LA TRANSFORMACIÓN DE LA EVALUACIÓN EN ENTORNOS DIGITALES

Vicente Marlon Villa Villa, Ricardo Geovanni Lalangui Sarango, Franklin Santiago Sosa Cifuentes, Cynthia Shakira Enríquez Fierro & Martha Susana Yauli Flores

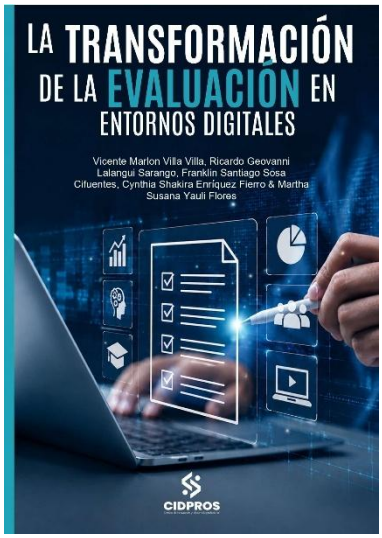


CIDPROS
Centro de innovación y desarrollo profesional

La transformación de la evaluación en entornos digitales

Vicente Marlon Villa Villa, Ricardo Geovanni Lalangui Sarango, Franklin Santiago Sosa Cifuentes, Cynthia Shakira Enríquez Fierro & Martha Susana Yauli Flores

CIDPROS
Centro de innovación y desarrollo profesional



Datos bibliográficos

ISBN	978-9907-9574-1-9
Título del libro	La transformación de la evaluación en entornos digitales
Autores	Vicente Marlon Villa Villa Ricardo Geovanni Lalangui Sarango Franklin Santiago Sosa Cifuentes Cynthia Shakira Enríquez Fierro Martha Susana Yauli Flores
Editorial	CIDPROS EDITORIAL
Materia	370 - Educación
Público objetivo	Profesional / académico
Año	2026
Número de edición	1
Tamaño	2.1Mb
Soporte	Libro digital descargable
Formato	PDF (.pdf)
Idioma	Español
DOI	https://doi.org/10.67166/0xwxwy14

Hecho en Ecuador / Made in Ecuador

Dr. Vicente Marlon Villa Villa, Ph.D.

Universidad Nacional de Chimborazo
mvilla@unach.edu.ec
https://orcid.org/0000-0002-4292-2391
Riobamba, Chimborazo, Ecuador

Semblanza



Ph.D. Vicente Marlon Villa Villa es docente e investigador de la Universidad Nacional de Chimborazo (Ecuador), donde se desempeña como Profesor Titular Principal a Tiempo Completo en la Facultad de Ciencias Políticas y Administrativas. Posee una sólida formación académica que incluye un Doctorado (Ph.D.) en Educación por la Universidad Nacional Mayor de San Marcos (Perú), un Doctorado en Ciencias de la Educación, una Maestría en Matemática Básica, una Maestría en Docencia Universitaria e Investigación Educativa, además de estudios de Postdoctorado en Estrategias Didácticas para la Investigación y la distinción de Doctor Honoris Causa.

Cuenta con más de veintiocho años de experiencia en la educación superior, desarrollando actividades de docencia, investigación, gestión universitaria y vinculación con la sociedad. Ha contribuido significativamente al fortalecimiento institucional de la Universidad Nacional de Chimborazo, participando en la creación y consolidación de programas académicos en las áreas de Contabilidad y Auditoría, Economía e Ingeniería Comercial, así como en diversos procesos de dirección y administración universitaria.

Su producción científica comprende libros, capítulos de libro y artículos publicados en revistas científicas indexadas nacionales e internacionales, abordando temáticas relacionadas con matemática aplicada, estadística, investigación científica, inteligencia artificial, contabilidad, finanzas, tributación, innovación educativa y transformación digital. Sus investigaciones integran metodologías cuantitativas y herramientas tecnológicas para generar soluciones a problemas académicos y organizacionales, con especial énfasis en la toma de decisiones basada en evidencia.

Ha participado como evaluador de revistas científicas, miembro de comités científicos, director y tribunal de trabajos de investigación de posgrado, además de impartir docencia en programas de maestría. Su trayectoria se caracteriza por el compromiso con la excelencia académica, el liderazgo institucional, la innovación pedagógica y la formación de profesionales con pensamiento crítico y capacidad investigativa. Actualmente, además de su labor docente e investigadora, desempeña funciones de representación gremial y previsional universitaria, contribuyendo al fortalecimiento de la comunidad académica ecuatoriana.

Dr. Ricardo Geovanni Lalangui Sarango, MSc.

Universidad Bolivariana del Ecuador (UBE)
rilasdie33@gmail.com
https://orcid.org/0000-0002-7614-5682
Machala, El Oro, Ecuador

Semblanza



Dr. Ricardo Geovanni Lalangui Sarango, docente investigador por 35 años, he sido partícipe de la formación de inumerables conglomerados de estudiantes de Educación Báisca Superior, bachillerato, pregrado y posgrado en distintas universidades nacionales y extranjeras, originario de Zamora Chinchipe, cuento con seis titulos de posgrado 5 nacionales y uno extranjero, dos titulos de pregrado: Licenciado y profesor en Ciencias de la Educación, mención Químico Biológicas, otrogado por la Universidad Nacional de Loja. Actualmente doctorante del Doctorado de Ciencias de la Educación ofertado por la Universidad de Santander de Máxico, en este contexto se consolida un perfil profesional forjado en el límite superior de mejoramiento de la calidad educativa, la docencia, la investigación y vinculación con la comunidad.

Actualmente formo parte de planta de docentes virtuales de la Universidad Bolivariana del Ecuador, institución líder en la formación profesionalizante en el contexto nacional e internacional; además, cumplo las funciones de Cordinador Académico virtual de Redeemed International Christian University (IRCU) del State of Florida que oferta formación de pregrado y posgrado en los países de habla hispana. Este desarrollo epistémico avala mi trayectoria: Director de investigación Educativa del LICJAM, docente de Ciencias Naturales, Biología, Anatomía, Química en Básica Superior y Bachillerato, a nivel universitario he impartido múltiples módulos relacionados con la didáctica, pedagógía, interdisciplinariedad, curriculo, filosofía, investigación, asesoría de tesis de pregrado y posgrado, tutor de artículos científicos, he cumplido las funciones de director de área, rector y vicerrector en unidades educativas, coordinador académico de pregrado y posgrado en universidades; además de facilitador de cursos del Ministerio de Educación. Autor y coautor de 2 libros, dos capítulos de libros y de más de 10 artículos científicos indexada a Latindex catálogo 2.0 y Scopus Q1, esto respalda la sólida producción intelectual orientada a forjar la transformación educativa.

Experticia académica ampliamente reconocida por Unidades Educativas, universidades nacionales e internacionales, participación en varios eventos científicos y académicos, relacionado con: Paradigmas, modelos, tipos y diseño de investigación de cuantitativa y cualitativa; Metodología para la enseñanza y aprendizaje cooperativa y colaborativa; Desarrollo de estruturas mentales congruentes con el desarrollo del pensamiento científico, crítico y creativo; Praxis pedagógica: Teoría de la equilibración de Jean Piaget, Teoría sociocultural de L.S. Vigotsky, Teoría del descubrimiento de Jerome Bruner, Teoría del aprendizaje por descubrimiento David Ausubel.

Mgs. Franklin Santiago Sosa Cifuentes

Investigador Independiente
franklinsosac_@hotmail.com
https://orcid.org/0009-0009-3959-3206
Riobamba, Chimborazo, Ecuador

Semblanza



Franklin Santiago Sosa Cifuentes es un Psicólogo Educativo, de nacionalidad ecuatoriana nacido el 09 de junio de 1994 en la ciudad de Riobamba, cuya vida ha estado marcado por su perseverancia, en un entorno urbano que fortaleció su sensibilidad intra e interpersonal enfocando en contribuir positivamente a las personas y la comunidad para el servicio del bien común. Desde temprana edad desarrolló una profunda empatía, respeto y tolerancia por el comportamiento humano, las emociones y los procesos mentales de los seres humanos.

Cuenta con una sólida formación académica: es Licenciado en Psicología Educativa, Orientador Vocacional y Familiar desde 2017, además obtuvo el título de Magíster en Intervención Psicológica en el Desarrollo y la Educación, el cual le permite generar espacios de apoyo psicopedagógico en las distintas etapas del desarrollo humano, así como estoy plenamente capacitado en las distintas especialidades de problemas de aprendizaje. Además, es especializado en las áreas de actividades de capacitación y administración educativa.

A lo largo de su trayectoria, ha destacado tanto en el ejercicio libre de la profesión dado que, cuenta con el centro de capacitaciones virtuales conocido como REPSICO-ECUADOR, el cual se desempeña como propietario y presidente, permitiendo ejecutar capacitaciones a nivel nacional e internacional, brindado ponencias en la Universidad de los Hemisferios y al Instituto Superior Tecnológico José Ortega y Gasset. Finalmente, fue miembro de la Federación de Estudiantes Universitarios del Ecuador para la Directiva de la Asociación Estudiantil de la Carrera de Psicología Educativa en la Universidad Nacional de Chimborazo de la ciudad de Riobamba.

PhD. Cynthia Shakira Enríquez Fierro

Universidad Internacional del Ecuador – UIDE
cyenriquezfi@uide.edu.ec
https://orcid.org/0009-0002-5389-9892
Quito – Pichincha - Ecuador

Semblanza



Cynthia Shakira Enríquez Fierro es docente universitaria, investigadora y estratega en comunicación, marketing y publicidad, con experiencia académica y profesional en diversas universidades e instituciones de educación superior. Su trayectoria se enfoca en áreas como comunicación digital, branding, reputación corporativa, comportamiento del consumidor, storytelling, publicidad estratégica y análisis de medios digitales. Ha participado en proyectos académicos, científicos y de formación vinculados a investigación, innovación y desarrollo de estrategias comunicacionales aplicadas a contextos contemporáneos. Su trabajo integra enfoques críticos, tecnológicos y creativos, especialmente en temas relacionados con transformación digital, eWOM, activismo de marca, medios digitales y comunicación estratégica, aportando tanto en docencia como en dirección y acompañamiento de proyectos académicos y científicos.

MSc. Martha Susana Yauli Flores

Unida Educativa Los Andes Distrito 18D05. Zona 3 MINEDUC-EC. Píllaro – Ecuador
susyflor2484@gmail.com
https://orcid.org/0009-0003-4282-7049
Ambato, Tungurahua, Ecuador

Semblanza



Martha Susana Yauli Flores, es una jurista y docente ecuatoriana nacida el 24 de marzo de 1984 en la ciudad de Ambato, su trayectoria profesional ha estado marcada por una firme vocación de servicio, iniciando su carrera en el ejercicio del Derecho, lo que le permitió conocer de cerca las distintas realidades sociales. Posteriormente, encontró en la docencia una nueva forma de servir a la sociedad, trasladando su experiencia jurídica al ámbito educativo. Como docente, asumió la responsabilidad de formar ciudadanos íntegros, críticos y comprometidos con los valores de la justicia, la solidaridad y la responsabilidad social y promoviendo el pensamiento crítico, analítico y el respeto por los derechos humanos.

Cuenta con una sólida formación académica: es Técnica Ejecutiva Asistente Jurídico (2005), Tecnóloga Ayudante Judicial (2007) y Abogada de los Tribunales y Juzgados del Ecuador desde 2008 por la Universidad Regional Autónoma de los Andes (UNIANDES) . Posteriormente obtuvo el título de Magíster en Educación en la Universidad Tecnológica Indoamérica (UTI) Ecuador en el 2020.

A lo largo de su trayectoria, ha integrado el ejercicio del Derecho con la docencia, promoviendo una formación basada en la ética, el pensamiento crítico y el respeto a los derechos humanos. Su vocación educativa se evidencia en artículos como: “La enseñanza de la Historia y Ciencias Sociales: un análisis de las metodologías y recursos utilizados en las aulas”; “La integración de la tecnología en la Educación Básica: un análisis de su impacto en el aprendizaje”; “Ciencias Sociales y el desarrollo del pensamiento crítico en los estudiantes de básica superior en Ecuador” y Uso de Tecnologías Digitales en la enseñanza de Ciencias Sociales en el nivel de Bachillerato en los colegios del Ecuador”. Madre y profesional comprometida con una educación dedicada a la formación de jovenes íntegros, promoviendo el conocimiento, la ética y la responsabilidad social como pilares para la construcción de una sociedad más justa y solidaria.



