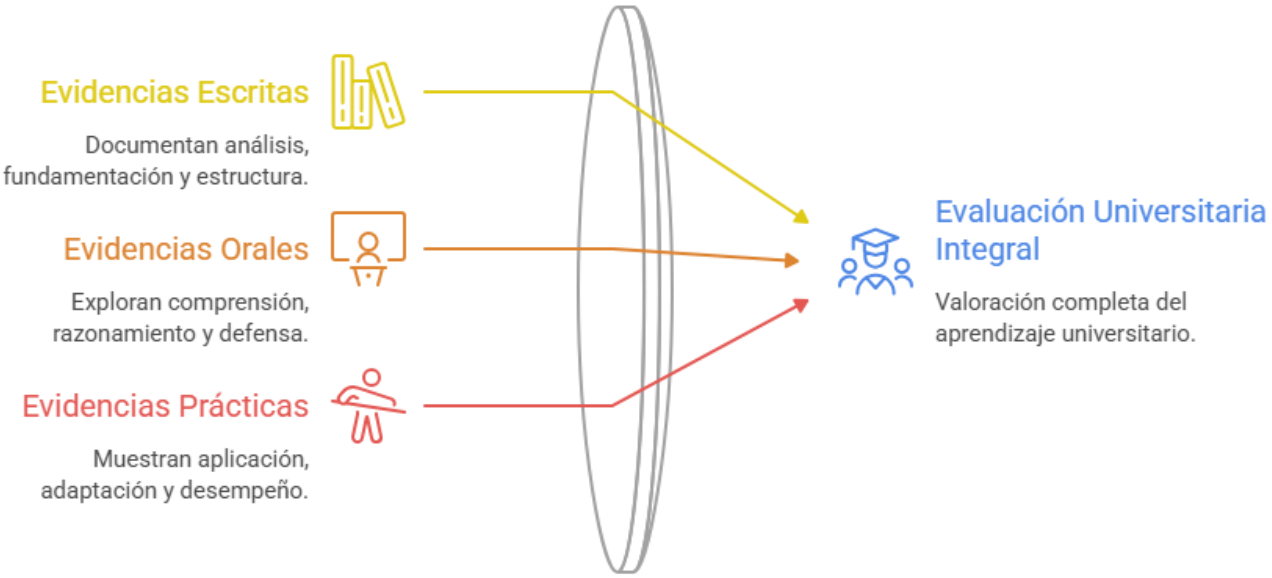
También pueden aparecer discrepancias entre las evidencias, y estas diferencias ofrecen información útil. Un estudiante puede presentar un informe de alta calidad, pero mostrar dificultades para explicar sus decisiones oralmente; otro puede demostrar gran comprensión durante una interacción y tener limitaciones para estructurarla por escrito. Stephenson et al. (2025) muestran que el desempeño oral posee características propias que deben interpretarse cuidadosamente. Las divergencias no deberían entenderse automáticamente como fallos, sino como señales que permiten identificar dimensiones diferenciadas del aprendizaje y orientar decisiones posteriores de retroalimentación o apoyo.

En escenarios donde existe inteligencia artificial generativa, la integración de modalidades adquiere un valor adicional. Kofinas et al. (2025) muestran que estas tecnologías pueden influir en la autenticidad de determinados productos, especialmente cuando la evaluación depende únicamente de entregas elaboradas fuera de interacción directa. Combinar una producción escrita con una explicación oral o una demostración práctica permite ampliar las evidencias sin convertir la evaluación en un mecanismo de detección tecnológica. El interés se desplaza hacia la coherencia entre lo producido, lo explicado y lo ejecutado por el estudiante.

Gedera (2026) señala que las evaluaciones auténticas deben mantener la agencia del estudiante incluso cuando incorporan herramientas tecnológicas. Esta idea resulta central para integrar evidencias, porque cada modalidad puede mostrar una dimensión diferente de esa agencia. El estudiante puede justificar decisiones en un documento, defenderlas oralmente y ponerlas en práctica en una situación determinada. La consistencia entre estas formas de desempeño permite interpretar con mayor precisión su participación intelectual y profesional, especialmente cuando el uso de inteligencia artificial forma parte legítima del entorno de aprendizaje.

Figura 9

Integración de modalidades de evaluación en la educación universitaria



Nota. Se representa la complementariedad de las evidencias escritas, orales y prácticas para valorar integralmente la comprensión, el razonamiento y la aplicación del conocimiento estudiantil.

En escenarios donde existe inteligencia artificial generativa, la integración de modalidades adquiere un valor adicional. Kofinas et al. (2025) muestran que estas tecnologías pueden influir en la autenticidad de determinados productos, especialmente cuando la evaluación depende únicamente de entregas elaboradas fuera de interacción directa. Combinar una producción escrita con una explicación oral o una demostración práctica permite ampliar las evidencias sin convertir la evaluación en un mecanismo de detección tecnológica. El interés se desplaza hacia la coherencia entre lo producido, lo explicado y lo ejecutado por el estudiante.

Gedera (2026) señala que las evaluaciones auténticas deben mantener la agencia del estudiante incluso cuando incorporan herramientas tecnológicas. Esta idea resulta central para integrar evidencias, porque cada modalidad puede mostrar una dimensión diferente de esa agencia. El estudiante puede justificar decisiones en un documento, defenderlas oralmente y ponerlas en práctica en una situación determinada. La consistencia entre estas formas de desempeño permite interpretar con mayor precisión su participación intelectual y profesional, especialmente cuando el uso de inteligencia artificial forma parte legítima del entorno de aprendizaje.

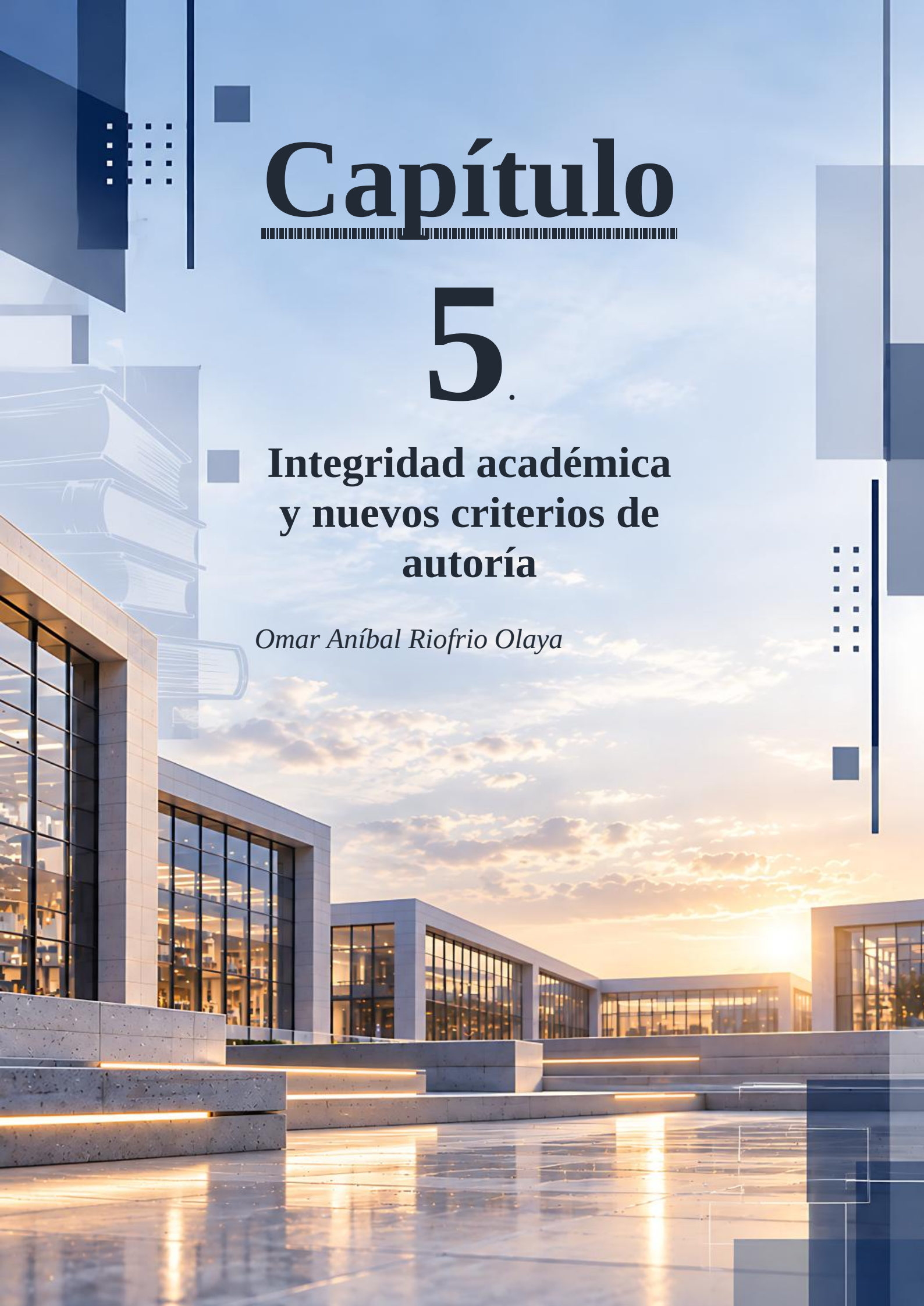
La integración de evidencias escritas, orales y prácticas configura una evaluación universitaria más equilibrada al reconocer que el aprendizaje se manifiesta mediante

distintas formas de actuación. Bearman et al. (2024) permiten vincular esta integración con la necesidad de emitir juicios apoyados en evidencias diversas, mientras que Heil e Ifenthaler (2023) muestran la importancia de seleccionar formatos adecuados al propósito evaluativo. La combinación estratégica de modalidades reduce la dependencia de una única evidencia y permite valorar comprensión, argumentación, aplicación y transferencia dentro de una visión más completa del desempeño universitario.

Capítulo 5.

Integridad académica y nuevos criterios de autoría

Omar Aníbal Riofrio Olaya



5.1 Transformación del concepto de integridad académica ante la inteligencia artificial generativa

La integridad académica ha sido entendida tradicionalmente como un conjunto de principios vinculados con honestidad, responsabilidad y respeto por la autoría dentro de los procesos educativos. La expansión de la inteligencia artificial generativa obliga a revisar esta comprensión porque introduce formas de producción que no encajan completamente en categorías clásicas como copia, plagio o fraude. Tan y Maravilla (2024) sostienen que estas tecnologías no necesariamente debilitan la educación, pero sí exigen nuevos marcos para interpretar cómo participan en el trabajo académico y qué responsabilidades deben conservar quienes las utilizan.