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).

Todos los derechos reservados © 2026

Sinopsis

La transformación de la evaluación en entornos digitales analiza los cambios conceptuales, metodológicos y tecnológicos que han reconfigurado la forma de valorar el aprendizaje. La obra examina el tránsito desde modelos tradicionales, centrados en pruebas y calificaciones, hacia enfoques más continuos, formativos, auténticos y sustentados en evidencias diversas. A lo largo de cinco capítulos se estudian los fundamentos de la evaluación educativa, las herramientas tecnológicas disponibles, la retroalimentación en espacios virtuales, la incorporación de la inteligencia artificial y la analítica del aprendizaje, así como los principales desafíos éticos contemporáneos.

El libro aborda plataformas virtuales, cuestionarios, rúbricas digitales, portafolios electrónicos, gamificación, simuladores, sistemas automatizados de calificación, retroalimentación adaptativa y modelos de predicción del desempeño. También examina la privacidad, la protección de datos, los sesgos algorítmicos, la integridad académica, la brecha digital, la transparencia y la gobernanza de las decisiones evaluativas.

Desde una perspectiva crítica, se sostiene que la innovación tecnológica no garantiza por sí misma calidad educativa. Su valor depende de la coherencia pedagógica, la validez de las evidencias, la accesibilidad, la equidad y la supervisión humana. La obra propone comprender la evaluación digital como un proceso orientado a mejorar el aprendizaje, fortalecer la responsabilidad institucional y preservar los derechos de los estudiantes.

Palabras clave: evaluación digital; inteligencia artificial; retroalimentación formativa; analítica del aprendizaje; ética educativa.

Synopsis

The Transformation of Assessment in Digital Environments analyzes the conceptual, methodological, and technological changes that have reshaped the ways in which learning is assessed. The book examines the transition from traditional models centered on tests and grades toward more continuous, formative, authentic approaches supported by diverse forms of evidence. Across five chapters, it explores the foundations of educational assessment, available technological tools, feedback in virtual environments, the incorporation of artificial intelligence and learning analytics, and the main contemporary ethical challenges.

The book addresses virtual platforms, questionnaires, digital rubrics, electronic portfolios, gamification, simulators, automated grading systems, adaptive feedback, and performance prediction models. It also examines privacy, data protection, algorithmic bias, academic integrity, the digital divide, transparency, and the governance of assessment decisions.

From a critical perspective, the book argues that technological innovation does not guarantee educational quality by itself. Its value depends on pedagogical coherence, the validity of evidence, accessibility, equity, and human oversight. The work proposes understanding digital assessment as a process aimed at improving learning, strengthening institutional accountability, and safeguarding students' rights.

Keywords: digital assessment; artificial intelligence; formative feedback; learning analytics; educational ethics.

Índice de Contenido

Sinopsis 10

Synopsis 11

Introducción 14

Capítulo 01 Evolución de la evaluación educativa en contextos digitales 17

1.1 Fundamentos y propósitos de la evaluación educativa 18

1.2 De la evaluación tradicional a la evaluación digital 20

1.3 Transformaciones de los modelos, técnicas e instrumentos evaluativos 24

1.4 Evaluación del aprendizaje en modalidades virtuales e híbridas 27

1.5 Desarrollo de competencias digitales para la evaluación educativa 30

1.6 Nuevas tendencias y perspectivas de la evaluación en entornos digitales 33

Capítulo 02 Herramientas tecnológicas para la evaluación del aprendizaje 37

2.1 Plataformas virtuales y sistemas de gestión del aprendizaje 38

2.2 Cuestionarios, pruebas y exámenes digitales 41

2.3 Rúbricas digitales, listas de cotejo y escalas de valoración 43

2.4 Portafolios electrónicos y evidencias digitales de aprendizaje 46

2.5 Herramientas interactivas, gamificación y evaluación mediante juegos 49

2.6 Evaluación audiovisual, simuladores y laboratorios virtuales 52

2.7 Criterios para la selección e implementación de herramientas evaluativas 55

Capítulo 03 Evaluación formativa y retroalimentación en entornos virtuales 59

3.1 Fundamentos de la evaluación formativa digital 60

3.2 Diagnóstico y seguimiento continuo del aprendizaje 62

3.3 Retroalimentación efectiva, oportuna y personalizada 65

3.4 Autoevaluación y desarrollo de la autorregulación 68

3.5 Coevaluación y aprendizaje colaborativo en espacios virtuales 71

3.6 Evaluación auténtica mediante proyectos y resolución de problemas 74

3.7 Integración de la evaluación formativa en la planificación docente 77

Capítulo 04 Inteligencia artificial, analítica de datos y automatización evaluativa 81

4.1 Fundamentos de la inteligencia artificial aplicada a la evaluación 82
4.2 Analítica del aprendizaje y seguimiento de trayectorias estudiantiles 84
4.3 Sistemas automatizados de calificación y corrección 87
4.4 Retroalimentación adaptativa mediante sistemas inteligentes 90
4.5 Evaluación personalizada y aprendizaje adaptativo 93
4.6 Predicción del desempeño y detección temprana de dificultades 96
4.7 Alcances, limitaciones y supervisión humana de la evaluación automatizada 99
Capítulo 05 Retos éticos y desafíos de la evaluación digital contemporánea 103
5.1 Privacidad, protección de datos y consentimiento informado 104
5.2 Sesgos algorítmicos, equidad e inclusión educativa 106
5.3 Integridad académica, plagio y uso responsable de la inteligencia artificial 109
5.4 Brecha digital y desigualdades en el acceso a la evaluación 113
5.5 Transparencia, explicabilidad y responsabilidad en las decisiones evaluativas 115
5.6 Gobernanza y perspectivas futuras de la evaluación digital 118
Bibliografía 122

Introducción

La evaluación educativa constituye un componente esencial de los procesos de enseñanza y aprendizaje, debido a que permite recopilar, interpretar y comunicar evidencias sobre el desarrollo académico. Su función no se limita a asignar calificaciones, sino que comprende el diagnóstico, el seguimiento, la retroalimentación y la toma de decisiones pedagógicas. En este sentido, evaluar supone establecer relaciones coherentes entre objetivos, actividades, criterios y resultados. García (2024) sostiene que la evaluación adquiere verdadero valor cuando se orienta hacia la mejora del aprendizaje y la comprensión de los desempeños estudiantiles.

La expansión de las tecnologías digitales ha transformado progresivamente las condiciones en que se desarrollan las prácticas evaluativas. Plataformas virtuales, cuestionarios automatizados, rúbricas digitales, portafolios electrónicos, simulaciones y sistemas inteligentes han ampliado las posibilidades para recopilar evidencias en distintos momentos y formatos. Heil y Ifenthaler (2023) señalan que la evaluación en línea integra métodos, instrumentos y modalidades diversas. Sin embargo, la incorporación tecnológica no garantiza por sí misma calidad pedagógica, pues sus resultados dependen de la pertinencia de las tareas, los criterios empleados y las interpretaciones realizadas.

La transformación digital de la evaluación no consiste únicamente en sustituir instrumentos impresos por aplicaciones electrónicas. Implica modificar la temporalidad, la interacción, la producción de evidencias y la comunicación de resultados. Las plataformas permiten registrar intentos, entregas, revisiones, tiempos y formas de participación, lo que favorece una observación más continua del aprendizaje. No obstante, Muñoz et al. (2021) advierten que los registros digitales requieren interpretación contextual, porque la frecuencia de acceso, la permanencia en línea o la cantidad de interacciones no representan automáticamente comprensión, compromiso o desempeño académico.

Uno de los principales aportes de los entornos digitales se relaciona con el fortalecimiento de la evaluación formativa. La posibilidad de recoger evidencias de manera continua facilita identificar dificultades, ofrecer apoyos y ajustar la enseñanza antes de finalizar el proceso. Huauya et al. (2023) vinculan esta orientación con el desarrollo de competencias y con una participación más activa del estudiante. La retroalimentación adquiere aquí una función central, siempre que sea clara, oportuna y vinculada con acciones concretas de mejora, en lugar de limitarse a comunicar errores, puntuaciones o juicios generales sobre el desempeño.

Las tecnologías también han diversificado las evidencias mediante proyectos multimedia, portafolios, producciones audiovisuales, simulaciones y actividades colaborativas. Estas alternativas permiten valorar procesos, decisiones, argumentos y aplicaciones del conocimiento que difícilmente pueden observarse mediante pruebas cerradas. Zhang y Tur (2024) destacan que los portafolios electrónicos favorecen el seguimiento de producciones y reflexiones desarrolladas a lo largo del tiempo. Sin embargo, la variedad instrumental exige criterios claros, porque la calidad técnica, el dominio de una plataforma o la presentación visual no deben confundirse con el aprendizaje disciplinar evaluado.

La inteligencia artificial y la analítica del aprendizaje representan una etapa avanzada de esta transformación. Los sistemas automatizados pueden clasificar respuestas, asignar puntuaciones, generar retroalimentación, recomendar actividades y detectar patrones relacionados con el desempeño. Wang et al. (2024) identifican la evaluación como uno de los campos de mayor crecimiento dentro de la inteligencia artificial educativa. A pesar de estas posibilidades, los resultados automatizados dependen de datos, modelos y criterios previamente definidos. Por ello, la precisión técnica debe acompañarse de validación pedagógica, explicabilidad y supervisión humana permanente.

La evaluación digital también introduce riesgos vinculados con privacidad, protección de datos, integridad académica y sesgos algorítmicos. Las plataformas pueden recopilar información sobre respuestas, calificaciones, comportamientos, tiempos e interacciones, generando perfiles detallados de las trayectorias estudiantiles. Korir et al. (2023) muestran que los estudiantes expresan preocupaciones sobre la recopilación, utilización y compartición de estos datos. De manera complementaria, Baker y Hawn (2022) advierten que los sistemas algorítmicos pueden reproducir desigualdades presentes en los datos, los criterios o las condiciones institucionales de implementación.

La equidad constituye una condición indispensable para interpretar correctamente los resultados de la evaluación digital. Las diferencias en conectividad, dispositivos, competencias tecnológicas, accesibilidad y condiciones domésticas pueden afectar la participación sin reflejar necesariamente el nivel real de aprendizaje. Mathrani et al. (2022) explican que la brecha digital comprende dimensiones sociales, educativas e institucionales, especialmente relevantes en países en desarrollo. Por ello, la evaluación debe incorporar alternativas razonables, formatos accesibles, flexibilidad ante dificultades técnicas y criterios capaces de distinguir entre limitaciones de acceso y desempeño académico.

Este libro analiza la transformación de la evaluación en entornos digitales mediante cinco ejes articulados. El primero examina su evolución conceptual y metodológica; el segundo aborda las herramientas tecnológicas empleadas; el tercero profundiza en la evaluación formativa y la retroalimentación; el cuarto analiza la inteligencia artificial, la analítica y la automatización; y el quinto estudia los desafíos éticos, sociales y de gobernanza. En conjunto, la obra sostiene que la innovación evaluativa debe fortalecer la validez, la equidad, la accesibilidad y el aprendizaje, manteniendo la responsabilidad pedagógica en manos humanas.

Capítulo

01

Evolución de la evaluación educativa en contextos digitales



1.1 Fundamentos y propósitos de la evaluación educativa

La evaluación educativa constituye un proceso sistemático de interpretación de evidencias orientado a comprender el aprendizaje y fundamentar decisiones pedagógicas. Su propósito no se limita a asignar calificaciones, pues también permite identificar avances, dificultades, necesidades de apoyo y condiciones que influyen en el desempeño. García (2024) sostiene que la evaluación adquiere valor cuando se vincula con la mejora del aprendizaje y con formas auténticas de valoración. Bajo esta lógica, evaluar supone relacionar objetivos, criterios, evidencias y decisiones dentro de una secuencia coherente, integral y formativa.