La transformación conceptual comienza por reconocer que un producto puede ser original en términos de coincidencia textual y, al mismo tiempo, contener una participación humana limitada. Holden et al. (2021) muestran que la integridad en la evaluación depende de condiciones capaces de preservar la confianza en aquello que el estudiante presenta como evidencia de aprendizaje. Con sistemas generativos, esta confianza ya no puede apoyarse únicamente en determinar si existe copia de una fuente previa. Resulta necesario considerar también cuánto control intelectual ejerció el estudiante sobre el contenido y qué decisiones asumió como propias.

Alessio y Messinger (2021) destacan que las percepciones de estudiantes y docentes sobre integridad pueden variar según las condiciones tecnológicas en las que se desarrolla el aprendizaje y la evaluación. Esta variabilidad adquiere mayor relevancia con la inteligencia artificial generativa, porque prácticas consideradas aceptables en un contexto pueden interpretarse de manera diferente en otro. Utilizar una herramienta para corregir estilo, generar ideas o producir una respuesta completa implica grados distintos de intervención. La integridad académica necesita, por tanto, criterios capaces de diferenciar usos y no depender de prohibiciones uniformes.

Navarrete-Cazales et al. (2026) sitúan la discusión de autoría e integridad académica frente a la inteligencia artificial generativa dentro de un escenario donde las fronteras entre producción humana y asistencia tecnológica se vuelven más complejas. Esto obliga a incorporar elementos como transparencia, responsabilidad y trazabilidad de las decisiones. La integridad deja de concentrarse exclusivamente en evitar apropiaciones indebidas y comienza a incluir la obligación de reconocer cómo fue construido el

producto académico. Esta ampliación permite interpretar con mayor precisión la participación real del estudiante dentro de procesos mediados por inteligencia artificial.

Tabla 16

Transformaciones del concepto de integridad académica ante la inteligencia artificial generativa

Dimensión	Enfoque tradicional	Enfoque ampliado ante IA generativa	Implicación evaluativa
Autoría	Producción atribuida al estudiante	Responsabilidad sobre contenido humano y asistido	Exige reconocer niveles de intervención tecnológica
Originalidad	Ausencia de copia o plagio	Apropiación intelectual y transformación del contenido	Requiere valorar control sobre lo producido
Transparencia	Citación de fuentes utilizadas	Declaración de herramientas y formas de apoyo	Amplía las obligaciones de información
Responsabilidad	Responder por el trabajo presentado	Verificar y asumir decisiones asistidas por IA	Mantiene la responsabilidad humana
Honestidad	No presentar trabajo ajeno como propio	No ocultar sustitución significativa del trabajo académico	Obliga a diferenciar asistencia y delegación
Confianza	Basada en autoría presumida	Basada en evidencias, criterios y declaraciones claras	Fortalece la interpretación del aprendizaje

Nota. El enfoque ampliado no elimina los principios tradicionales de integridad, sino que incorpora nuevas responsabilidades relacionadas con transparencia, control intelectual y participación de herramientas generativas.

La responsabilidad constituye uno de los componentes que más se fortalece dentro de esta transformación. Carvajal et al. (2026) muestran que estudiantes y docentes pueden mantener percepciones diferentes sobre la relación entre inteligencia artificial e integridad académica, lo que evidencia la necesidad de establecer expectativas claras. El estudiante continúa siendo responsable del producto presentado aun cuando utilice sistemas generativos. Esta responsabilidad implica revisar exactitud, coherencia, pertinencia y adecuación del contenido, evitando asumir que la participación tecnológica reduce la obligación de responder académicamente por los resultados entregados.

La integridad académica también necesita diferenciar entre asistencia y sustitución. Una herramienta puede apoyar actividades secundarias sin reemplazar el razonamiento que constituye el objetivo de una evaluación, pero también puede asumir funciones

centrales hasta disminuir sustancialmente la participación del estudiante. Tan y Maravilla (2024) plantean que el impacto de la inteligencia artificial depende de cómo se integre dentro del aprendizaje. Por ello, el problema de integridad no puede definirse únicamente por la presencia de IA, sino por el grado en que su intervención altera aquello que el estudiante debía demostrar mediante su propio desempeño.

Holden et al. (2021) permiten comprender que la integridad académica requiere condiciones institucionales y evaluativas que reduzcan ambigüedades sobre lo que se considera comportamiento aceptable. En entornos generativos, estas condiciones deben incluir orientaciones específicas sobre funciones permitidas, formas de declaración y límites de asistencia. La ausencia de reglas claras puede producir interpretaciones contradictorias entre asignaturas o docentes. La transformación del concepto de integridad exige entonces pasar de expectativas implícitas a criterios comunicados de manera explícita, especialmente cuando distintas herramientas pueden intervenir en múltiples fases del trabajo académico.

Otra modificación importante se relaciona con la confianza. Tradicionalmente, esta podía apoyarse en la presunción de que el estudiante había elaborado personalmente aquello que presentaba. Alessio y Messinger (2021) muestran que la integridad académica depende también de percepciones compartidas entre quienes participan en el proceso educativo. Con inteligencia artificial generativa, la confianza necesita sostenerse en reglas comprensibles, prácticas de transparencia y diseños evaluativos que hagan visible la participación del estudiante. De esta manera, confianza e integridad dejan de depender exclusivamente de vigilancia y comienzan a construirse mediante expectativas y evidencias consistentes.

Navarrete-Cazales et al. (2026) permiten ampliar esta discusión hacia la noción de autoría, porque los sistemas generativos desafían la relación directa entre escribir un contenido y asumir su creación intelectual. La autoría académica necesita asociarse no solo con producción textual, sino con responsabilidad sobre decisiones, selección, interpretación y validación. Un estudiante puede utilizar apoyo tecnológico sin perder necesariamente su condición de autor, siempre que mantenga control significativo sobre el trabajo. La integridad depende entonces de que esa participación pueda explicarse y resulte compatible con los objetivos formativos establecidos.

La transformación del concepto también exige evitar enfoques basados exclusivamente en detección y sanción. Carvajal et al. (2026) evidencian que la integración de inteligencia artificial en educación superior plantea desafíos que requieren considerar percepciones, prácticas y responsabilidades. Una política centrada únicamente en identificar uso tecnológico puede resultar insuficiente si no distingue entre apoyo legítimo y sustitución indebida. La integridad académica necesita orientarse hacia prácticas preventivas y formativas que ayuden a comprender por qué determinadas decisiones afectan la validez de una evaluación y qué responsabilidades acompañan el uso de herramientas generativas.

La inteligencia artificial generativa amplía, por tanto, el significado de integridad académica desde una lógica centrada en evitar plagio hacia otra que incorpora autoría, transparencia, responsabilidad y control intelectual. Tan y Maravilla (2024) permiten comprender que estas tecnologías pueden coexistir con prácticas educativas íntegras cuando su uso se articula con propósitos claros, mientras que Navarrete-Cazales et al. (2026) destacan la necesidad de revisar los criterios de autoría. Esta transformación prepara el análisis de conductas específicas donde el problema ya no reside únicamente en copiar, sino en utilizar sistemas generativos sin declarar su participación.

5.2 Del plagio tradicional al uso no declarado de sistemas generativos

El plagio tradicional se ha definido principalmente a partir de la apropiación de ideas, palabras o producciones ajenas sin el reconocimiento correspondiente. La inteligencia artificial generativa introduce una dificultad adicional porque permite producir contenido nuevo sin copiar literalmente una fuente identificable. Holden et al. (2021) muestran que la integridad académica en entornos digitales exige considerar prácticas que van más allá de la reproducción textual. En este escenario, el problema ya no consiste únicamente en detectar coincidencias, sino en determinar cuándo una producción presentada como propia ha sido delegada significativamente a un sistema automatizado.

El uso no declarado de sistemas generativos puede adoptar formas muy diferentes, desde pequeñas intervenciones de apoyo hasta la producción sustancial de una tarea académica. Tan y Maravilla (2024) sostienen que la inteligencia artificial no necesariamente conduce a prácticas deshonestas, pero su utilización requiere criterios capaces de distinguir entre asistencia y sustitución. Una herramienta empleada para

corregir estilo no plantea las mismas implicaciones que otra utilizada para construir argumentos, resolver un problema o redactar una entrega completa. La valoración ética depende, por tanto, del nivel y finalidad de la intervención tecnológica.

Carvajal et al. (2026) evidencian que estudiantes y docentes pueden interpretar de manera distinta los límites aceptables del uso de inteligencia artificial en educación superior. Esta diversidad de percepciones explica por qué el uso no declarado puede convertirse en una fuente de conflictos académicos incluso cuando no existe intención explícita de engañar. Si las reglas institucionales son ambiguas, el estudiante puede asumir que determinada asistencia resulta permitida mientras el docente la interpreta como sustitución indebida. La integridad exige reducir estas zonas grises mediante expectativas claras y contextualizadas.

A diferencia del plagio convencional, el contenido generado por inteligencia artificial puede no corresponder a un autor humano fácilmente identificable ni derivar de una fuente concreta que pueda citarse de forma tradicional. Zdravkova (2026) destaca precisamente la ambigüedad que estas tecnologías introducen en los conceptos de autoría e integridad. Esta característica dificulta aplicar de manera automática los marcos diseñados para detectar copia textual. Una producción puede ser inédita y, al mismo tiempo, no representar adecuadamente el trabajo intelectual del estudiante, lo que obliga a revisar cómo se interpreta la originalidad académica.

El uso no declarado adquiere especial gravedad cuando la inteligencia artificial realiza aquello que constituye el núcleo de la competencia evaluada. Si una actividad pretende valorar argumentación, razonamiento, programación, traducción o escritura, delegar esas operaciones puede alterar directamente la validez de la evaluación. Cabrera et al. (2026) muestran que la garantía de integridad en educación superior requiere examinar cómo intervienen las tecnologías generativas en el trabajo estudiantil. La cuestión central no es únicamente si se utilizó IA, sino si su participación sustituyó capacidades que debían ser demostradas personalmente.