Los fundamentos de la evaluación se apoyan en la correspondencia entre aquello que se pretende enseñar, las tareas propuestas y los criterios utilizados para valorar los resultados. Una evaluación pertinente debe representar de manera suficiente el aprendizaje esperado y evitar que la medición se reduzca a respuestas aisladas o descontextualizadas. Gu (2021) plantea que la validez de la evaluación formativa depende de argumentos que conecten las evidencias obtenidas con las interpretaciones pedagógicas. Por ello, la calidad evaluativa exige justificar tanto los instrumentos empleados como las decisiones derivadas de sus resultados.

El propósito diagnóstico permite reconocer los conocimientos previos, las concepciones iniciales y las dificultades que pueden condicionar nuevos aprendizajes. Esta función favorece una planificación ajustada a las características reales del grupo, siempre que los resultados no se utilicen para etiquetar de manera rígida al estudiantado. La evaluación diagnóstica adquiere sentido cuando orienta apoyos, secuencias y estrategias diferenciadas, en lugar de convertirse en una clasificación temprana. En términos pedagógicos, su utilidad reside en anticipar necesidades y reducir la distancia entre las exigencias curriculares y las posibilidades de participación.

La función formativa desplaza el centro de la evaluación desde la certificación hacia el acompañamiento continuo del aprendizaje. Huauya et al. (2023) relacionan la evaluación formativa con el desarrollo de competencias, debido a que permite ajustar la enseñanza mientras el proceso todavía está en curso. Esta orientación exige recoger evidencias frecuentes, interpretarlas con criterios explícitos y comunicar información útil para mejorar. No basta con aumentar el número de actividades evaluativas; resulta necesario que cada evidencia contribuya a comprender cómo aprende el estudiante y qué acciones pueden fortalecer su progreso.

La retroalimentación representa uno de los mecanismos centrales mediante los cuales la evaluación cumple una función pedagógica. Vera (2022) la concibe como una herramienta

efectiva cuando ofrece información comprensible, oportuna y vinculada con posibilidades concretas de mejora. Su eficacia disminuye si se limita a señalar errores, emitir juicios generales o proporcionar respuestas sin explicación. Una retroalimentación rigurosa debe mostrar la relación entre el desempeño alcanzado, los criterios establecidos y los pasos siguientes. Así, el estudiante puede reconocer brechas, revisar estrategias y asumir un papel más activo en su aprendizaje.

La evaluación auténtica amplía la valoración hacia tareas que demandan aplicar conocimientos en situaciones significativas, resolver problemas, argumentar decisiones o producir resultados con sentido contextual. García (2024) vincula este enfoque con la mejora del aprendizaje porque permite observar desempeños complejos que difícilmente se expresan mediante pruebas centradas únicamente en la reproducción de información. Su propósito consiste en aproximar la evaluación a las formas reales de uso del conocimiento. No obstante, su implementación requiere criterios claros, evidencias comparables y condiciones que garanticen coherencia entre la tarea, la competencia y la valoración.

La función sumativa mantiene relevancia porque certifica logros, informa resultados finales y sustenta decisiones de promoción o acreditación. Sin embargo, su predominio puede reducir la evaluación a un momento terminal desvinculado del proceso de aprendizaje. Ibarra et al. (2026) muestran que la evaluación educativa ha evolucionado desde modelos centrados en la medición hacia concepciones más amplias, orientadas a interpretar procesos y resultados. En consecuencia, la evaluación sumativa resulta más sólida cuando se integra con evidencias acumuladas y no cuando depende exclusivamente de una prueba única o de un desempeño aislado.

La equidad constituye un fundamento transversal, dado que toda práctica evaluativa distribuye oportunidades, reconoce capacidades y puede producir consecuencias académicas diferenciadas. Evaluar de manera justa no significa aplicar idénticas condiciones sin considerar barreras, sino garantizar que los instrumentos permitan demostrar los aprendizajes relevantes. Las directrices del CAST (2024) proponen múltiples formas de participación, representación y expresión, principios que fortalecen la accesibilidad de la evaluación. Esta apertura no elimina los criterios académicos; procura que las diferencias en el acceso, la comunicación o la interacción no distorsionen indebidamente el desempeño.

La participación del estudiante redefine los propósitos evaluativos al incorporar autoevaluación, coevaluación y comprensión de los criterios. Estas prácticas favorecen la

autorregulación porque obligan a interpretar expectativas, valorar evidencias y justificar decisiones sobre la calidad del propio trabajo. Romero et al. (2025) destacan el uso de rúbricas en la evaluación formativa, debido a su capacidad para hacer visibles dimensiones y niveles de desempeño. No obstante, una rúbrica solo resulta formativa cuando se analiza, se discute y se utiliza para revisar producciones, no cuando funciona como una lista cerrada de verificación.

En contextos digitales, los fundamentos de la evaluación permanecen vinculados con la validez, la pertinencia y la mejora del aprendizaje, aunque cambien los medios de recopilación y análisis de evidencias. Heil y Ifenthaler (2023) muestran que la evaluación en línea incorpora modalidades, instrumentos y condiciones que requieren decisiones pedagógicas. La tecnología puede ampliar la frecuencia, trazabilidad y variedad de los datos, pero no garantiza por sí una evaluación de calidad. El propósito educativo debe orientar la selección de plataformas, tareas y criterios, y no quedar subordinado a sus posibilidades técnicas.

La integridad académica forma parte de los propósitos evaluativos porque la evidencia solo es interpretable cuando existe confianza razonable sobre su autoría y las condiciones en que fue producida. Holden et al. (2021) señalan que la evaluación en línea plantea desafíos particulares relacionados con autenticidad, supervisión y diseño de tareas. Responder a estos problemas únicamente mediante control puede afectar la confianza y la accesibilidad. Una estrategia más consistente combina criterios transparentes, actividades significativas, acompañamiento docente y formas de evaluación que reduzcan la utilidad de respuestas mecánicas o ajenas.

En conjunto, la evaluación educativa cumple funciones diagnósticas, formativas, sumativas, reguladoras y certificadoras que deben articularse en lugar de operar como prácticas desconectadas. Su fundamento central radica en producir evidencias válidas para comprender el aprendizaje y orientar decisiones justificadas. La calidad no depende del número de instrumentos ni del uso de tecnologías, sino de la coherencia entre propósitos, criterios, tareas, retroalimentación y condiciones de participación. Concebida de este modo, la evaluación deja de ser un acto terminal y se integra como componente permanente de la enseñanza y del desarrollo educativo.

1.2 De la evaluación tradicional a la evaluación digital

La evaluación tradicional se configuró durante décadas alrededor de pruebas estandarizadas, calificaciones numéricas y momentos específicos de comprobación del aprendizaje. Este modelo privilegió la medición de resultados observables y la comparación del rendimiento, con una participación limitada del estudiante en la interpretación de sus propios avances. Ibarra et

al. (2026) explican que la evolución de la evaluación educativa ha implicado un tránsito desde enfoques centrados en la medición hacia concepciones más amplias. Dicho cambio responde a la necesidad de comprender procesos, competencias y condiciones de aprendizaje.

El paso hacia formas digitales de evaluación no debe entenderse únicamente como la sustitución del papel por una pantalla. La transformación alcanza el diseño de actividades, la recopilación de evidencias, la comunicación de resultados y la posibilidad de ofrecer seguimiento continuo. Heil y Ifenthaler (2023) identifican en la evaluación en línea una diversidad de enfoques, métodos e instrumentos que modifican las prácticas convencionales. No obstante, la incorporación tecnológica conserva problemas pedagógicos previos cuando mantiene tareas memorísticas, criterios poco claros o decisiones basadas exclusivamente en una calificación final.

Las primeras experiencias de digitalización reprodujeron con frecuencia la estructura de los exámenes tradicionales mediante cuestionarios automatizados y respuestas cerradas. Aunque estas herramientas facilitaron la administración, corrección y almacenamiento de resultados, su aporte fue limitado cuando no modificaron el propósito evaluativo. Topuz et al. (2022) observaron que los sistemas de evaluación en línea adquirieron especial relevancia durante la enseñanza remota de emergencia. Este crecimiento evidenció que la disponibilidad tecnológica podía sostener la continuidad educativa, pero también que la innovación metodológica no se producía automáticamente por emplear plataformas digitales.

A diferencia de la evaluación convencional, los entornos digitales permiten registrar múltiples interacciones asociadas con tiempos de respuesta, secuencias de trabajo, revisiones y participación. Tales datos amplían las posibilidades de interpretar el aprendizaje más allá del producto final, siempre que exista una relación clara entre los registros y los objetivos educativos. Muñoz et al. (2021) señalan que la analítica del aprendizaje ofrece herramientas para examinar información generada durante la actividad académica. Su valor, sin embargo, depende de evitar que la acumulación de datos sustituya el juicio pedagógico o simplifique procesos complejos.

La retroalimentación constituye uno de los ámbitos donde la transición digital presenta mayores posibilidades. Los sistemas tecnológicos pueden ofrecer respuestas inmediatas, personalizadas o adaptadas al tipo de error, lo que reduce el tiempo entre la actuación del estudiante y la orientación recibida. Banihashem et al. (2022) muestran que la analítica del aprendizaje puede fortalecer las prácticas de retroalimentación en educación superior. Esta

capacidad no elimina la intervención docente, pues los comentarios automatizados necesitan criterios adecuados, explicaciones comprensibles y oportunidades reales de revisión para producir mejoras significativas en el desempeño.

También se amplía el repertorio de evidencias mediante portafolios electrónicos, simulaciones, proyectos multimedia y tareas desarrolladas en entornos interactivos. Sillat et al. (2021) destacan la diversidad de métodos utilizados para evaluar competencias digitales en educación superior, lo cual demuestra que ciertos aprendizajes requieren instrumentos distintos de las pruebas tradicionales. La evaluación digital puede representar procesos, decisiones y producciones complejas con mayor amplitud. Sin embargo, esa variedad exige definir criterios consistentes que permitan valorar la calidad de los desempeños sin confundir dominio tecnológico con dominio conceptual o disciplinar.

La evolución hacia la evaluación digital modifica, además, la temporalidad del proceso. En lugar de concentrar la valoración en un examen terminal, las plataformas permiten recopilar evidencias sucesivas y observar cambios a lo largo del tiempo. Liao et al. (2022) subrayan la necesidad de mantener y supervisar la calidad en evaluaciones digitales administradas de manera continua. Esta vigilancia resulta esencial porque las condiciones técnicas, los bancos de preguntas y los patrones de respuesta pueden variar. La continuidad evaluativa requiere controles de calidad, revisión periódica y criterios estables de interpretación.

Tabla 1
Transformaciones centrales entre la evaluación tradicional y la evaluación digital

Dimensión	Predominio tradicional	Transformación digital	Implicación pedagógica
Temporalidad	Aplicación en momentos finales	Recopilación continua de evidencias	Permite ajustes durante el aprendizaje
Tipo de evidencia	Respuestas escritas y productos terminados	Registros de interacción, procesos y producciones multimodales	Amplía la comprensión del desempeño
Retroalimentación	General y posterior a la calificación	Inmediata, automatizada o personalizada	Favorece revisión y autorregulación
Participación estudiantil	Recepción pasiva del resultado	Seguimiento del progreso y uso de criterios visibles	Incrementa la responsabilidad sobre el aprendizaje

Dimensión	Predominio tradicional	Transformación digital	Implicación pedagógica
Gestión de resultados	Registro manual y fragmentado	Almacenamiento, trazabilidad y análisis integrado	Facilita decisiones fundamentadas en múltiples evidencias
Riesgos predominantes	Subjetividad y limitación instrumental	Privacidad, brecha digital, sesgos y dependencia tecnológica	Exige gobernanza, accesibilidad y supervisión humana

Nota. La tabla contrasta cambios estructurales derivados de la digitalización y sus principales implicaciones para la práctica evaluativa, con base en las fuentes utilizadas.

La comparación muestra que la evaluación digital no constituye un modelo opuesto a la evaluación tradicional, sino una reorganización de sus procedimientos, tiempos y fuentes de evidencia. Sus ventajas aparecen cuando la tecnología fortalece la función formativa y permite interpretar trayectorias de aprendizaje con mayor detalle. Tsai et al. (2021) advierten que la adopción de analítica depende de factores institucionales, experiencias previas y prioridades educativas. Por tanto, la transformación requiere capacidades docentes, infraestructura adecuada y acuerdos sobre el uso pedagógico de la información generada.

El tránsito digital también introduce desigualdades que pueden afectar la validez de los resultados. Las diferencias en conectividad, dispositivos, habilidades tecnológicas y espacios de estudio condicionan la participación y pueden confundirse con diferencias de aprendizaje. Mathrani et al. (2022) analizan la brecha digital en países en desarrollo y muestran que el acceso tecnológico forma parte de un problema más amplio de condiciones educativas. Van de Werfhorst et al. (2022) añaden que la preparación digital desigual de estudiantes e instituciones influye en las oportunidades de educación en línea.

La integridad académica adquiere nuevas características cuando la evaluación se realiza a distancia, utiliza recursos digitales o incorpora herramientas automatizadas. Holden et al. (2021) identifican desafíos relacionados con la autenticidad de las respuestas y las condiciones de realización de las pruebas en línea. Una respuesta exclusivamente basada en vigilancia puede generar tensiones con la privacidad, la confianza y la inclusión. Por ello, la transición digital debe acompañarse de tareas auténticas, criterios transparentes y diseños que valoren razonamiento, aplicación y elaboración personal, reduciendo la dependencia de controles invasivos.

La evaluación digital representa una etapa de transformación pedagógica cuyo alcance depende de la coherencia entre tecnología, propósitos y decisiones educativas. No toda práctica digital

supera las limitaciones del enfoque tradicional, del mismo modo que una evaluación presencial no carece necesariamente de valor formativo. La diferencia sustantiva reside en cómo se seleccionan las evidencias, se interpretan los datos y se comunica la retroalimentación. La evolución más sólida combina eficiencia técnica, calidad pedagógica, accesibilidad e integridad, manteniendo el juicio docente como componente indispensable de toda decisión evaluativa.

1.3 Transformaciones de los modelos, técnicas e instrumentos evaluativos

La transformación de la evaluación educativa ha implicado una revisión progresiva de sus modelos, técnicas e instrumentos. El enfoque centrado exclusivamente en medir resultados ha cedido espacio a propuestas orientadas a interpretar procesos, desempeños y trayectorias de aprendizaje. Ibarra et al. (2026) sitúan esta evolución dentro de un cambio conceptual más amplio, en el que evaluar deja de equivaler únicamente a comprobar conocimientos adquiridos. Esta reconfiguración modifica la función del docente, amplía la participación estudiantil y exige evidencias capaces de representar aprendizajes complejos con mayor pertinencia.

Los modelos tradicionales se apoyaron principalmente en pruebas escritas, escalas de calificación y procedimientos estandarizados. Su fortaleza residía en la facilidad para administrar resultados comparables, aunque tendían a privilegiar respuestas cerradas y productos finales. La evaluación formativa introdujo una lógica diferente, al utilizar la evidencia para ajustar la enseñanza y acompañar el progreso. Huauya et al. (2023) relacionan este enfoque con el desarrollo de competencias, ya que permite identificar dificultades, orientar mejoras y vincular la valoración con decisiones pedagógicas tomadas durante el proceso educativo.

La evaluación auténtica representó otra transformación relevante al desplazar la atención hacia tareas contextualizadas, resolución de problemas y aplicación del conocimiento. García (2024) sostiene que este enfoque favorece aprendizajes más significativos porque exige demostrar comprensión mediante actuaciones vinculadas con situaciones reales o plausibles. Su adopción ha impulsado técnicas como proyectos, estudios de caso, presentaciones, portafolios y producciones integradoras. Estas alternativas no eliminan las pruebas convencionales, pero cuestionan su suficiencia cuando se pretende valorar razonamiento, transferencia, creatividad, argumentación o capacidad para tomar decisiones fundamentadas.

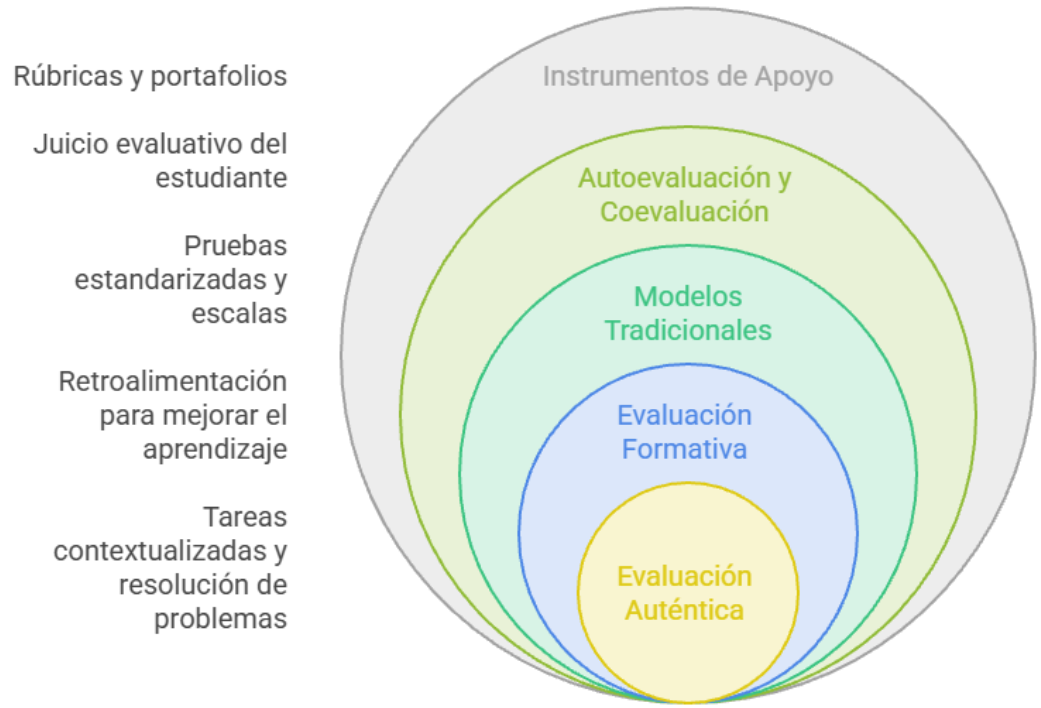
La ampliación de los modelos evaluativos también ha fortalecido la autoevaluación y la coevaluación. Ambas técnicas trasladan parte del juicio evaluativo al estudiante, quien debe

analizar criterios, identificar fortalezas y justificar valoraciones sobre su propio trabajo o el de sus pares. Santana et al. (2025) vinculan la evaluación entre pares con el desarrollo de competencias transversales dentro de proyectos integradores. No obstante, estas prácticas requieren orientación, criterios explícitos y acompañamiento docente, pues la participación estudiantil no garantiza por sí sola valoraciones consistentes, equitativas o pedagógicamente útiles.

Entre los instrumentos que han adquirido mayor presencia se encuentran las rúbricas, debido a su capacidad para organizar criterios, niveles de desempeño y expectativas de calidad. Romero et al. (2025) destacan su predominio dentro de la evaluación formativa en educación general básica. Su utilidad se incrementa cuando permiten comprender qué se espera, comparar evidencias y orientar revisiones. Sin embargo, una rúbrica excesivamente rígida puede fragmentar actuaciones complejas o convertir la evaluación en una verificación mecánica, especialmente si sus descriptores no representan adecuadamente la naturaleza del aprendizaje valorado.

Los portafolios han transformado la manera de reunir y analizar evidencias al integrar productos, borradores, reflexiones y muestras de progreso. En formato digital, esta técnica amplía las posibilidades de incorporar recursos audiovisuales, hipervínculos y registros desarrollados en distintos momentos. Zhang y Tur (2024) muestran que el uso de portafolios electrónicos adquirió relevancia durante la pandemia y ofrece aprendizajes para prácticas posteriores. Su valor radica menos en acumular archivos que en seleccionar evidencias, justificar su pertinencia y relacionarlas con criterios claros de desarrollo académico.

Figura 1
Evolución de la evaluación educativa



Nota. La figura sintetiza la evolución de la evaluación educativa desde modelos tradicionales hacia enfoques formativos, auténticos y participativos.

Las simulaciones y los entornos virtuales han introducido instrumentos capaces de representar situaciones difíciles de reproducir mediante pruebas convencionales. Edgar et al. (2022) encontraron que la educación basada en simulación virtual puede generar valor percibido e impacto sobre el aprendizaje. Wu et al. (2022) también identifican una expansión de estas prácticas en la formación médica. Desde el punto de vista evaluativo, las simulaciones permiten observar decisiones, secuencias de actuación y respuestas frente a escenarios complejos, aunque requieren criterios válidos para distinguir el dominio conceptual de la familiaridad tecnológica.

La evaluación automatizada constituye una de las transformaciones más visibles de los instrumentos digitales. Los sistemas de puntuación pueden procesar respuestas extensas, reconocer patrones y generar resultados con mayor rapidez. Maestrales et al. (2021) utilizaron aprendizaje automático para calificar evaluaciones multidimensionales en química y física, mientras Haudek y Zhai (2024) examinaron cómo las características del constructo influyen en la puntuación automatizada de argumentaciones científicas. Estos avances muestran posibilidades relevantes, pero también evidencian que el rendimiento técnico depende de la definición precisa del aprendizaje que se desea valorar.

La retroalimentación automatizada ha ampliado la función de los instrumentos al convertir la evaluación en una intervención inmediata sobre el aprendizaje. Lee et al. (2021) analizaron un sistema capaz de apoyar la revisión de argumentos científicos mediante comentarios generados a partir de datos de simulación. De forma complementaria, Wilson et al. (2021) estudiaron el uso de puntuación y retroalimentación automatizadas en escritura escolar. Estas herramientas pueden aumentar la frecuencia del acompañamiento, aunque su aporte depende de la claridad de los mensajes y de las oportunidades ofrecidas para corregir, revisar y volver a producir.

La analítica del aprendizaje ha incorporado técnicas orientadas a interpretar grandes volúmenes de registros generados en plataformas educativas. González y Lugo (2021) muestran que su utilización puede fortalecer la práctica docente al proporcionar información sobre participación y desempeño. Muñoz et al. (2021) advierten, sin embargo, que la analítica debe entenderse como una herramienta de apoyo y no como sustituto de la interpretación pedagógica. Los indicadores digitales adquieren sentido cuando se relacionan con objetivos, contextos y evidencias cualitativas; aislados, pueden producir lecturas reduccionistas sobre procesos educativos complejos.

En conjunto, las transformaciones de los modelos, técnicas e instrumentos han ampliado el alcance de la evaluación y diversificado las formas de producir evidencia. La evolución no consiste en abandonar completamente los recursos tradicionales, sino en seleccionar procedimientos coherentes con los aprendizajes que se pretende comprender. Pruebas, rúbricas, portafolios, simulaciones, analítica y sistemas automatizados ofrecen posibilidades distintas y también limitaciones específicas. La calidad evaluativa depende de articular estas opciones con criterios válidos, propósitos formativos, accesibilidad, supervisión docente y oportunidades reales para que los resultados mejoren el aprendizaje.

1.4 Evaluación del aprendizaje en modalidades virtuales e híbridas

La evaluación del aprendizaje en modalidades virtuales e híbridas exige reconsiderar la relación entre presencia, participación y evidencia académica. La distancia física modifica las formas de observar el desempeño, acompañar los procesos y verificar la autoría de las producciones. Heil y Ifenthaler (2023) muestran que la evaluación en línea integra métodos diversos, desde pruebas automatizadas hasta actividades colaborativas y tareas auténticas. Esta variedad amplía las posibilidades de valoración, pero también demanda criterios pedagógicos capaces de distinguir entre participación tecnológica, cumplimiento de actividades y aprendizaje efectivamente alcanzado.