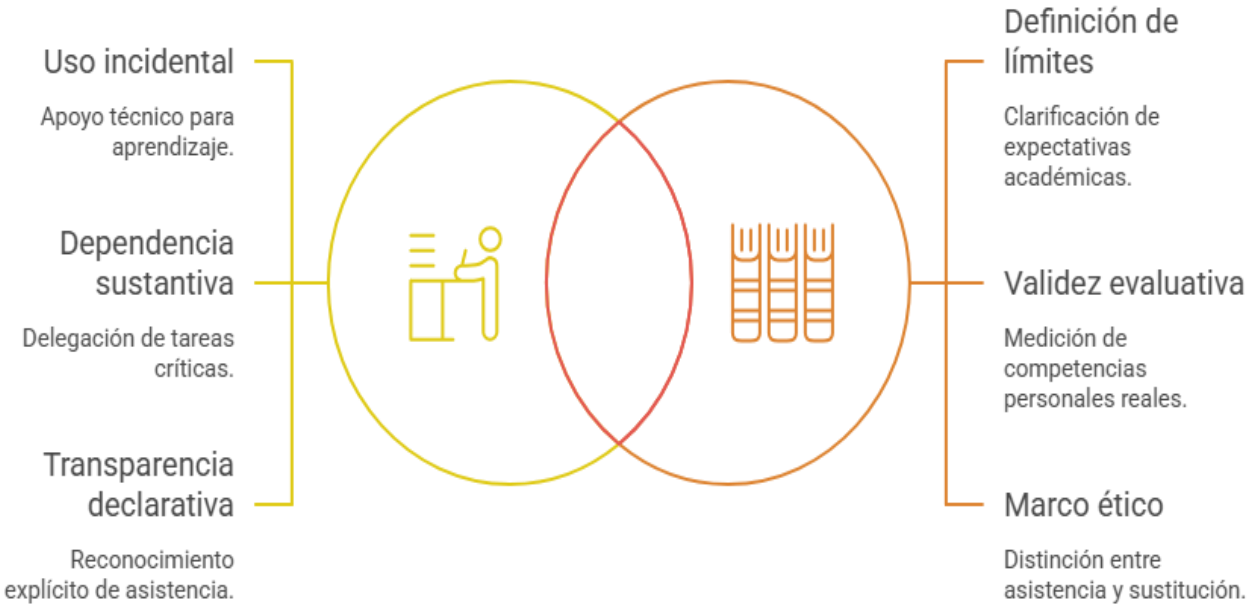
Holden et al. (2021) permiten comprender que la confiabilidad de una evaluación depende de que las evidencias presentadas correspondan razonablemente al desempeño del estudiante. Desde esta perspectiva, ocultar un uso significativo de inteligencia artificial puede afectar la interpretación del aprendizaje aun cuando el contenido sea correcto. El problema no radica exclusivamente en la herramienta, sino en la

discrepancia entre lo que el producto aparenta demostrar y aquello que realmente realizó quien lo presenta. La no declaración puede así comprometer la validez del juicio académico sobre conocimientos y competencias.

También es necesario diferenciar entre uso incidental y dependencia sustantiva. Tan y Maravilla (2024) plantean que los sistemas generativos pueden integrarse de manera constructiva dentro de la educación cuando mantienen la participación y responsabilidad humanas. Consultar una explicación, solicitar ejemplos o revisar una formulación puede formar parte de un proceso legítimo, mientras que delegar decisiones centrales modifica la naturaleza del trabajo. La integridad académica requiere reconocer estas diferencias y evitar clasificar todo uso de inteligencia artificial dentro de una única categoría de conducta indebida.

Figura 10

Integridad académica digital y uso responsable de la inteligencia artificial



Nota. Se destacan las condiciones éticas y evaluativas que permiten diferenciar el apoyo tecnológico legítimo de la dependencia académica, promoviendo la transparencia y la autenticidad del aprendizaje.

Carvajal et al. (2026) muestran que el debate sobre integridad académica en contextos virtuales está estrechamente relacionado con las percepciones de quienes participan en ellos. Esta dimensión resulta importante porque la conducta no declarada puede normalizarse cuando los estudiantes consideran que el uso de sistemas generativos es equivalente a emplear cualquier herramienta digital. Frente a esta percepción, las universidades necesitan explicar por qué determinadas intervenciones tecnológicas

afectan la evidencia de aprendizaje y otras no. La claridad pedagógica resulta más efectiva que depender únicamente de prohibiciones generales o sanciones posteriores.

El uso no declarado también presenta dificultades desde la perspectiva de la detección. Los sistemas diseñados para identificar similitud textual fueron creados para responder principalmente a formas tradicionales de plagio y no necesariamente pueden determinar con precisión si un contenido fue generado por inteligencia artificial. Zdravkova (2026) señala que esta ambigüedad tecnológica complica la aplicación de criterios de integridad. Por ello, una política basada exclusivamente en detectores puede producir conclusiones poco confiables. La evaluación necesita apoyarse en evidencias más amplias y en reglas claras sobre responsabilidad y declaración.

Cabrera et al. (2026) permiten situar este problema dentro de un contexto institucional más amplio, donde la integridad requiere combinar prevención, orientación y criterios consistentes. El uso no declarado puede disminuir cuando los estudiantes conocen de antemano qué formas de asistencia son aceptables, cuáles deben informarse y cuáles resultan incompatibles con el propósito de una evaluación. Esta claridad permite trasladar la discusión desde la pregunta "¿se utilizó inteligencia artificial?" hacia cuestiones más relevantes: "¿para qué se utilizó?", "¿qué parte del trabajo realizó?" y "¿qué responsabilidad conserva el estudiante?".

Otra diferencia respecto del plagio tradicional es que el uso de inteligencia artificial puede ocurrir en múltiples momentos de una misma actividad. Un estudiante puede generar ideas, reformular párrafos, consultar alternativas o construir una versión inicial que posteriormente modifica. Holden et al. (2021) permiten comprender que las prácticas de integridad deben adaptarse a las condiciones reales de producción académica. Por ello, no siempre resulta adecuado clasificar una tarea completa como legítima o ilegítima sin considerar el tipo de intervención tecnológica. La evaluación necesita interpretar grados de asistencia y su impacto sobre la evidencia obtenida.

El tránsito desde el plagio tradicional hacia el uso no declarado de sistemas generativos amplía el campo de la integridad académica. Tan y Maravilla (2024) permiten comprender que el problema no puede reducirse a prohibir herramientas, mientras que Zdravkova (2026) evidencia la necesidad de revisar categorías clásicas ante nuevas formas de producción. La cuestión central pasa a ser cuánto control intelectual conserva el estudiante, qué intervención tecnológica ocurrió y si fue comunicada de acuerdo con

las reglas establecidas. Este cambio conduce directamente al análisis de la autoría humana y la responsabilidad sobre los productos académicos.

5.3 Autoría humana y responsabilidad sobre los productos académicos

La autoría humana adquiere nuevas implicaciones cuando un producto académico se construye con apoyo de sistemas generativos capaces de producir textos, argumentos, estructuras o propuestas completas. Navarrete-Cazales et al. (2026) plantean que la inteligencia artificial obliga a revisar los criterios tradicionales de autoría porque ya no siempre existe correspondencia directa entre quien presenta un trabajo y quien genera cada uno de sus componentes. En este contexto, la autoría debe vincularse con el control intelectual del proceso, la capacidad para tomar decisiones y la responsabilidad sobre aquello que finalmente se incorpora al producto académico.

Savvidou (2026) muestra que los estudiantes negocian de manera compleja la relación entre asistencia tecnológica, responsabilidad y sentido de autoría cuando utilizan inteligencia artificial para escribir. Esta realidad sugiere que la autoría no puede reducirse a quién redactó cada frase, sino que debe considerar quién define el propósito, selecciona contenidos, revisa resultados y asume las consecuencias de las decisiones tomadas. Un estudiante puede utilizar herramientas generativas y conservar una participación autoral significativa si mantiene control sobre la orientación del trabajo y transforma críticamente los aportes tecnológicos.

La responsabilidad sobre el producto constituye un criterio central para delimitar la autoría humana. Culebro et al. (2025) muestran que el uso de inteligencia artificial en la edición y producción de textos universitarios introduce preguntas sobre quién puede considerarse autor cuando una parte del contenido surge de procesos automatizados. Desde una perspectiva académica, el estudiante continúa siendo responsable de verificar exactitud, pertinencia, coherencia y adecuación disciplinar. La participación de una herramienta no desplaza esta obligación, porque el producto presentado sigue formando parte de una evaluación asociada con su aprendizaje y desempeño.

Zdravkova (2026) advierte que la inteligencia artificial generativa introduce ambigüedades importantes en torno a la integridad y la autoría, especialmente cuando la contribución tecnológica adquiere un peso sustancial. Esta ambigüedad obliga a establecer criterios más precisos para identificar cuándo existe apoyo legítimo y cuándo se produce una sustitución del trabajo intelectual. La autoría humana conserva sentido

cuando el estudiante puede explicar qué decisiones tomó, qué partes modificó y qué responsabilidad asume sobre el contenido, incluso si determinadas fases del proceso fueron asistidas tecnológicamente.

Tabla 17
Criterios para reconocer autoría humana en productos académicos asistidos por inteligencia artificial

Criterio	Evidencia esperada	Indicador de autoría humana	Riesgo cuando está ausente
Intencionalidad	Definición clara del propósito del trabajo	El estudiante orienta la producción	Dependencia de resultados generados
Selección	Elección razonada de contenidos y alternativas	Existe juicio sobre qué incorporar	Aceptación automática
Transformación	Modificación, adaptación o reorganización	El contenido es reelaborado críticamente	Reproducción pasiva
Verificación	Revisión de exactitud y pertinencia	El estudiante valida lo utilizado	Incorporación de errores o afirmaciones no comprobadas
Justificación	Explicación de decisiones relevantes	Existe apropiación intelectual	Dificultad para sostener el producto
Responsabilidad	Asunción de consecuencias académicas	El estudiante responde por el resultado final	Dilución de responsabilidad

Nota. Los criterios permiten valorar la participación autoral del estudiante más allá de la producción textual directa, considerando control intelectual, revisión y responsabilidad sobre el producto final.

La transformación del concepto de autoría también obliga a diferenciar entre contribución técnica y contribución intelectual. Una herramienta puede corregir estilo, reorganizar un texto o sugerir formulaciones sin necesariamente asumir el núcleo de la tarea. Navarrete-Cazales et al. (2026) permiten comprender que la autoría académica debe asociarse con la responsabilidad sobre las decisiones sustantivas y con la capacidad para demostrar apropiación del contenido. Esta distinción resulta especialmente importante cuando el producto final presenta una calidad formal elevada que podría haber sido obtenida con apoyo tecnológico sin implicar una comprensión equivalente.

Savvidou (2026) muestra que los estudiantes pueden percibir la inteligencia artificial como colaboradora, herramienta o asistente, dependiendo del grado y finalidad de su intervención. Estas percepciones influyen en la forma en que entienden su propia

responsabilidad sobre el producto. Sin embargo, desde la evaluación académica, la autoría necesita mantenerse vinculada con la capacidad para responder por lo presentado. Si el estudiante no puede explicar, defender o revisar críticamente una idea incorporada mediante IA, su participación autoral se vuelve más débil, aunque formalmente figure como único responsable del trabajo.

La noción de autoría humana también requiere evitar dos simplificaciones. La primera consiste en asumir que cualquier uso de IA elimina automáticamente la autoría; la segunda, en considerar que basta con presentar el producto para conservarla. Culebro et al. (2025) muestran que la participación tecnológica puede formar parte de procesos de edición y construcción académica sin anular necesariamente la responsabilidad del estudiante. La clave se encuentra en determinar si existe control significativo sobre las decisiones y si la persona puede sostener intelectualmente aquello que entrega como resultado de su trabajo.

La responsabilidad también implica reconocer límites. Un estudiante puede utilizar inteligencia artificial para explorar alternativas y, aun así, debe asumir los errores, omisiones o sesgos que incorpora al producto final. Zdravkova (2026) permite comprender que la ambigüedad de los sistemas generativos no exime al usuario de responsabilidad académica. La autoría humana se fortalece cuando existe conciencia de que aceptar una sugerencia generada equivale a incorporarla como propia dentro de la tarea. Esto exige revisar, contrastar y decidir con criterios disciplinarios antes de integrar cualquier contenido al trabajo evaluado.

La autoría en contextos mediados por inteligencia artificial puede definirse, por tanto, como la combinación de intención, selección, transformación, verificación, justificación y responsabilidad. Navarrete-Cazales et al. (2026) muestran que la relación entre autoría e integridad necesita adaptarse a nuevas formas de producción, mientras que Savvidou (2026) destaca la importancia de cómo los estudiantes interpretan su papel frente a herramientas generativas. Esta perspectiva permite superar la idea de autoría como simple producción textual y prepara el análisis de la transparencia y declaración del uso de inteligencia artificial dentro del trabajo académico.

5.4 Transparencia y declaración del uso de inteligencia artificial

La transparencia en el uso de inteligencia artificial constituye una condición necesaria para preservar la confianza en los productos académicos cuando intervienen sistemas

generativos. Sharma (2026) plantea que la evaluación debe hacer visible el juicio del estudiante y las decisiones que sustentan el trabajo presentado. Desde esta perspectiva, declarar el uso de inteligencia artificial no responde únicamente a una exigencia formal, sino a la necesidad de mostrar cómo se relacionaron los aportes tecnológicos con la producción humana. La transparencia permite interpretar mejor la responsabilidad académica y evita que la asistencia tecnológica permanezca oculta dentro del producto final.

La declaración debe proporcionar información suficiente para comprender la función que cumplió la inteligencia artificial dentro de la actividad. Indicar únicamente que se utilizó una herramienta resulta insuficiente si no se explica para qué fue empleada. Savvidou (2026) muestra que los estudiantes pueden utilizar sistemas generativos de maneras muy diferentes durante la escritura académica, desde apoyo para organizar ideas hasta generación de contenido sustantivo. Por ello, una declaración útil debería especificar el propósito del uso y el tipo de intervención realizada, permitiendo distinguir entre asistencia limitada y participación tecnológica más amplia.

Navarrete-Cazales et al. (2026) sostienen que la relación entre autoría e integridad académica se vuelve más compleja cuando intervienen herramientas generativas. Esta complejidad hace necesario incorporar la transparencia como parte de la responsabilidad autoral. Si el estudiante declara que utilizó inteligencia artificial para reformular, comparar alternativas o desarrollar una parte determinada del trabajo, facilita la interpretación de aquello que corresponde a su propia contribución intelectual. La transparencia no disminuye automáticamente la autoría; puede, por el contrario, fortalecerla cuando permite reconocer cómo se mantuvo el control sobre las decisiones académicas.

La declaración también debe ser proporcional al nivel de intervención tecnológica. UNESCO (2023) plantea que el uso de inteligencia artificial en educación debe orientarse mediante principios claros y centrados en la responsabilidad humana. En una actividad donde la IA tuvo una participación mínima, una declaración breve puede resultar suficiente; cuando intervino en fases sustantivas, se requiere mayor detalle. Esta proporcionalidad evita tanto la omisión de información relevante como la creación de procedimientos excesivamente burocráticos que dificulten el trabajo académico sin aportar mejoras significativas a la interpretación de la evaluación.

Temper et al. (2025) destacan la necesidad de establecer marcos institucionales capaces de regular el uso de inteligencia artificial en educación superior. Desde esta perspectiva, la transparencia no debería depender exclusivamente de decisiones individuales de cada docente. Las universidades necesitan definir criterios comunes sobre cuándo declarar el uso, qué información proporcionar y cómo incorporar esa declaración dentro de los productos académicos. La existencia de reglas compartidas reduce contradicciones entre asignaturas y facilita que los estudiantes comprendan que declarar la IA forma parte de una práctica institucional y no de una exigencia arbitraria.

La transparencia debe diferenciarse del registro detallado del proceso de aprendizaje. En una declaración académica no siempre resulta necesario presentar todos los prompts, capturas o interacciones realizadas con una herramienta. Sharma (2026) permite comprender que lo relevante es hacer visible el juicio y la participación del estudiante. Por ello, la declaración puede centrarse en los usos que influyeron significativamente en el producto final y en las decisiones asociadas con esos aportes. Esta función ética y comunicativa distingue la transparencia de otros mecanismos orientados específicamente a reconstruir la trazabilidad del proceso.

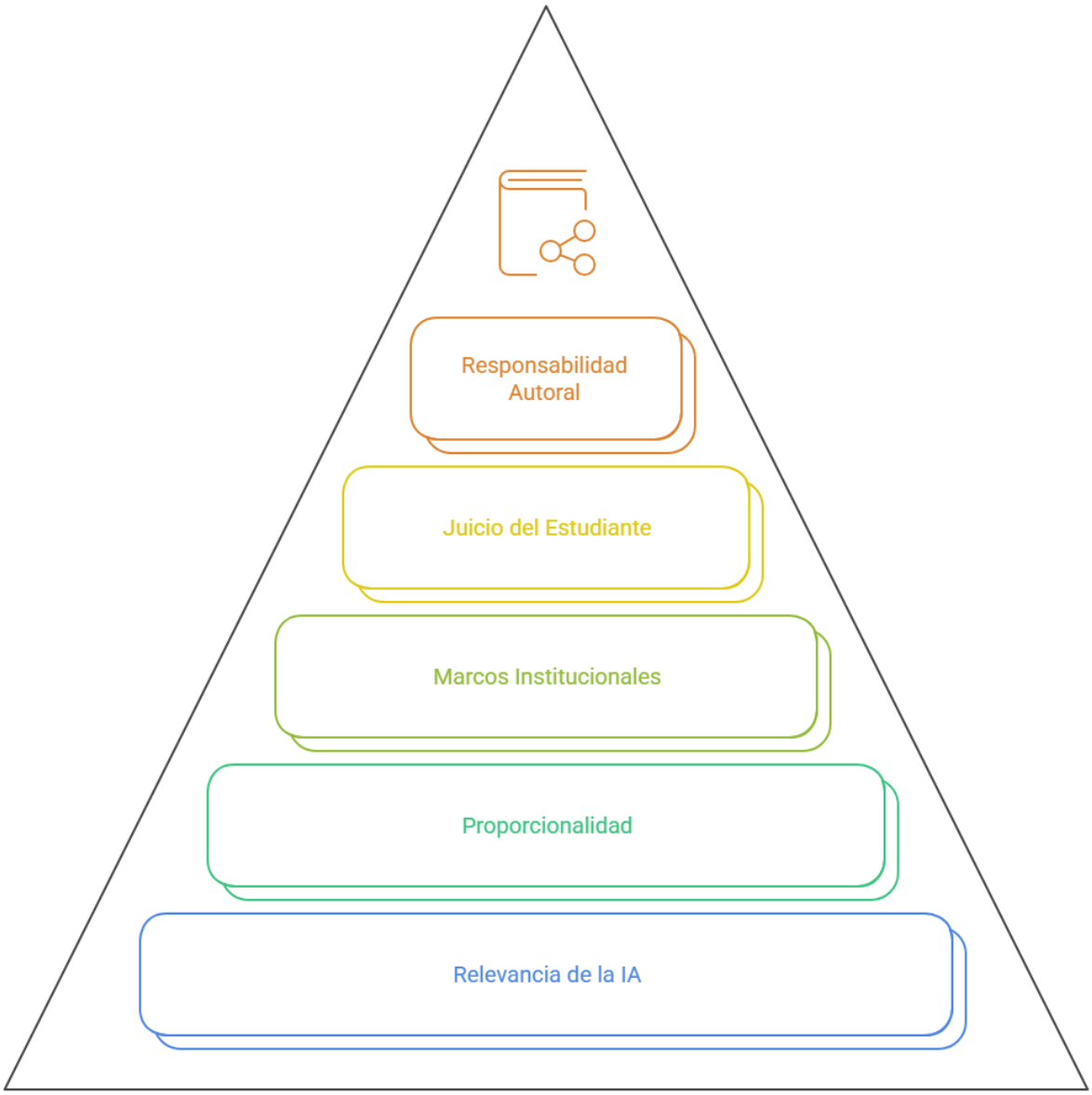
Savvidou (2026) muestra que la percepción de autoría puede modificarse según el nivel de dependencia respecto de la inteligencia artificial. Esta relación refuerza la importancia de que la declaración permita comprender no solo qué herramienta se utilizó, sino cuánto influyó en la elaboración del producto. Una breve descripción sobre si la IA ayudó a explorar alternativas, revisar redacción, estructurar contenido o generar partes específicas puede proporcionar contexto suficiente. La transparencia adquiere valor cuando permite interpretar con mayor precisión la relación entre asistencia tecnológica, participación humana y responsabilidad sobre el resultado final.

La declaración también puede contribuir a reducir ambigüedades sobre la legitimidad del uso tecnológico. Navarrete-Cazales et al. (2026) señalan que las nuevas condiciones de producción académica requieren criterios capaces de distinguir diferentes formas de participación de la inteligencia artificial. Cuando el estudiante conoce que determinados usos están permitidos siempre que sean informados, disminuye la necesidad de ocultarlos. La transparencia puede favorecer una relación más abierta con la tecnología, donde el cumplimiento académico no depende de aparentar ausencia de IA, sino de utilizarla dentro de límites conocidos y asumir responsabilidad sobre su incorporación.