En los entornos virtuales, la evaluación puede desarrollarse de manera sincrónica o asincrónica, lo que transforma la temporalidad de la recogida de evidencias. Las actividades sincrónicas permiten observar respuestas inmediatas, interacción oral y resolución de problemas en tiempo real, mientras las asincrónicas favorecen reflexión, revisión y producción progresiva. La combinación de ambas modalidades contribuye a representar mejor los aprendizajes, siempre que las tareas respondan a propósitos diferenciados y no se dupliquen innecesariamente (Topuz et al., 2022). La calidad depende, por tanto, del diseño y no del formato utilizado.

Las modalidades híbridas plantean una dificultad adicional: integrar evidencias obtenidas en espacios presenciales y digitales sin tratarlas como componentes separados. Tsai et al. (2021) señalan que la adopción de analítica del aprendizaje depende de prioridades institucionales, experiencias y capacidades organizativas. En consecuencia, la evaluación híbrida requiere acuerdos sobre qué información se recoge, cómo se interpreta y de qué manera se articula con la observación docente. La fragmentación entre plataformas, clases presenciales y actividades autónomas puede producir valoraciones parciales si no existe una planificación común.

La retroalimentación adquiere una función decisiva cuando el contacto presencial es limitado o discontinuo. En estos contextos, el estudiante necesita comprender con claridad qué logró, qué dificultades presenta y qué acciones puede realizar para mejorar. La analítica puede apoyar este proceso mediante información sobre avances, errores y patrones de participación, aunque su eficacia depende de la calidad de los mensajes y de las oportunidades de revisión (Banihashem et al., 2022). Una retroalimentación rápida pero superficial no necesariamente orienta el aprendizaje ni fortalece la autonomía.

La evaluación en línea también favorece el uso de portafolios electrónicos, proyectos multimedia y actividades desarrolladas en varias etapas. Zhang y Tur (2024) destacan que los portafolios digitales permiten reunir producciones, reflexiones y evidencias acumuladas durante el proceso formativo. Su aplicación en modalidades virtuales e híbridas facilita el seguimiento longitudinal y la integración de formatos diversos. Sin embargo, su valor no reside en almacenar archivos, sino en seleccionar muestras pertinentes, justificar su relación con los criterios y demostrar transformaciones en la calidad del trabajo realizado.

La participación estudiantil puede observarse mediante registros de acceso, tiempo de conexión, intervenciones y entregas, pero estos indicadores no deben confundirse automáticamente con aprendizaje. La analítica ofrece información útil sobre comportamientos dentro de las plataformas, aunque los datos requieren interpretación contextual y contraste con

evidencias de desempeño (Muñoz et al., 2021). Un estudiante puede permanecer conectado sin participar de manera significativa, mientras otro puede desarrollar aprendizajes relevantes mediante actividades realizadas fuera del sistema. La evaluación debe evitar conclusiones simplificadas basadas exclusivamente en métricas de actividad.

Las diferencias de conectividad, disponibilidad de dispositivos y competencias digitales condicionan la validez de la evaluación virtual. Mathrani et al. (2022) muestran que la brecha digital en países en desarrollo no se limita al acceso técnico, sino que incluye condiciones familiares, institucionales y educativas. Cuando estas desigualdades no se consideran, los resultados pueden reflejar limitaciones de acceso más que niveles reales de aprendizaje. Por ello, la evaluación debe ofrecer alternativas razonables de participación, tiempos flexibles y formatos accesibles, sin reducir las exigencias académicas esenciales.

La accesibilidad constituye otro criterio central en modalidades virtuales e híbridas, porque las plataformas y recursos pueden generar barreras para estudiantes con distintas necesidades. Los errores de diseño web, la falta de compatibilidad con tecnologías de apoyo y el uso exclusivo de determinados formatos restringen la participación (Scanlon et al., 2021). Las directrices del CAST (2024) proponen múltiples formas de implicación, representación y expresión, lo que permite diversificar las evidencias sin perder coherencia. Evaluar de manera inclusiva supone mantener criterios comunes y ampliar las vías para demostrarlos.

La integridad académica representa uno de los desafíos más discutidos en la evaluación a distancia. Holden et al. (2021) identifican problemas relacionados con autenticidad, supervisión y diseño de las pruebas en línea. Sin embargo, la vigilancia intensiva no constituye la única respuesta posible y puede entrar en tensión con la privacidad o la confianza. Las tareas contextualizadas, las defensas orales, los procesos de revisión y las producciones progresivas dificultan la reproducción mecánica de respuestas. Una evaluación sólida combina prevención, acompañamiento y criterios transparentes, en lugar de depender únicamente del control tecnológico.

La evaluación en modalidades virtuales e híbridas requiere articular flexibilidad, validez, accesibilidad, retroalimentación e integridad dentro de un mismo diseño. Su eficacia no depende de trasladar instrumentos presenciales a una plataforma, sino de seleccionar evidencias coherentes con los objetivos y las condiciones de aprendizaje. Las tecnologías pueden ampliar el seguimiento, diversificar tareas y mejorar la comunicación, pero también introducir desigualdades y nuevas formas de exclusión. El juicio docente continúa siendo

indispensable para interpretar datos, contextualizar resultados y tomar decisiones pedagógicas justificadas.

1.5 Desarrollo de competencias digitales para la evaluación educativa

El desarrollo de competencias digitales para la evaluación educativa implica mucho más que manejar plataformas, cargar calificaciones o aplicar cuestionarios en línea. Supone comprender cómo seleccionar herramientas, diseñar evidencias válidas, interpretar datos y comunicar resultados de manera pedagógicamente pertinente. Sillat et al. (2021) muestran que la evaluación de la competencia digital en educación superior comprende métodos diversos y dimensiones que exceden el dominio instrumental. Por ello, la formación evaluativa debe integrar conocimientos tecnológicos, pedagógicos, éticos y analíticos dentro de decisiones coherentes con los propósitos de aprendizaje.

La primera dimensión corresponde a la capacidad de diseñar actividades digitales alineadas con objetivos, criterios y resultados esperados. Una herramienta resulta pertinente cuando permite representar el aprendizaje que se desea valorar, no únicamente porque automatiza tareas o facilita su administración. La competencia docente exige distinguir entre posibilidades técnicas y utilidad pedagógica, especialmente cuando una plataforma condiciona el tipo de respuesta disponible. La adopción tecnológica debe responder a prioridades educativas claras y a condiciones institucionales que permitan sostener su uso responsable (Tsai et al., 2021).

También se requiere competencia para seleccionar instrumentos adecuados entre pruebas en línea, rúbricas digitales, portafolios, simulaciones, foros y producciones multimedia. Cada recurso genera evidencias diferentes y presenta limitaciones específicas. Zhang y Tur (2024) destacan que los portafolios electrónicos permiten documentar procesos, reflexiones y productos acumulados, mientras las simulaciones favorecen la observación de decisiones en escenarios complejos. Elegir un instrumento exige valorar su correspondencia con la competencia evaluada, las condiciones de aplicación, la accesibilidad y la posibilidad de obtener información útil para orientar mejoras posteriores.

La interpretación de datos constituye otra capacidad central, debido a que los entornos digitales producen registros continuos sobre accesos, entregas, tiempos, errores e interacciones. Estos datos pueden apoyar la identificación de patrones, pero no representan por sí solos comprensión, esfuerzo o logro académico. La analítica del aprendizaje requiere relacionar indicadores cuantitativos con evidencias cualitativas y conocimiento del contexto educativo (Muñoz et al., 2021). Una lectura competente evita convertir métricas aisladas en juicios

definitivos y reconoce que toda interpretación depende de criterios, supuestos y decisiones pedagógicas previas.

La retroalimentación digital demanda habilidades para transformar resultados en orientaciones comprensibles y accionables. Banihashem et al. (2022) señalan que la analítica puede fortalecer las prácticas de retroalimentación cuando ayuda a ofrecer información oportuna y vinculada con el progreso. Sin embargo, la rapidez no garantiza calidad. El docente debe valorar qué información necesita el estudiante, cómo presentarla y qué oportunidad tendrá para revisar su trabajo. Esta competencia incluye combinar comentarios automáticos con intervención humana, evitando mensajes genéricos que describan errores sin explicar vías concretas de mejora.

La accesibilidad debe incorporarse como parte de la competencia digital evaluativa y no como una adaptación posterior. Los recursos pueden excluir cuando dependen de un único formato, presentan problemas de navegación o no funcionan con tecnologías de apoyo. Las barreras de diseño web afectan directamente las posibilidades de participación de las personas con discapacidad (Scanlon et al., 2021). Las directrices del CAST (2024) respaldan la diversificación de formas de representación, participación y expresión, de modo que el estudiante pueda demostrar el aprendizaje sin que condiciones técnicas irrelevantes distorsionen su desempeño.

La competencia ética comprende el tratamiento responsable de datos, la protección de la privacidad y la revisión crítica de sistemas automatizados. Korir et al. (2023) evidencian que las preocupaciones estudiantiles abarcan la recopilación, utilización y compartición de información en procesos de analítica. Por tanto, no basta con conocer cómo funciona una plataforma; es necesario determinar qué datos son pertinentes, quién puede acceder a ellos y durante cuánto tiempo deben conservarse. La transparencia sobre estas prácticas fortalece la confianza y limita usos que excedan los propósitos educativos establecidos.

Tabla 2
Dimensiones de la competencia digital aplicadas a la evaluación educativa

Dimensión	Capacidad requerida	Aplicación evaluativa	Riesgo ante un desarrollo insuficiente
Pedagógica	Alinear objetivos, evidencias y criterios	Diseño de tareas coherentes con el aprendizaje esperado	Uso tecnológico sin valor formativo

Dimensión	Capacidad requerida	Aplicación evaluativa	Riesgo ante un desarrollo insuficiente
Instrumental	Seleccionar y gestionar herramientas digitales	Aplicación de pruebas, rúbricas, portafolios y simulaciones	Dependencia de funciones predeterminadas
Analítica	Interpretar registros y resultados contextualizados	Identificación de avances, dificultades y patrones	Confusión entre actividad digital y aprendizaje
Comunicativa	Proporcionar retroalimentación clara y oportuna	Orientación de revisiones y decisiones de mejora	Comentarios automáticos poco comprensibles
Inclusiva	Garantizar accesibilidad y alternativas de participación	Diversificación de formatos y medios de respuesta	Exclusión por barreras tecnológicas
Ética	Proteger datos y explicar su utilización	Gestión responsable de información académica	Vulneración de privacidad y pérdida de confianza
Crítica	Examinar sesgos y límites de la automatización	Supervisión de puntuaciones y recomendaciones	Aceptación acrítica de decisiones algorítmicas

Nota. La tabla organiza las capacidades que intervienen en una práctica evaluativa digital integral y los principales efectos derivados de su desarrollo insuficiente, con base en la bibliografía empleada.

La tabla evidencia que la competencia digital evaluativa no puede reducirse al aprendizaje de aplicaciones específicas. Sus dimensiones se encuentran conectadas: una herramienta técnicamente eficiente puede generar resultados pedagógicamente débiles, éticamente problemáticos o poco accesibles. Stewart et al. (2023) muestran que las creencias y comprensiones del profesorado influyen en la manera de interpretar la datificación educativa. En consecuencia, la formación docente debe incluir espacios para discutir supuestos, límites y consecuencias de las tecnologías, además de desarrollar destrezas operativas para utilizarlas.

La incorporación de inteligencia artificial añade la necesidad de comprender cómo se producen recomendaciones, puntuaciones y clasificaciones automatizadas. La explicabilidad permite examinar las razones asociadas con una decisión y valorar si los resultados son consistentes con los criterios educativos (Khosravi et al., 2022). Esta competencia crítica resulta indispensable porque los sistemas pueden reproducir sesgos, simplificar constructos o generar interpretaciones difíciles de cuestionar. La UNESCO (2023) subraya la importancia de utilizar la inteligencia artificial generativa con supervisión humana, responsabilidad y claridad sobre sus limitaciones educativas.

Las desigualdades de acceso también condicionan el desarrollo de estas competencias. La preparación digital varía entre estudiantes, docentes e instituciones, de modo que una misma práctica evaluativa puede producir oportunidades distintas de participación (Van de Werfhorst et al., 2022). Mathrani et al. (2022) muestran que la brecha digital incluye factores tecnológicos, sociales e institucionales, especialmente relevantes en países en desarrollo. Fortalecer competencias requiere, por tanto, formación continua, infraestructura, acompañamiento y alternativas accesibles; de lo contrario, la digitalización puede ampliar diferencias que la evaluación debería identificar y reducir.