UNESCO (2023) plantea que las instituciones educativas necesitan desarrollar orientaciones que promuevan un uso responsable de la inteligencia artificial generativa. La declaración puede formar parte de esta orientación si se concibe como práctica educativa y no únicamente como control. Explicar de manera breve cómo se utilizó una herramienta obliga al estudiante a reconocer su función dentro del trabajo y puede favorecer una mayor conciencia sobre las decisiones tomadas. Así, la transparencia no solo informa al evaluador, sino que también contribuye a que el estudiante examine críticamente la relación entre tecnología y aprendizaje.

Figura 11

Jerarquía de principios para la transparencia en el uso académico de inteligencia artificial



Nota. Se organizan los principios que orientan la integración responsable de la inteligencia artificial, desde su pertinencia educativa hasta la responsabilidad autoral del estudiante.

La estandarización de formatos puede facilitar la declaración cuando existe diversidad de herramientas y usos. Temper et al. (2025) permiten comprender que las instituciones necesitan criterios consistentes para regular la inteligencia artificial en educación superior. Un formato sencillo puede incluir nombre de la herramienta, propósito, alcance de la intervención y responsabilidad asumida sobre el contenido final. No es necesario convertir esta declaración en un documento extenso; su función principal consiste en ofrecer información relevante y comparable que permita comprender cómo participó la IA dentro del producto evaluado.

La transparencia también debe preservar condiciones de equidad. Si algunos docentes exigen declaraciones detalladas y otros no establecen ninguna orientación, los estudiantes pueden enfrentar expectativas inconsistentes dentro de una misma institución. Sharma (2026) muestra que la evaluación en contextos aumentados por inteligencia artificial requiere criterios claros y visibles. Por ello, las reglas de declaración deberían comunicarse antes de la actividad y aplicarse de manera coherente. Esta claridad permite que todos los estudiantes conozcan las mismas condiciones y reduce interpretaciones posteriores basadas en criterios que no fueron informados previamente.

La declaración del uso de inteligencia artificial constituye, por tanto, una práctica vinculada con transparencia, responsabilidad y autoría académica. Navarrete-Cazales et al. (2026) permiten situarla dentro de los nuevos criterios de integridad, mientras que Savvidou (2026) muestra cómo la presencia de sistemas generativos modifica la percepción de participación autoral. Declarar no significa renunciar a la autoría, sino hacer explícita la relación entre estudiante y herramienta. Esta transparencia permite avanzar hacia una delimitación más clara de los límites éticos entre asistencia tecnológica, colaboración y sustitución del trabajo académico.

5.5 Límites éticos entre asistencia tecnológica, colaboración y sustitución del trabajo académico

Los límites éticos entre asistencia tecnológica, colaboración y sustitución del trabajo académico dependen del grado en que la inteligencia artificial interviene sobre las capacidades que el estudiante debe demostrar. Tan y Maravilla (2024) sostienen que el uso de sistemas generativos no tiene por qué debilitar la educación cuando existe participación humana significativa y responsabilidad sobre el resultado. Desde esta

perspectiva, la cuestión ética no se resuelve preguntando únicamente si se utilizó inteligencia artificial, sino determinando qué función cumplió, qué decisiones permanecieron bajo control del estudiante y cuánto aprendizaje puede atribuirse razonablemente a su actuación.

La asistencia tecnológica puede considerarse compatible con la integridad académica cuando apoya tareas secundarias sin reemplazar el núcleo intelectual de la actividad. Revisar claridad, explorar ejemplos, ordenar información o recibir sugerencias puede facilitar el trabajo siempre que el estudiante mantenga capacidad para evaluar y transformar los resultados. Savvidou (2026) muestra que los estudiantes distinguen diferentes formas de interacción con herramientas generativas y relacionan esas diferencias con su percepción de autoría. El límite comienza a desplazarse cuando la herramienta deja de apoyar el proceso y comienza a ejecutar aquello que constituye el aprendizaje esperado.

La colaboración con inteligencia artificial representa una categoría más compleja porque supone una interacción en la que estudiante y sistema contribuyen al desarrollo del producto. Zdravkova (2026) señala que las tecnologías generativas introducen ambigüedades importantes en los conceptos de integridad y participación académica. Una colaboración puede resultar legítima cuando el estudiante conserva capacidad para seleccionar, verificar, reformular y justificar los aportes recibidos. Sin embargo, pierde legitimidad cuando la contribución tecnológica domina las decisiones esenciales y el estudiante se limita a aceptar resultados sin ejercer un control intelectual suficientemente visible sobre el contenido final.

Chen et al. (2025) muestran que las percepciones éticas forman parte de la alfabetización en inteligencia artificial y condicionan la manera en que los estudiantes interpretan su utilización. Esta dimensión resulta relevante porque los límites no siempre son evidentes para quien utiliza una herramienta. Un estudiante puede considerar legítimo generar un borrador completo si posteriormente realiza algunos cambios, mientras que el propósito evaluativo puede exigir precisamente demostrar capacidad de escritura autónoma. La delimitación ética necesita, por tanto, relacionarse con los resultados de aprendizaje y no únicamente con la cantidad de contenido modificado posteriormente.

Tabla 18

Continuo ético del uso de inteligencia artificial en actividades académicas

Nivel de intervención	Característica principal	Participación del estudiante	Valoración ética orientativa
Asistencia básica	La IA apoya tareas secundarias	El estudiante conserva las decisiones centrales	Generalmente compatible si está permitida
Asistencia analítica	La IA propone alternativas o explicaciones	El estudiante compara, verifica y selecciona	Compatible cuando existe juicio propio
Colaboración supervisada	La IA contribuye a partes relevantes del trabajo	El estudiante transforma y justifica los aportes	Requiere transparencia y criterios claros
Dependencia elevada	La IA desarrolla una parte sustantiva de la tarea	La participación humana se reduce significativamente	Puede comprometer la validez de la evaluación
Sustitución	La IA ejecuta la competencia que debía demostrar el estudiante	El estudiante principalmente acepta o entrega	Incompatible con la finalidad evaluativa

Nota. La valoración depende del propósito de la actividad, las reglas previamente comunicadas y el grado en que la intervención tecnológica afecta las capacidades que deben ser demostradas por el estudiante.

Temper et al. (2025) destacan la necesidad de establecer marcos institucionales que permitan regular el uso de inteligencia artificial de acuerdo con criterios comprensibles. Estos marcos resultan necesarios porque una misma práctica puede tener implicaciones diferentes según la actividad. Utilizar IA para corregir redacción en un informe técnico puede ser aceptable cuando la competencia evaluada es el análisis, pero podría resultar problemática en una asignatura donde la escritura constituye el desempeño central. El límite ético no puede definirse de manera universal sin considerar qué aprendizaje pretende demostrar cada evaluación.

La sustitución del trabajo académico ocurre cuando la herramienta realiza de manera predominante aquello que debía ejecutar el estudiante para demostrar aprendizaje. Tan y Maravilla (2024) permiten comprender que la presencia de inteligencia artificial no constituye por sí misma el problema; la dificultad aparece cuando reduce o elimina la participación intelectual necesaria para validar la evidencia. Si una actividad pretende valorar razonamiento, argumentación o resolución de problemas y el sistema produce directamente esas respuestas, el producto deja de representar de manera suficiente las capacidades personales que la evaluación buscaba observar.

Savvidou (2026) muestra que la percepción de responsabilidad cambia según cuánto control conserva el estudiante durante la producción asistida. Esta relación permite establecer un criterio práctico: cuanto mayor sea la contribución de la inteligencia artificial, mayor debe ser la capacidad del estudiante para revisar, explicar y asumir responsabilidad sobre aquello que incorpora. La colaboración tecnológica puede mantener legitimidad cuando existe apropiación intelectual verificable. En cambio, si el estudiante no puede explicar las razones que sostienen el contenido generado, la relación con el producto se aproxima más a la sustitución que a una colaboración académicamente significativa.

Los límites éticos también deben considerar la transparencia. Zdravkova (2026) evidencia que la ambigüedad generada por estos sistemas dificulta aplicar categorías tradicionales de integridad, por lo que ocultar una intervención sustantiva puede resultar problemático incluso cuando el producto no contiene plagio convencional. Declarar el uso no convierte automáticamente una práctica en aceptable, pero permite interpretarla según las reglas de la actividad. Una intervención que sustituye la competencia evaluada continúa siendo inadecuada aunque sea declarada; la transparencia ayuda a valorar la situación, pero no reemplaza la necesidad de preservar el aprendizaje esperado.

Chen et al. (2025) permiten relacionar estos límites con la alfabetización ética en inteligencia artificial. Los estudiantes necesitan comprender no solo cómo utilizar una herramienta, sino también cuándo su uso modifica el significado de una evidencia académica. Esta comprensión requiere reconocer que distintas actividades toleran diferentes niveles de asistencia y que la legitimidad depende del propósito formativo. Enseñar a distinguir apoyo, colaboración y sustitución resulta más útil que ofrecer prohibiciones generales, porque proporciona criterios que pueden aplicarse ante nuevas herramientas y situaciones que todavía no han sido reguladas de manera específica.

Los límites éticos entre asistencia, colaboración y sustitución deben construirse, por tanto, a partir de tres preguntas: qué debía demostrar el estudiante, cuánto realizó efectivamente la inteligencia artificial y qué control intelectual conservó la persona sobre el resultado. Temper et al. (2025) muestran la importancia de establecer marcos institucionales claros, mientras que Tan y Maravilla (2024) permiten reconocer que la integración tecnológica puede coexistir con una educación íntegra. Esta delimitación ofrece una base para desarrollar políticas institucionales capaces de orientar el uso

responsable de inteligencia artificial sin depender únicamente de prohibiciones generales.

5.6 Políticas institucionales y criterios para el uso responsable de inteligencia artificial

Las políticas institucionales sobre inteligencia artificial deben establecer un marco común que permita orientar decisiones académicas sin depender exclusivamente de criterios individuales de cada docente. Temper et al. (2025) sostienen que las instituciones de educación superior necesitan desarrollar marcos regulatorios capaces de definir usos aceptables, responsabilidades y mecanismos de supervisión. Esta necesidad responde a la rápida incorporación de sistemas generativos en actividades de aprendizaje y evaluación. Una política pertinente debe ofrecer claridad suficiente para reducir ambigüedades, pero también conservar flexibilidad para adaptarse a disciplinas, niveles formativos y tipos de tareas con exigencias diferentes.

La regulación institucional necesita partir del propósito educativo y no de la herramienta específica utilizada. Una política demasiado vinculada con plataformas concretas corre el riesgo de quedar obsoleta rápidamente, mientras que criterios basados en aprendizaje, responsabilidad y transparencia pueden conservar mayor vigencia. UNESCO (2023) plantea que la gobernanza educativa de la inteligencia artificial debe mantener un enfoque centrado en las personas y en la protección de los procesos formativos. Por ello, las universidades necesitan definir principios generales capaces de orientar el uso de diferentes tecnologías sin depender de sus características comerciales particulares.

Cabrera et al. (2026) muestran que la garantía de integridad académica en la educación superior sudamericana requiere respuestas institucionales capaces de reconocer las particularidades del uso de inteligencia artificial generativa. Esta necesidad incluye establecer diferencias entre actividades donde la tecnología puede utilizarse libremente, aquellas donde su uso debe ser limitado y aquellas en las que determinadas capacidades necesitan demostrarse sin asistencia. Una política clara puede evitar que cada asignatura construya reglas completamente diferentes y facilitar que estudiantes y docentes comprendan de manera consistente qué comportamientos resultan compatibles con los objetivos académicos.

La definición de niveles de uso constituye uno de los criterios más útiles para organizar estas políticas. Una institución puede diferenciar entre usos permitidos, usos permitidos con declaración, usos restringidos y usos incompatibles con determinadas evaluaciones. Esta clasificación permite superar modelos basados únicamente en prohibiciones generales y facilita decisiones adaptadas al tipo de competencia que se evalúa. Cuando las reglas especifican qué funciones puede cumplir la inteligencia artificial dentro de cada actividad, los estudiantes disponen de referencias más claras para actuar y los docentes pueden interpretar con mayor coherencia los productos presentados (Carvajal et al., 2026).

Temper et al. (2025) destacan que cualquier marco institucional necesita definir responsabilidades para estudiantes, docentes y autoridades. El estudiante debe conocer las condiciones de uso y asumir responsabilidad sobre los contenidos presentados; el docente necesita diseñar evaluaciones coherentes con las reglas establecidas; y la institución debe proporcionar formación, orientación y mecanismos de revisión. Esta distribución evita concentrar toda la responsabilidad en una sola parte y reconoce que el uso responsable de inteligencia artificial depende de una cultura compartida. La política adquiere mayor legitimidad cuando las obligaciones se encuentran claramente distribuidas y comunicadas.

La formación constituye otro componente indispensable. Una política que establece restricciones sin desarrollar capacidades puede generar cumplimiento superficial o interpretaciones incorrectas. UNESCO (2023) enfatiza la importancia de fortalecer competencias para utilizar inteligencia artificial de manera crítica, ética y segura. En el ámbito universitario, esto implica ofrecer formación sobre verificación, transparencia, sesgos, privacidad, autoría y límites de asistencia. Las reglas resultan más efectivas cuando quienes deben aplicarlas comprenden por qué existen y cómo se relacionan con la calidad del aprendizaje, en lugar de percibirlas únicamente como mecanismos disciplinarios.

Las políticas institucionales también necesitan incorporar criterios de explicabilidad cuando sistemas automatizados participan en procesos evaluativos. Khosravi et al. (2022) señalan que la inteligencia artificial aplicada a educación debe ofrecer condiciones que permitan comprender cómo se generan determinados resultados o recomendaciones. Si una herramienta interviene en calificación, retroalimentación o análisis de desempeño, docentes y estudiantes deberían conocer sus límites y la forma

en que sus resultados son utilizados. La responsabilidad final sobre decisiones académicas relevantes debe permanecer bajo supervisión humana, especialmente cuando existen consecuencias significativas para la trayectoria estudiantil.

Otro aspecto necesario es la revisión periódica de las políticas. La rápida evolución de los sistemas generativos puede modificar en poco tiempo las posibilidades de uso y las dificultades asociadas con determinadas actividades. Cabrera et al. (2026) permiten comprender que la regulación institucional necesita responder a contextos cambiantes y a prácticas reales de estudiantes y docentes. Una política estática puede volverse insuficiente si no incorpora mecanismos para revisar experiencias, identificar nuevos riesgos y ajustar criterios. Esta actualización debe basarse en evidencia institucional y no únicamente en reacciones frente a tecnologías emergentes.

La coherencia entre política y evaluación resulta igualmente importante. Carvajal et al. (2026) evidencian que las percepciones de docentes y estudiantes pueden diferir en torno a integridad y uso de inteligencia artificial, lo que puede generar tensiones cuando las reglas no son suficientemente claras. Una institución debería procurar que las orientaciones generales puedan traducirse en decisiones concretas dentro de las asignaturas. Esto implica que las consignas evaluativas indiquen cómo se aplica la política en cada caso, evitando que los estudiantes tengan que interpretar normas generales sin conocer su relación con una tarea específica.

Las políticas deben considerar además criterios de equidad, privacidad y acceso. UNESCO (2023) advierte que la incorporación de inteligencia artificial en educación plantea desafíos relacionados con derechos, protección de datos y desigualdades tecnológicas. Una institución no debería exigir determinadas herramientas si parte del estudiantado enfrenta barreras para acceder a ellas, ni debería promover plataformas sin revisar las implicaciones asociadas con información personal o académica. El uso responsable requiere condiciones donde las decisiones tecnológicas no generen desventajas injustificadas ni comprometan la protección de datos dentro de los procesos educativos.

Una política institucional sólida combina principios generales, niveles de uso, transparencia, formación, supervisión humana y mecanismos de actualización. Temper et al. (2025) muestran la importancia de establecer marcos regulatorios coherentes para la educación superior, mientras que Khosravi et al. (2022) subrayan la necesidad de

mantener comprensibles las decisiones mediadas por inteligencia artificial. La finalidad no debería ser controlar cada interacción tecnológica, sino crear condiciones para que docentes y estudiantes actúen con criterios compartidos. Esta base institucional permite avanzar hacia una cultura universitaria donde integridad, responsabilidad y autoría se sostengan de manera consistente.

5.7 Hacia una cultura de integridad, responsabilidad y autoría académica en la universidad

Construir una cultura de integridad académica implica superar una visión centrada exclusivamente en normas, infracciones y sanciones. Alessio y Messinger (2021) muestran que las percepciones sobre integridad pueden variar entre estudiantes y docentes, por lo que la existencia de reglas formales no garantiza por sí sola comportamientos coherentes. En contextos donde la inteligencia artificial generativa participa crecientemente en el trabajo universitario, la cultura institucional necesita apoyarse en valores compartidos, expectativas comprensibles y prácticas educativas que permitan entender por qué la honestidad, la responsabilidad y la autoría continúan siendo esenciales para validar el aprendizaje.

La responsabilidad académica adquiere mayor relevancia cuando los productos pueden construirse mediante diferentes niveles de asistencia tecnológica. Holden et al. (2021) permiten comprender que la integridad de la evaluación depende de condiciones que preserven la confianza en las evidencias presentadas por los estudiantes. Esta responsabilidad no consiste únicamente en evitar conductas prohibidas, sino en asumir las decisiones realizadas durante la producción, revisar la calidad del contenido y responder por aquello que se entrega. Una cultura íntegra necesita que estudiantes y docentes comprendan la responsabilidad como parte del aprendizaje y no como una exigencia externa.

Tan y Maravilla (2024) sostienen que la inteligencia artificial generativa no tiene que convertirse necesariamente en una amenaza para la educación si su incorporación mantiene criterios claros de responsabilidad y participación humana. Esta perspectiva permite orientar la cultura institucional hacia un uso crítico de la tecnología en lugar de construirla únicamente alrededor de la prohibición. La universidad necesita formar estudiantes capaces de reconocer cuándo una herramienta amplía sus posibilidades de aprendizaje y cuándo comienza a sustituir capacidades que deberían demostrar

personalmente, integrando así criterio ético y autonomía dentro de la práctica académica cotidiana.

La autoría académica también necesita consolidarse como una responsabilidad consciente. Navarrete-Cazales et al. (2026) plantean que la inteligencia artificial generativa modifica las fronteras tradicionales entre producción humana y asistencia tecnológica, haciendo necesario revisar cómo se comprende la participación autoral. Una cultura universitaria sólida debe enseñar que ser autor no significa solamente figurar como responsable de un documento, sino poder explicar decisiones, justificar contenidos y asumir consecuencias sobre aquello que se presenta. La autoría se fortalece cuando existe apropiación intelectual y no únicamente atribución formal del producto.

Tabla 19

Componentes de una cultura universitaria de integridad, responsabilidad y autoría académica

Componente	Manifestación institucional	Práctica esperada en estudiantes	Resultado cultural esperado
Integridad	Criterios claros y coherentes	Actuar conforme a reglas conocidas	Mayor confianza académica
Responsabilidad	Expectativas sobre revisión y rendición de cuentas	Responder por decisiones y contenidos	Compromiso con el propio aprendizaje
Autoría	Reconocimiento de contribuciones humanas y tecnológicas	Explicar y sostener el trabajo presentado	Apropiación intelectual
Transparencia	Orientaciones sobre declaración de IA	Informar usos relevantes de herramientas	Menor ambigüedad sobre la producción
Formación	Desarrollo de alfabetización ética y digital	Utilizar IA con criterio	Autonomía responsable
Confianza	Evaluaciones basadas en evidencias diversas	Participar activamente en los procesos	Relaciones académicas menos punitivas

Nota. Los componentes se articulan entre sí y requieren coherencia entre políticas institucionales, prácticas docentes y responsabilidades estudiantiles para consolidar una cultura de integridad sostenible.

Carvajal et al. (2026) evidencian que estudiantes y docentes pueden mantener percepciones diferentes sobre el uso de inteligencia artificial y su relación con la integridad académica. Estas diferencias muestran que una cultura institucional no puede construirse únicamente mediante documentos normativos. Se requiere diálogo, formación y comunicación permanente sobre situaciones concretas que permitan

interpretar los límites de uso. Cuando las expectativas se discuten y se relacionan con los objetivos de aprendizaje, disminuye la posibilidad de que estudiantes y profesores operen con criterios contradictorios frente a una misma práctica tecnológica.

La confianza constituye otro elemento central. Cabrera et al. (2026) permiten situar la integridad dentro de un escenario donde las universidades necesitan responder institucionalmente a nuevas formas de producción académica. Una cultura basada únicamente en sospecha y detección puede deteriorar la relación entre docentes y estudiantes, especialmente cuando cualquier producto bien elaborado es considerado potencialmente artificial. La confianza no implica ausencia de controles, sino construcción de mecanismos transparentes, evidencias diversas y responsabilidades compartidas. Evaluar con claridad permite sostener relaciones académicas donde la supervisión no sustituya el vínculo formativo.