El desarrollo de competencias digitales para evaluar constituye un proceso continuo que integra diseño pedagógico, selección instrumental, interpretación de datos, comunicación, inclusión y responsabilidad ética. Ninguna de estas capacidades resulta suficiente de forma aislada, pues la calidad depende de su articulación dentro de decisiones contextualizadas. La competencia evaluativa se demuestra cuando la tecnología mejora la producción y comprensión de evidencias sin desplazar el juicio profesional. Su finalidad no consiste en automatizar la evaluación, sino en utilizar los recursos digitales para fortalecer la validez, la equidad y el aprendizaje.

1.6 Nuevas tendencias y perspectivas de la evaluación en entornos digitales

Las nuevas tendencias de la evaluación digital se caracterizan por una transición desde instrumentos aislados hacia ecosistemas capaces de recopilar, integrar e interpretar evidencias de manera continua. Esta evolución combina analítica del aprendizaje, inteligencia artificial, retroalimentación automatizada y entornos adaptativos. Wang et al. (2024) identifican una expansión sostenida de aplicaciones de inteligencia artificial en educación, entre ellas la evaluación y el seguimiento del rendimiento. El cambio no radica únicamente en la sofisticación tecnológica, sino en la posibilidad de personalizar decisiones y responder con mayor rapidez a necesidades específicas.

La evaluación adaptativa constituye una de las perspectivas con mayor proyección, porque ajusta tareas, niveles de dificultad y rutas de aprendizaje según el desempeño registrado. Los sistemas adaptativos pueden ofrecer experiencias diferenciadas y reducir la aplicación uniforme de actividades que no responden a las necesidades del estudiante (Aparicio y Aparicio, 2024). No obstante, su calidad depende de la precisión con que se definan los criterios de adaptación. Una personalización mal diseñada puede limitar oportunidades, reforzar trayectorias predeterminadas o confundir rapidez de respuesta con comprensión profunda.

La analítica del aprendizaje seguirá ocupando un lugar central en la transformación de la evaluación. Su aporte consiste en convertir registros dispersos en información útil para reconocer patrones, anticipar dificultades y orientar intervenciones. Alfredo et al. (2024) señalan que los enfoques centrados en las personas buscan equilibrar las posibilidades de la analítica y la inteligencia artificial con las necesidades humanas. Esta orientación resulta decisiva, pues la evaluación no puede reducir al estudiante a un conjunto de indicadores ni sustituir la interpretación pedagógica por relaciones estadísticas automatizadas.

La automatización de la retroalimentación representa otra tendencia relevante, especialmente en tareas que requieren revisión frecuente y respuestas oportunas. Los sistemas pueden identificar errores, sugerir mejoras y ofrecer comentarios personalizados durante el proceso de aprendizaje (de Schipper et al., 2021). Gombert et al. (2024) muestran que la evaluación automatizada de ensayos puede evolucionar hacia formas de retroalimentación más informativas. Aun así, la eficacia de estos recursos depende de la calidad de los criterios, la claridad de los mensajes y la posibilidad de complementar la respuesta automática con orientación docente.

La puntuación automatizada también continuará ampliándose en áreas donde existen grandes volúmenes de respuestas y necesidad de resultados rápidos. Zhai et al. (2021) advierten que la validez de las evaluaciones basadas en aprendizaje automático debe analizarse mediante redes de inferencias que conecten respuestas, puntuaciones e interpretaciones. Esta precaución evita asumir que un alto nivel de precisión técnica equivale a una medición educativa válida. Los modelos automatizados necesitan supervisión constante, revisión del constructor y controles capaces de detectar cambios en los patrones de respuesta.

Las evaluaciones multimodales constituyen una perspectiva emergente porque integran texto, audio, video, simulaciones, interacciones y producciones digitales. Este enfoque permite observar aprendizajes que no se expresan plenamente mediante respuestas escritas convencionales. La simulación virtual, por ejemplo, facilita el análisis de decisiones, procedimientos y reacciones ante escenarios complejos (Edgar et al., 2022; Wu et al., 2022). Su expansión plantea retos metodológicos relacionados con la comparabilidad, la definición de criterios y la necesidad de distinguir entre desempeño disciplinar, dominio tecnológico y familiaridad con el entorno utilizado.

La evaluación continua apoyada en plataformas digitales modificará la relación entre seguimiento y certificación. Liao et al. (2022) sostienen que las evaluaciones administradas de

forma permanente requieren procedimientos específicos para supervisar su calidad. Esta tendencia permite reunir evidencias sucesivas, detectar cambios y evitar que una única prueba determine todo el resultado académico. Sin embargo, la continuidad puede convertirse en vigilancia permanente si no se delimitan los propósitos de la recopilación. La frecuencia de los datos debe responder a necesidades pedagógicas y no a la capacidad técnica de almacenarlos.

La privacidad y la protección de datos se consolidan como condiciones esenciales para el desarrollo de futuras prácticas evaluativas. Los estudiantes expresan preocupaciones sobre la recopilación, el uso y la compartición de información académica generada mediante analítica (Korir et al., 2023). Estas inquietudes presentan variaciones culturales, lo que demuestra que la aceptación tecnológica no es uniforme entre contextos (Viberg et al., 2024). La evaluación digital deberá establecer principios de minimización de datos, transparencia, consentimiento informado y acceso restringido, conforme con las obligaciones de protección aplicables en Ecuador.

La explicabilidad de los sistemas automatizados será determinante para sostener la confianza en decisiones que afectan calificaciones, recomendaciones y trayectorias educativas. Khosravi et al. (2022) destacan que la inteligencia artificial explicable permite examinar cómo se generan determinados resultados y facilita su revisión. Sin explicaciones comprensibles, docentes y estudiantes pueden quedar subordinados a puntuaciones difíciles de cuestionar. La transparencia no exige revelar todos los componentes técnicos, pero sí comunicar los criterios relevantes, las limitaciones del sistema y las vías disponibles para solicitar revisión humana.

La inclusión y la accesibilidad también definirán el valor de las tendencias emergentes. La personalización tecnológica carece de sentido cuando las plataformas reproducen barreras de navegación, comunicación o participación. Los problemas de accesibilidad web pueden excluir a personas con discapacidad incluso cuando los contenidos académicos son adecuados (Scanlon et al., 2021). Las orientaciones del CAST (2024) promueven múltiples formas de implicación, representación y expresión. Aplicadas a la evaluación, estas directrices permiten diversificar los medios para demostrar aprendizaje sin modificar injustificadamente los criterios esenciales de desempeño.

El uso de inteligencia artificial generativa plantea nuevas posibilidades para formular tareas, producir retroalimentación y apoyar procesos de escritura. Ossa y Willatt (2023) examinan su empleo para retroalimentar escritura académica en la formación inicial docente. Al mismo tiempo, su incorporación exige revisar las nociones de autoría, originalidad e integridad. La

inteligencia artificial no tiene que debilitar necesariamente la educación si se establecen criterios claros sobre usos permitidos, atribución y responsabilidad (Tan y Maravilla, 2024). Las futuras evaluaciones deberán valorar procesos y justificar decisiones, no solo productos finales.

Las perspectivas de la evaluación digital apuntan hacia sistemas más adaptativos, continuos, multimodales e integrados con inteligencia artificial. Su desarrollo, sin embargo, no garantiza por sí mismo mayor calidad ni equidad. La UNESCO (2023) plantea que el uso educativo de la inteligencia artificial debe conservar supervisión humana y atención a sus riesgos. Las tendencias más prometedoras serán aquellas que combinen innovación, validez, accesibilidad, privacidad y explicabilidad. El futuro de la evaluación dependerá menos de automatizar decisiones que de construir criterios responsables para utilizar, interpretar y revisar las tecnologías disponibles.

Capítulo

02

Herramientas tecnológicas para la evaluación del aprendizaje



2.1 Plataformas virtuales y sistemas de gestión del aprendizaje

Las plataformas virtuales y los sistemas de gestión del aprendizaje constituyen infraestructuras digitales destinadas a organizar contenidos, actividades, interacciones y evidencias educativas dentro de un entorno integrado. Su relevancia para la evaluación radica en que permiten reunir información sobre participación, desempeño, entregas y progresos en distintos momentos del proceso formativo. Heil y Ifenthaler (2023) muestran que la evaluación en línea comprende múltiples métodos y condiciones de aplicación. Por ello, estas plataformas no deben entenderse únicamente como repositorios, sino como espacios donde se articulan enseñanza, aprendizaje y valoración.

La función evaluativa de estos sistemas depende de la coherencia entre los objetivos de aprendizaje, las actividades configuradas y los criterios utilizados para interpretar resultados. Una plataforma puede administrar cuestionarios, tareas, foros o rúbricas, pero la disponibilidad de estas funciones no garantiza una evaluación pertinente. La calidad surge cuando el recurso seleccionado representa de manera adecuada el conocimiento o la competencia esperada (Sillat et al., 2021). De lo contrario, la estructura predeterminada del sistema puede condicionar el diseño y favorecer actividades fáciles de automatizar, aunque pedagógicamente limitadas.

Los sistemas de gestión del aprendizaje facilitan la evaluación continua al conservar registros de entregas, revisiones, calificaciones y comentarios. Esta trazabilidad permite observar el desarrollo del estudiante y comparar evidencias producidas en diferentes etapas. Liao et al. (2022) señalan que las evaluaciones digitales administradas continuamente requieren mecanismos para mantener y supervisar su calidad. En consecuencia, la acumulación de información debe acompañarse de procedimientos que revisen la consistencia de los instrumentos, la estabilidad de los criterios y la pertinencia de las interpretaciones realizadas a partir de los registros disponibles.

La integración de cuestionarios automatizados representa una de las funciones más utilizadas en las plataformas virtuales. Estos instrumentos agilizan la aplicación, corrección y comunicación de resultados, especialmente cuando se trabaja con grupos numerosos. Durante la expansión de la enseñanza remota, los sistemas de evaluación en línea adquirieron una presencia creciente por su capacidad para sostener actividades académicas a distancia (Topuz et al., 2022). Sin embargo, cuando predominan preguntas cerradas y respuestas memorísticas, la digitalización reproduce las limitaciones del examen convencional sin ampliar la comprensión de los procesos de aprendizaje.

Las tareas abiertas, los foros y los portafolios electrónicos amplían el tipo de evidencias que pueden recopilarse dentro de un sistema de gestión. Zhang y Tur (2024) destacan que los portafolios digitales permiten integrar productos, reflexiones y muestras de progreso generadas a lo largo del tiempo. Los foros, a su vez, pueden evidenciar argumentación, interacción y construcción colectiva de conocimientos. Su valoración exige criterios específicos, porque la cantidad de intervenciones no equivale necesariamente a calidad académica. La plataforma organiza la evidencia, pero el sentido evaluativo procede del diseño pedagógico.

La incorporación de rúbricas digitales contribuye a transparentar expectativas, dimensiones de desempeño y niveles de logro. Estos instrumentos pueden vincularse directamente con tareas, proyectos y producciones alojadas en la plataforma, facilitando la comunicación de resultados. El uso de rúbricas fortalece la evaluación formativa cuando permite interpretar la calidad del trabajo y orientar procesos de revisión (Romero et al., 2025). No obstante, su aplicación pierde profundidad si se limita a seleccionar casillas o asignar puntuaciones, sin proporcionar explicaciones que ayuden al estudiante a comprender las diferencias entre niveles.

La retroalimentación constituye otra función decisiva de los sistemas de gestión del aprendizaje. Las plataformas permiten entregar comentarios escritos, archivos corregidos, grabaciones y respuestas automáticas en tiempos más breves. Banihashem et al. (2022) sostienen que la analítica del aprendizaje puede fortalecer las prácticas de retroalimentación cuando facilita información pertinente sobre el progreso. Esta posibilidad debe vincularse con oportunidades de revisión y mejora. Si el comentario llega después de cerrar la actividad o se reduce a una calificación, la eficiencia tecnológica no se traduce en un beneficio formativo significativo.