Alessio y Messinger (2021) muestran que las percepciones de integridad pueden fortalecerse cuando existen expectativas coherentes dentro de la comunidad educativa. Esta coherencia requiere que docentes, programas y autoridades eviten mensajes contradictorios sobre el uso de inteligencia artificial. Si una práctica es permitida en una asignatura y sancionada en otra sin explicación pedagógica, los estudiantes pueden interpretar la integridad como un conjunto arbitrario de reglas. Una cultura sostenible necesita principios comunes, pero también capacidad para explicar por qué determinadas actividades requieren diferentes niveles de autonomía o asistencia tecnológica.

La formación continua debe involucrar tanto a estudiantes como a docentes. Tan y Maravilla (2024) permiten comprender que la respuesta educativa frente a la inteligencia artificial puede orientarse hacia el fortalecimiento de capacidades críticas en lugar de limitarse al control. Los estudiantes necesitan aprender a utilizar tecnologías sin delegar indiscriminadamente su trabajo intelectual, mientras los docentes requieren herramientas para diseñar evaluaciones, comunicar expectativas y reconocer diferentes formas de asistencia. Esta formación compartida ayuda a convertir la integridad en una práctica cotidiana vinculada con decisiones reales y no en un contenido aislado de inducción institucional.

Navarrete-Cazales et al. (2026) sitúan la autoría y la integridad dentro de un escenario donde la producción académica necesita ser comprendida desde nuevas relaciones entre

tecnología y responsabilidad humana. Una cultura universitaria madura debe permitir el uso razonado de herramientas emergentes sin renunciar a la exigencia de participación intelectual, transparencia y rendición de cuentas. La finalidad no consiste en preservar prácticas anteriores sin cambios, sino en mantener aquello que da sentido a la evaluación: que el producto presentado represente aprendizaje, decisiones y capacidades que puedan ser atribuidas responsablemente al estudiante.

La consolidación de una cultura de integridad, responsabilidad y autoría requiere coherencia entre evaluación, formación, políticas y prácticas institucionales. Cabrera et al. (2026) muestran la importancia de respuestas articuladas dentro de la educación superior, mientras que Holden et al. (2021) recuerdan que la confianza evaluativa depende de condiciones capaces de proteger la validez de las evidencias. Frente a la inteligencia artificial generativa, la universidad necesita pasar de una lógica centrada en detectar incumplimientos hacia otra orientada a formar estudiantes capaces de utilizar tecnología con criterio, transparencia y responsabilidad sobre su propio aprendizaje.

Conclusión

La expansión de la inteligencia artificial generativa ha colocado a la evaluación universitaria frente a una transformación que no puede resolverse mediante ajustes superficiales. Las prácticas centradas exclusivamente en controlar contenidos, verificar respuestas finales o asumir que todo producto presentado corresponde íntegramente al trabajo individual del estudiante han perdido parte de su capacidad para representar el aprendizaje. Esta situación no significa que la evaluación universitaria haya dejado de ser válida, sino que necesita reorganizar sus propósitos, evidencias y criterios para responder a un entorno donde producir información resulta cada vez más sencillo, pero demostrar comprensión continúa siendo una responsabilidad humana.

Este cambio exige desplazar la atención desde la reproducción de conocimientos hacia la capacidad para utilizarlos, interpretarlos, relacionarlos y defenderlos. La universidad necesita evaluar menos aquello que una herramienta puede generar de manera inmediata y prestar mayor atención a las decisiones que el estudiante adopta frente a problemas, información y escenarios específicos. El valor de una evaluación ya no depende únicamente de la dificultad para obtener una respuesta, sino de su capacidad para mostrar razonamiento, criterio, autonomía y transferencia. Esta transformación convierte la evaluación en un espacio donde aprender y demostrar aprendizaje deben permanecer estrechamente vinculados.

La evaluación auténtica adquiere especial relevancia dentro de este escenario porque permite situar al estudiante frente a tareas que requieren decisiones contextualizadas. Problemas reales, casos, proyectos, productos aplicados y desafíos profesionales ofrecen mayores posibilidades para observar cómo se movilizan conocimientos y competencias. Sin embargo, la autenticidad no depende simplemente de convertir una actividad tradicional en una situación aparentemente práctica. Requiere que la tarea tenga propósito, restricciones, criterios y consecuencias coherentes con aquello que se desea valorar. La inteligencia artificial puede participar en estas actividades, pero su presencia debe enriquecer el razonamiento y no reemplazarlo.

La transformación de la evaluación también implica dejar de considerar el producto final como evidencia suficiente. Un ensayo, informe, presentación o proyecto terminado puede mostrar calidad, pero no necesariamente permite comprender cómo fue construido ni qué aprendizajes intervinieron en su elaboración. Por ello, las

evidencias de proceso adquieren una función estratégica. Borradores, revisiones, bitácoras, portafolios, decisiones justificadas, retroalimentaciones incorporadas y registros de cambios permiten observar la evolución del pensamiento. Este desplazamiento hacia el proceso permite valorar el aprendizaje como construcción progresiva y disminuye la dependencia de una única entrega como representación definitiva del desempeño.

La trazabilidad no debe entenderse como vigilancia permanente del estudiante, sino como posibilidad de reconstruir razonablemente cómo evolucionó su trabajo. En entornos donde la inteligencia artificial puede intervenir en múltiples momentos, resulta más relevante conocer qué decisiones fueron tomadas, qué alternativas se descartaron y cómo se modificaron los productos a partir de nueva información o retroalimentación. Una evaluación basada en evidencias progresivas permite distinguir entre producir un resultado y comprender aquello que se produce. También fortalece la metacognición, porque invita al estudiante a identificar qué aprendió, qué dificultades enfrentó y cómo transformó su desempeño.

Junto con las evidencias de proceso, la recuperación de modalidades orales, prácticas y situadas amplía las posibilidades de comprobar el aprendizaje de manera directa. La capacidad para explicar, argumentar, responder preguntas, ejecutar procedimientos, resolver tareas y actuar frente a situaciones profesionales ofrece información que difícilmente puede obtenerse a través de una única producción escrita. Estas modalidades no deberían utilizarse como mecanismos de sospecha frente a la inteligencia artificial, sino como formas complementarias de evaluación. Su valor consiste en permitir que el estudiante haga visible su comprensión mediante acciones, decisiones y explicaciones desarrolladas frente a condiciones concretas.

La integración de evidencias escritas, orales y prácticas permite construir juicios más equilibrados sobre el desempeño. Cada modalidad muestra dimensiones diferentes y ninguna necesita funcionar como prueba absoluta de aprendizaje. Un texto puede evidenciar organización conceptual; una defensa oral puede mostrar apropiación y capacidad argumentativa; una demostración práctica puede revelar transferencia y adaptación. Cuando estas evidencias convergen, fortalecen la interpretación del aprendizaje; cuando presentan diferencias, permiten identificar necesidades específicas de desarrollo. Esta complementariedad contribuye a una evaluación menos dependiente de formatos únicos y más sensible a la complejidad de las competencias universitarias.

La inteligencia artificial generativa también obliga a revisar las nociones de integridad y autoría académica. El problema ya no puede limitarse a determinar si existe plagio o copia textual. Un contenido puede ser formalmente original y, sin embargo, representar una participación humana insuficiente. Por ello, la integridad necesita incorporar criterios de responsabilidad, transparencia, control intelectual y capacidad para responder por aquello que se presenta. Ser autor de un producto académico supone conservar participación significativa en las decisiones, verificar los contenidos utilizados, justificar las elecciones realizadas y asumir las consecuencias académicas del resultado entregado.

Esta perspectiva requiere diferenciar con claridad entre asistencia tecnológica, colaboración y sustitución. La inteligencia artificial puede apoyar la exploración de ideas, facilitar revisiones, ofrecer alternativas o acompañar determinadas fases del trabajo sin necesariamente anular la autoría del estudiante. La dificultad aparece cuando ejecuta precisamente aquello que la evaluación pretendía observar. Los límites no deberían definirse únicamente por la presencia o ausencia de una herramienta, sino por el propósito de la actividad, el grado de intervención tecnológica y el nivel de control intelectual conservado. Esta diferenciación necesita políticas institucionales claras, coherentes y suficientemente flexibles.

La universidad se encuentra, por tanto, ante una oportunidad para fortalecer la evaluación en lugar de reducirla a una respuesta defensiva frente a nuevas tecnologías. Intentar eliminar la inteligencia artificial generativa de todos los procesos académicos resultaría insuficiente y difícilmente sostenible. El desafío consiste en diseñar evaluaciones donde el aprendizaje humano pueda seguir siendo reconocido mediante decisiones, razonamientos, procesos, explicaciones, actuaciones y responsabilidades verificables. La respuesta universitaria a la inteligencia artificial generativa no consiste en impedir que la tecnología forme parte del aprendizaje, sino en rediseñar la evaluación para que aquello verdaderamente importante vuelva a hacerse visible y demostrable: la capacidad humana para comprender, interpretar, decidir, crear, justificar y responder por el conocimiento que utiliza.

Bibliografía

Alessio, H., y Messinger, J. (2021). Faculty and student perceptions of academic integrity in technology-assisted learning and testing. *Frontiers in Education*, 6. https://doi.org/10.3389/feduc.2021.629220

Alfredo, R., Echeverria, V., Jin, Y., Yan, L., Swiecki, Z., Gašević, D., y Martinez-Maldonado, R. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *Computers and Education: Artificial Intelligence, 6*, 100215. https://doi.org/10.1016/j.caeai.2024.100215

Avalos, A. (2024). Impacto de la inteligencia artificial en la evaluación y retroalimentación educativa. *Revista Retos para la investigación, 3*(1), 19-32. https://doi.org/10.62465/rri.v3n1.2024.72

Banihashem, S., Noroozi, O., van Ginkel, S., Macfadyen, L., y Biemans, H. (2022). A systematic review of the role of learning analytics in enhancing feedback practices in higher education. *Educational Research Review, 37*, 100489. https://doi.org/10.1016/j.edurev.2022.100489

Bearman, M., Tai, J., Dawson, P., Boud, D., y 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

Brogan, M., Gooding, B., y Feld, F. (2024). The viva voce: A post-pandemic assessment. *The Law Teacher, 58*(4), 442-468. https://doi.org/10.1080/03069400.2024.2359865

Cabrera, J., Llanos, R., y Bonilla, L. (2026). Generative AI and the assurance of academic integrity in the context of South American higher education. *Frontiers in Education, 11*, 1796556. https://doi.org/10.3389/feduc.2026.1796556

Carvajal, H., Tusa, F., Samaniego, R., y 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), 67. https://doi.org/10.3390/higheredu5030067

Chen, K., Tallant, A., y Selig, I. (2025). Exploring generative AI literacy in higher education: Student adoption, interaction, evaluation and ethical perceptions.

Information and Learning Sciences, 126(1-2), 132-148. https://doi.org/10.1108/ILS-10-2023-0160

Chhetri Thapa, G., y Lewis, S. (2026). Preserving epistemic authenticity: Process-oriented assessment in the age of generative AI. *Assessment & Evaluation in Higher Education*. https://doi.org/10.1080/02602938.2026.2736679

Cobo Dorado, J. (2021). La bitácora como estrategia de enseñanza e instrumento de aprendizaje. En A. Parra Puentes, y M. Carrillo Salazar (Edits.), *Herramientas pedagógicas: Manual para la creación de textos en el aula* (pp. 163-177). Universidad Santiago de Cali. https://doi.org/10.35985/9786287501591.8

Culebro, K., Gazca, L., Garizurieta, J., y Vargas, I. (2025). Estudio de autoría en la IA generativa en la edición de textos en estudiantes universitarios. *Interconectando Saberes* (19), 79-90. https://doi.org/10.25009/is.v0i19.2934

de Schipper, E., Feskens, R., y Keuning, J. (2021). Personalized and automated feedback in summative assessment using recommender systems. *Frontiers in Education, 6*, 652070. https://doi.org/10.3389/feduc.2021.652070

Edgar, A., Macfarlane, S., Kiddell, E., Armitage, J., y Wood-Bradley, R. (2022). The perceived value and impact of virtual simulation-based education on students' learning: A mixed methods study. *BMC Medical Education, 22*, 823. https://doi.org/10.1186/s12909-022-03912-8

Galindo-Domínguez, H., Delgado, N., Sainz, M., y Expósito, E. (2024). Un análisis experimental de la relación entre las evaluaciones proporcionadas por la inteligencia artificial y las proporcionadas por los docentes en formación. *EDUTEC. Revista Electrónica de Tecnología Educativa*(89), 84-104. https://doi.org/10.21556/edutec.2024.89.3509

García, G. (2024). La evaluación como herramienta para mejorar los aprendizajes: la retroalimentación y la evaluación auténtica. *Revista Latinoamericana Ogmios, 4*(9), 17-32. https://doi.org/10.53595/rlo.v4.i9.091

Gedera, D. (2026). A conceptual framework for authentic AI-enabled assessment (AAA) in higher education. *Future in Educational Research*, 1-11. https://doi.org/10.1002/fer3.70049

Gombert, S., Fink, A., Giorgashvili, T., Jivet, I., Di Mitri, D., Yau, J., . . . Drachsler, H. (2024). From the automated assessment of student essay content to highly informative feedback: A case study. *International Journal of Artificial Intelligence in Education, 34*(4), 1378-1416. https://doi.org/10.1007/s40593-023-00387-6

Gu, P. (2021). An argument-based framework for validating formative assessment in the classroom. *Frontiers in Education, 6*, 605999. https://doi.org/10.3389/feduc.2021.605999

Heil, J., y Ifenthaler, D. (2023). Online assessment in higher education: A systematic review. *Online Learning, 27*(1), 187-218. https://doi.org/10.24059/olj.v27i1.3398

Holden, O., Norris, M., y Kuhlmeier, V. (2021). Academic integrity in online assessment: A research review. *Frontiers in Education, 6*, 639814. https://doi.org/10.3389/feduc.2021.639814

Hornberger, M., Bewersdorff, A., y Nerdel, C. (2023). What do university students know about artificial intelligence? Development and validation of an AI literacy test. *Computers and Education: Artificial Intelligence, 5*, 100165. https://doi.org/10.1016/j.caeai.2023.100165

Huauya, L., Vacas-Gonzales, F., y Solis, B. (2023). Evaluación formativa y desarrollo de competencias en el proceso educativo: Una revisión sistemática. *Horizontes Revista de Investigación en Ciencias de la Educación, 7*(30), 2020-2044. https://doi.org/10.33996/revistahorizontes.v7i30.646

Ibarra, D., Armenta, L., y Cotta, A. (2026). Evolución de la evaluación educativa. *KNOW AND SHARE PSYCHOLOGY, 7*(1), 30-37. https://doi.org/10.25115/kasp.v7i1.10852

Jin, Y., Martinez-Maldonado, R., Gašević, D., y Yan, L. (2025). GLAT: The generative AI literacy assessment test. *Computers and Education: Artificial Intelligence, 9*, 100436. https://doi.org/10.1016/j.caeai.2025.100436

Khosravi, H., Buckingham, S., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., . . . Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and*

Education: Artificial Intelligence, *3*, 100074. https://doi.org/10.1016/j.caeai.2022.100074

Kofinas, A., Tsay, C., y Pike, D. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology*, *56*(6), 2522-2549. https://doi.org/10.1111/bjet.13585

Laamanen, M., Ladonlahti, T., Uotinen, S., Okada, A., Bañeres, D., y Koçdar, S. (2021). Acceptability of the e-authentication in higher education studies: Views of students with special educational needs and disabilities. *International Journal of Educational Technology in Higher Education*, *18*, 4. https://doi.org/10.1186/s41239-020-00236-9

Méndez, R., Donetch, M., Garrido, C., Rocha, H., y Fernández, J. (2024). Impacto de la aplicación del enfoque basado en competencias en el rendimiento académico de los alumnos de una universidad estatal chilena. *International Journal of Professional Business Review*, *9*(3), e04452. https://doi.org/10.26668/businessreview/2024.v9i3.4452

Navarrete-Cazales, Z., Cruz-Victoria, J., y Manzanilla-Granados, H. (2026). Autoría e integridad académica ante la inteligencia artificial generativa en educación superior. *Revista Educación Superior y Sociedad*, *37*(2), 260-283. https://doi.org/10.54674/ess.v37i2.1174

Naveda, J. (2025). Explorando el rol transformador de la inteligencia artificial en la evaluación formativa educativa. *Revista Tribunal*, *5*(10), 676-688. https://doi.org/10.59659/revistatribunal.v5i10.134

Ossa, C., y Willatt, C. (2023). Uso de inteligencia artificial generativa para retroalimentar escritura académica en procesos de formación inicial docente. *European journal of education and psychology*, *16*(2), 1-16. https://doi.org/10.32457/ejep.v16i2.2412

Pearce, J., y Chiavaroli, N. (2023). Rethinking assessment in response to generative artificial intelligence. *Medical Education*, *57*(10), 889-891. https://doi.org/10.1111/medu.15092

Pineda, K., Cortés, L., y Ramírez, A. (2024). Impacto de proyectos integradores en el desarrollo de competencias transversales y evaluación entre pares: Un acercamiento previo a la Nueva Escuela Mexicana. *CIENCIA ergo-sum, 32*. https://doi.org/10.30878/ces.v32n0a6

Romero, M., Salazar, J., Buñay, M., Cabrera, N., y Saltos, K. (2025). Las rúbricas como instrumento predominante en la evaluación formativa: un estudio en educación general básica, Ecuador 2025. *Ciencia Latina Revista Científica Multidisciplinar, 9*(2), 4846-4867. https://doi.org/10.37811/cl_rcm.v9i2.17259

Sánchez, M. (2023). La inteligencia artificial como recurso docente: usos y posibilidades para el profesorado. *Educar, 60*(1), 33-47. https://doi.org/10.5565/rev/educar.1810

Savvidou, C. (2026). Negotiating generative AI-assisted academic writing: Student perspectives on authorship, responsibility and learning. *Education Sciences, 16*(9), 1568. https://doi.org/10.3390/educsci16091568

Sharma, S. (2026). Educational integrity in GenAI-augmented assessment: Making judgement visible. *International Journal for Educational Integrity, 22*, 6. https://doi.org/10.1007/s40979-026-00216-y

Stephenson, Z., Johnson-Glauch, N., y Cruchley, S. (2025). Interventions and facilitators of oral assessment performance in higher education: A systematic review. *Assessment & Evaluation in Higher Education, 50*(7), 1140-1153. https://doi.org/10.1080/02602938.2025.2504621

Tan, M., y Maravilla, N. (2024). Shaping integrity: Why generative artificial intelligence does not have to undermine education. *Frontiers in Artificial Intelligence, 7*, 1471224. https://doi.org/10.3389/frai.2024.1471224

Temper, M., Tjoa, S., y David, L. (2025). Higher Education Act for AI (HEAT-AI): A framework to regulate the usage of AI in higher education institutions. *Frontiers in Education, 10*, 1505370. https://doi.org/10.3389/feduc.2025.1505370

Thanh, B. N., Vo, D., Nhat, M. N., Pham, T., Trung, H. T., y Xuan, S. H. (2023). Race with the machines: Assessing the capability of generative AI in solving authentic assessments. *Australasian Journal of Educational Technology, 39*(5), 59-81. https://doi.org/10.14742/ajet.8902

Tzirides, A., Zapata, G., Kastania, N., Saini, A., Castro, V., Ismael, S., . . . Kalantzis, M. (2024). Combining human and artificial intelligence for enhanced AI literacy in higher education. *Computers and Education Open, 6*, 100184. https://doi.org/10.1016/j.caeo.2024.100184

UNESCO. (2023). *Guía para el uso de IA generativa en educación e investigación.* https://www.unesco.org/es/articles/guia-para-el-uso-de-ia-generativa-en-educacion-e-investigacion

Vera, F. (2023). Integración de la inteligencia artificial en la educación superior: Desafíos y oportunidades. *Revista Científica Transformar*, *4*(1), 17-34. https://www.revistatransformar.cl/index.php/transformar/article/view/84

Villegas, G., y Carballo, E. (2026). Uso de la inteligencia artificial generativa en la evaluación del aprendizaje universitario: Una revisión sistemática de la literatura (2024-2025). *Innovaciones Educativas, 28*(45), 275-293. https://doi.org/10.22458/ie.v28i45.6535

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., y Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications, 252*, 124167. https://doi.org/10.1016/j.eswa.2024.124167

Zdravkova, K. (2026). Generative artificial intelligence and the ambiguity of academic integrity in higher education. *Education Sciences, 16*(7), 1120. https://doi.org/10.3390/educsci16071120

Zhang, P., y Tur, G. (2024). A systematic review of e-portfolio use during the pandemic: Inspiration for post-COVID-19 practices. *Open Praxis, 16*(3), 429-444. https://doi.org/10.55982/openpraxis.16.3.656



ISBN: 978-9907-848-17-5

CIDPROS
EDITORIAL