Los registros generados por las plataformas sustentan procesos de analítica del aprendizaje orientados a reconocer patrones de participación, dificultades y trayectorias. Muñoz et al. (2021) explican que estas tecnologías permiten examinar información procedente de la actividad académica. Aun así, los datos de conexión, permanencia o frecuencia de acceso no deben interpretarse como equivalentes directos del aprendizaje. Su análisis requiere contrastarlos con producciones, criterios y conocimiento del contexto. La competencia profesional consiste en transformar registros técnicos en evidencias pedagógicas sin formular conclusiones desproporcionadas ni descontextualizadas.

Figura 2

¿Cómo utilizar las plataformas virtuales para la evaluación?



Nota. La figura presenta diferentes recursos de las plataformas virtuales que facilitan la evaluación, la retroalimentación y el seguimiento del aprendizaje.

La adopción institucional de una plataforma está condicionada por infraestructura, formación docente, soporte técnico y acuerdos sobre sus usos. Los factores de adopción de la analítica incluyen experiencias previas, capacidades organizativas y prioridades educativas, según Tsai et al. (2021). Por tanto, la implementación no puede reducirse a adquirir licencias o habilitar cuentas. Se requieren orientaciones sobre diseño de actividades, interpretación de datos y comunicación de resultados. Sin una estrategia compartida, cada docente puede utilizar el sistema de forma aislada, generando prácticas fragmentadas y criterios evaluativos poco consistentes.

La accesibilidad y la brecha digital condicionan las oportunidades de participar en evaluaciones gestionadas mediante plataformas. Los problemas de navegación, los formatos incompatibles y la dependencia de conexiones estables pueden afectar el desempeño sin guardar relación con los aprendizajes valorados (Scanlon et al., 2021). Además, las desigualdades digitales comprenden diferencias de acceso, preparación y apoyo institucional, no solo disponibilidad de dispositivos (Van de Werfhorst et al., 2022). Una evaluación equitativa debe incorporar alternativas de participación y diseños accesibles, preservando al mismo tiempo los criterios académicos fundamentales.

Finalmente, la gestión de datos educativos plantea responsabilidades vinculadas con privacidad, transparencia y uso legítimo de la información. Korir et al. (2023) muestran que las preocupaciones estudiantiles abarcan la recopilación, utilización y compartición de datos generados por la analítica. En Ecuador, estas prácticas deben ajustarse a los principios establecidos por la Ley Orgánica de Protección de Datos Personales (2021). Las plataformas pueden fortalecer la evaluación cuando integran evidencias y retroalimentación, pero su utilización responsable exige limitar los datos recopilados, controlar los accesos y conservar la supervisión humana.

2.2 Cuestionarios, pruebas y exámenes digitales

Los cuestionarios, pruebas y exámenes digitales representan una de las formas más extendidas de evaluación mediada por tecnología. Su uso permite administrar actividades a distancia, automatizar parte de la corrección y almacenar resultados de manera organizada. Heil y Ifenthaler (2023) muestran que la evaluación en línea comprende una amplia variedad de formatos y procedimientos. Sin embargo, la digitalización de una prueba no modifica por sí misma su calidad pedagógica, pues un instrumento puede conservar enfoques memorísticos, criterios limitados y escasa utilidad formativa, aunque se aplique mediante una plataforma.

La principal diferencia entre un cuestionario digital y uno convencional radica en las posibilidades de configuración, adaptación y seguimiento. Las plataformas permiten combinar preguntas cerradas, respuestas breves, ejercicios de relación, secuencias aleatorias y retroalimentación inmediata. Estas funciones amplían la flexibilidad del instrumento, pero también pueden inducir a seleccionar tipos de pregunta por facilidad técnica y no por pertinencia educativa (Topuz et al., 2022). El diseño debe partir del aprendizaje esperado, de modo que cada reactivo produzca evidencia válida sobre conocimientos, habilidades o capacidades específicas.

Las pruebas de respuesta cerrada ofrecen ventajas en términos de rapidez, objetividad operativa y procesamiento de grandes volúmenes de información. No obstante, su utilidad disminuye cuando se emplean para valorar razonamiento complejo, argumentación o transferencia. La calidad de estas pruebas depende de la construcción de los ítems, la claridad de las opciones y la correspondencia con el contenido evaluado. Una selección deficiente puede medir reconocimiento superficial en lugar de comprensión. Por ello, la automatización de la puntuación no debe confundirse con validez ni con precisión pedagógica.

Las preguntas abiertas amplían las posibilidades de valorar explicación, análisis y producción propia. En entornos digitales, estas respuestas pueden procesarse mediante revisión docente o sistemas automatizados de puntuación. Maestrales et al. (2021) muestran que el aprendizaje automático puede utilizarse para calificar evaluaciones multidimensionales en ciencias. Sin embargo, la calidad de la puntuación depende de la definición del constructor y de las evidencias utilizadas para entrenar el sistema. La rapidez técnica no elimina la necesidad de comprobar si la calificación representa realmente el desempeño esperado.

La retroalimentación inmediata constituye uno de los principales aportes de los cuestionarios digitales. Cuando las respuestas incluyen explicaciones, pistas o indicaciones de mejora, el instrumento puede cumplir una función formativa y no solo calificadora. Los sistemas de retroalimentación automatizada pueden apoyar la revisión del trabajo del estudiante y orientar nuevos intentos (Lee et al., 2021). Su eficacia se reduce, en cambio, cuando se limita a informar si una respuesta es correcta o incorrecta. La retroalimentación debe relacionar el error con el criterio y ofrecer una oportunidad real de corrección.

La aleatorización de preguntas y opciones se utiliza con frecuencia para reducir la copia y diversificar las versiones de una prueba. Esta función puede fortalecer la integridad, siempre que las variantes mantengan niveles equivalentes de dificultad y representen los mismos objetivos. Liao et al. (2022) subrayan la necesidad de supervisar continuamente la calidad de las evaluaciones digitales administradas de manera reiterada. Si los bancos de ítems no se revisan, pueden aparecer desigualdades entre versiones, preguntas desactualizadas o patrones de respuesta que comprometan la interpretación de los resultados.

Los exámenes digitales también plantean desafíos relacionados con la autenticidad de las respuestas y las condiciones de aplicación. La evaluación en línea puede facilitar el acceso no autorizado a recursos, la colaboración no permitida o la suplantación, pero estos riesgos no se resuelven únicamente mediante vigilancia tecnológica (Holden et al., 2021). Las pruebas contextualizadas, la justificación de respuestas y la combinación con defensas orales reducen la dependencia del control. La integridad debe abordarse mediante diseño pedagógico, reglas claras y procedimientos proporcionales al nivel de riesgo.

Las condiciones técnicas influyen directamente en el desempeño cuando la evaluación depende de conectividad, dispositivos y familiaridad con la plataforma. Las desigualdades de preparación digital pueden afectar los resultados y producir interpretaciones injustas sobre el aprendizaje (Van de Werfhorst et al., 2022). Mathrani et al. (2022) muestran que la brecha

digital incluye factores sociales e institucionales, además del acceso a internet. Por ello, las pruebas deben contemplar tiempos razonables, procedimientos alternativos y protocolos frente a fallas, evitando que un problema técnico se convierta en una desventaja académica.

Los cuestionarios y exámenes digitales ofrecen eficiencia, trazabilidad y posibilidades de retroalimentación, pero su valor depende del diseño, la validez y las condiciones de aplicación. No todos los aprendizajes pueden representarse mediante respuestas automatizables, ni toda prueba en línea constituye una innovación evaluativa. La combinación de formatos, revisión de ítems, supervisión de calidad y atención a la equidad permite aprovechar sus ventajas sin reducir la evaluación a puntuaciones rápidas. La tecnología debe apoyar la interpretación pedagógica y no sustituir la reflexión sobre qué se evalúa y para qué.

2.3 Rúbricas digitales, listas de cotejo y escalas de valoración

Las rúbricas digitales, las listas de cotejo y las escalas de valoración permiten organizar criterios de evaluación y transformar observaciones cualitativas en evidencias sistemáticas sobre el desempeño. Aunque comparten una función estructuradora, responden a propósitos distintos: las rúbricas describen niveles de calidad, las listas verifican presencia o ausencia de elementos y las escalas estiman grados de logro. Su incorporación en plataformas virtuales facilita la aplicación, el registro y la comunicación de resultados, pero su utilidad depende de que los criterios representen con precisión el aprendizaje esperado y no solo aspectos fácilmente observables.

Las rúbricas digitales se caracterizan por articular criterios, descriptores y niveles de desempeño dentro de una matriz accesible para docentes y estudiantes. Romero et al. (2025) destacan su presencia predominante en la evaluación formativa de la educación general básica. Su valor radica en hacer visibles las expectativas y ofrecer una referencia común para interpretar la calidad de una producción. Sin embargo, una rúbrica no mejora automáticamente la evaluación: cuando sus descriptores son ambiguos, repetitivos o excesivamente generales, puede generar calificaciones aparentes de precisión sin una base pedagógica consistente.

Una ventaja central de las rúbricas digitales es su capacidad para apoyar la retroalimentación. Al vincular cada criterio con niveles concretos de desempeño, permiten señalar con mayor claridad dónde se encuentra el estudiante y qué aspectos requieren mejora. La retroalimentación resulta más efectiva cuando es comprensible, específica y orientada hacia acciones posteriores (Vera, 2022). En este marco, la rúbrica debe utilizarse antes, durante y

después de la tarea, no únicamente como instrumento final de calificación. Su lectura anticipada favorece planificación, autorregulación y revisión consciente del trabajo.

Las listas de cotejo responden a una lógica más simple, pues registran si una conducta, requisito o componente está presente. Su utilidad es mayor cuando se evalúan procedimientos, cumplimiento de etapas o elementos indispensables de una producción. No obstante, su carácter dicotómico limita la posibilidad de diferenciar grados de calidad. Una respuesta marcada como cumplida puede ocultar variaciones importantes en precisión, profundidad o coherencia. Por ello, las listas resultan más pertinentes como apoyo de seguimiento que como único instrumento para valorar aprendizajes complejos, argumentaciones o desempeños integrales.

Las escalas de valoración introducen niveles intermedios entre la presencia y la ausencia de un criterio. Mediante categorías, frecuencias o puntuaciones graduadas, permiten estimar intensidad, regularidad o dominio. Su aplicación requiere descriptores claros, porque términos como "siempre", "frecuentemente" o "adecuado" pueden interpretarse de manera distinta entre evaluadores. La validez de cualquier instrumento depende de la relación entre las evidencias observadas y las conclusiones que se extraen de ellas (Gu, 2021). Una escala ambigua puede convertir percepciones subjetivas en resultados numéricos aparentemente objetivos.

La digitalización de estos instrumentos facilita su integración con tareas, portafolios, foros y proyectos alojados en sistemas de gestión del aprendizaje. Las plataformas permiten asociar criterios con entregas concretas, conservar historiales y comparar resultados a lo largo del tiempo. Esta trazabilidad puede fortalecer la evaluación continua, siempre que los registros se utilicen para comprender avances y no solo para acumular calificaciones (Liao et al., 2022). La automatización del cálculo reduce trabajo operativo, pero no sustituye la observación, la interpretación ni la justificación del juicio docente.

Tabla 3
Diferencias funcionales entre rúbricas digitales, listas de cotejo y escalas de valoración

Instrumento	Tipo de juicio	Uso más pertinente	Principal fortaleza	Riesgo de uso inadecuado
Rúbrica digital	Cualitativo y graduado	Producciones complejas, proyectos, argumentaciones y desempeños	Explicita criterios y niveles de calidad	Fragmentar el desempeño o usar descriptores ambiguos

Instrumento	Tipo de juicio	Uso más pertinente	Principal fortaleza	Riesgo de uso inadecuado
Lista de cotejo	Dicotómico	Verificación de requisitos, pasos o componentes	Facilita seguimiento rápido y objetivo	Confundir cumplimiento con calidad
Escala de valoración	Graduado por frecuencia, intensidad o dominio	Observación de conductas, actitudes y progresión	Permite reconocer niveles intermedios	Convertir percepciones subjetivas en puntuaciones aparentes
Rúbrica con retroalimentación digital	Cualitativo, graduado y orientador	Revisión de borradores y mejora progresiva	Vincula resultado con acciones posteriores	Reducirse a selección automática de niveles
Lista de cotejo para autoevaluación	Dicotómico y reflexivo	Revisión previa a una entrega	Favorece control y autonomía	Promover verificación superficial sin análisis
Escala compartida entre evaluadores	Graduado y comparativo	Coevaluación y valoración colaborativa	Facilita contraste de perspectivas	Generar diferencias por falta de criterios comunes

Nota. La tabla distingue la lógica evaluativa de cada instrumento y sus condiciones de uso, con base en las fuentes empleadas.

La comparación evidencia que la elección del instrumento debe depender del tipo de evidencia y del nivel de complejidad del aprendizaje. Las rúbricas resultan más adecuadas para valorar calidad, mientras las listas permiten verificar componentes y las escalas estiman grados de manifestación. Utilizar un solo formato para todas las tareas reduce la sensibilidad evaluativa. La diversificación de instrumentos favorece una comprensión más amplia del desempeño, siempre que existan criterios coherentes y una relación directa con los objetivos establecidos.

Estos instrumentos también pueden favorecer la autoevaluación y la coevaluación cuando los criterios son comprensibles y se aplican con orientación previa. Santana et al. (2025) relacionan la evaluación entre pares con el desarrollo de competencias transversales en proyectos integradores. Las rúbricas digitales permiten comparar producciones, justificar juicios y discutir estándares de calidad. Las listas de cotejo ayudan a revisar requisitos antes de una entrega, mientras las escalas facilitan reconocer progresos. Sin acompañamiento, no obstante, la participación estudiantil puede reproducir valoraciones superficiales o basadas en afinidades personales.

La consistencia entre evaluadores representa un desafío común. Dos docentes pueden interpretar de manera diferente un mismo descriptor, especialmente cuando los criterios incluyen conceptos amplios como creatividad, coherencia o profundidad. El uso de ejemplos, sesiones de calibración y revisión conjunta contribuye a reducir discrepancias. La claridad de los criterios resulta esencial para sostener inferencias válidas sobre el desempeño (Gu, 2021). En entornos digitales, la plataforma puede registrar decisiones, pero no corrige por sí sola las diferencias de interpretación ni garantiza que los evaluadores comprendan los niveles de manera equivalente.

La accesibilidad también debe considerarse en el diseño de estos instrumentos. Una rúbrica extensa, una escala con lenguaje técnico o una lista presentada en un formato incompatible puede convertirse en una barrera. Las directrices del CAST (2024) promueven múltiples formas de representación y participación, lo que implica presentar criterios de manera clara, navegable y comprensible. La accesibilidad no supone simplificar el nivel académico, sino eliminar obstáculos irrelevantes para que todos los estudiantes puedan conocer las expectativas, interpretar la retroalimentación y utilizar los instrumentos durante la revisión de sus producciones.

En conjunto, las rúbricas digitales, listas de cotejo y escalas de valoración fortalecen la evaluación cuando sus funciones se distinguen y se articulan con objetivos claros. Su eficacia no reside en convertir juicios en casillas o porcentajes, sino en organizar evidencias, transparentar expectativas y orientar mejoras. Las plataformas amplían la trazabilidad y facilitan el registro, pero el valor pedagógico continúa dependiendo del diseño de criterios, la calidad de los descriptores y la interpretación docente. La selección adecuada evita tanto la simplificación excesiva como la falsa precisión evaluativa.

2.4 Portafolios electrónicos y evidencias digitales de aprendizaje

Los portafolios electrónicos constituyen herramientas destinadas a recopilar, organizar y presentar evidencias de aprendizaje producidas a lo largo del tiempo. A diferencia de una carpeta digital utilizada únicamente para almacenar archivos, su valor evaluativo depende de la selección intencional de trabajos, reflexiones y revisiones vinculadas con objetivos y criterios definidos. Zhang y Tur (2024) muestran que el uso de portafolios electrónicos se fortaleció durante la pandemia y ofrece orientaciones para prácticas posteriores. Su función principal consiste en representar procesos, progresos y resultados dentro de una secuencia coherente.

La estructura del portafolio permite integrar evidencias diversas, entre ellas textos, imágenes, grabaciones, presentaciones, proyectos, informes y productos multimedia. Esta variedad amplía las posibilidades de demostrar aprendizajes que no siempre pueden observarse mediante pruebas convencionales. La competencia digital puede evaluarse mediante métodos múltiples y no mediante un único tipo de instrumento (Sillat et al., 2021). Por ello, el portafolio resulta pertinente cuando se busca valorar desempeños complejos, integración de conocimientos, capacidad de producción y evolución de una competencia durante un periodo determinado.

La selección de evidencias constituye una fase central del proceso, porque no todo archivo producido por el estudiante posee el mismo valor evaluativo. El portafolio debe incluir muestras capaces de demostrar logros, dificultades, cambios y decisiones tomadas durante el aprendizaje. Zhang y Tur (2024) destacan que su uso adquiere mayor sentido cuando se relaciona con reflexión y desarrollo progresivo. La acumulación indiscriminada de documentos, en cambio, dificulta la interpretación y convierte la herramienta en un repositorio sin jerarquía, propósito formativo ni conexión suficiente con los criterios establecidos.

La reflexión escrita o audiovisual diferencia al portafolio de otras formas de registro. Mediante ella, el estudiante explica por qué seleccionó una evidencia, qué aprendió, qué dificultades enfrentó y cómo modificó su trabajo. Esta práctica favorece la autoevaluación porque exige interpretar el propio desempeño frente a criterios explícitos. La evaluación formativa puede fortalecer el desarrollo de competencias cuando ofrece oportunidades para analizar avances y realizar ajustes durante el proceso (Huauya et al., 2023). Sin reflexión, la evidencia muestra productos, pero revela poco sobre las decisiones que explican su evolución.

El portafolio electrónico también facilita la retroalimentación continua, ya que docentes y estudiantes pueden revisar versiones sucesivas de una misma producción. Los comentarios digitales permiten señalar cambios, formular preguntas y orientar nuevas decisiones sin esperar hasta la entrega final. La retroalimentación alcanza mayor efectividad cuando es oportuna, comprensible y vinculada con acciones concretas de mejora (Vera, 2022). Esta dinámica convierte la evaluación en un proceso iterativo, donde cada evidencia puede revisarse, ampliarse o sustituirse conforme el estudiante demuestra un aprendizaje más sólido.

La utilización de rúbricas fortalece la valoración del portafolio al establecer criterios comunes para examinar selección, calidad, reflexión, organización y progreso. Romero et al. (2025) destacan la importancia de estos instrumentos dentro de la evaluación formativa, especialmente por su capacidad para hacer visibles las expectativas. En un portafolio, la rúbrica debe valorar

tanto los productos incluidos como la coherencia del conjunto. Una calificación centrada únicamente en la presentación visual puede desviar la atención del aprendizaje y privilegiar habilidades técnicas que no corresponden al propósito académico principal.

Las evidencias digitales ofrecen trazabilidad, pues permiten conservar fechas, versiones, comentarios y cambios realizados durante el desarrollo de una tarea. Esta información amplía la comprensión del proceso y reduce la dependencia de un producto final aislado. Los sistemas de evaluación digital administrados de forma continua requieren mecanismos para mantener la calidad de las evidencias y de sus interpretaciones (Liao et al., 2022). La trazabilidad, por tanto, no equivale automáticamente a validez; los registros deben analizarse según criterios pedagógicos y dentro de las condiciones concretas en que fueron producidos.

El portafolio puede favorecer la integración entre aprendizaje presencial, virtual e híbrido. Las evidencias generadas en distintos espacios pueden reunirse dentro de una misma estructura y relacionarse con objetivos compartidos. Esta característica permite observar conexiones entre tareas, proyectos, prácticas y reflexiones desarrolladas durante el curso. Sin embargo, la integración exige una planificación que evite duplicaciones y acumulación excesiva. La adopción de herramientas digitales depende de prioridades pedagógicas, capacidades institucionales y experiencias de uso previamente construidas (Tsai et al., 2021), no solo de la disponibilidad de una plataforma.

La evaluación mediante portafolios también puede incorporar autoevaluación y coevaluación. Los estudiantes pueden revisar sus propias producciones, comentar evidencias de sus compañeros y justificar valoraciones con base en criterios comunes. Estas prácticas favorecen responsabilidad, argumentación y comprensión de la calidad del trabajo. La evaluación entre pares puede contribuir al desarrollo de competencias transversales cuando se integra en proyectos y procesos guiados (Santana et al., 2025). No obstante, requiere acompañamiento docente para evitar comentarios superficiales, comparaciones personales o juicios desvinculados de los criterios académicos.

La accesibilidad representa una condición indispensable para que el portafolio electrónico permita demostrar aprendizajes en igualdad de oportunidades. Las plataformas deben admitir formatos compatibles, navegación clara y alternativas para estudiantes con diferentes necesidades. Los errores de accesibilidad web pueden limitar la participación de personas con discapacidad, aunque el contenido académico sea pertinente (Scanlon et al., 2021). Las directrices del CAST (2024) respaldan el uso de múltiples formas de representación y

expresión, lo que permite diversificar evidencias sin reducir los criterios esenciales de valoración ni alterar los objetivos formativos.

La privacidad y la propiedad de las evidencias requieren especial atención, debido a que los portafolios pueden contener datos personales, imágenes, producciones originales y registros de desempeño. Las preocupaciones estudiantiles sobre recopilación, utilización y compartición de información se intensifican cuando los datos permanecen almacenados en sistemas digitales (Korir et al., 2023). En Ecuador, su gestión debe ajustarse a la Ley Orgánica de Protección de Datos Personales (2021). Las instituciones necesitan definir quién accede, cuánto tiempo se conserva la información y con qué finalidad puede utilizarse.

Los portafolios electrónicos fortalecen la evaluación cuando articulan selección, reflexión, retroalimentación, criterios y seguimiento progresivo. Su potencial no reside en digitalizar carpetas tradicionales, sino en construir una narrativa verificable del aprendizaje mediante evidencias relacionadas entre sí. Para que cumplan esta función, deben evitar la acumulación superficial, la sobrevaloración del diseño visual y la recopilación excesiva de datos. La calidad del portafolio depende de su coherencia con los objetivos, de la participación reflexiva del estudiante y de una interpretación docente capaz de contextualizar cada evidencia.

2.5 Herramientas interactivas, gamificación y evaluación mediante juegos

Las herramientas interactivas han ampliado las posibilidades de evaluar el aprendizaje mediante actividades que combinan respuesta inmediata, participación activa y seguimiento continuo. Plataformas como Educaplay y Kahoot permiten integrar preguntas, retos y dinámicas de competencia dentro de experiencias digitales estructuradas. Su valor pedagógico no depende únicamente del atractivo visual, sino de la relación entre las actividades propuestas y los objetivos evaluativos. La analítica sustentada en estas herramientas puede aportar información sobre desempeño, participación y dificultades recurrentes (Carrasco et al., 2025), siempre que los resultados se interpreten críticamente.

La gamificación consiste en incorporar elementos propios de los juegos, como puntos, insignias, niveles, recompensas o clasificaciones, dentro de actividades que no son juegos en sentido estricto. Martínez et al. (2025) vinculan esta estrategia con el fortalecimiento de la autorregulación en el aprendizaje. Desde la evaluación, estos componentes pueden incrementar la frecuencia de participación y ofrecer señales visibles de progreso. Sin embargo, el uso excesivo de recompensas externas puede desplazar la atención desde la comprensión hacia la obtención de puntuaciones, reduciendo la profundidad del aprendizaje evaluado.

La evaluación mediante juegos presenta una lógica distinta, pues el propio entorno lúdico se convierte en espacio para observar decisiones, estrategias, resolución de problemas y capacidad de adaptación. En estos contextos, el desempeño puede analizarse a partir de acciones, rutas seguidas y respuestas frente a situaciones progresivamente complejas. Las simulaciones virtuales muestran que los escenarios digitales pueden favorecer aprendizajes significativos y percepción de utilidad formativa (Edgar et al., 2022). No obstante, la interpretación exige distinguir entre habilidad para jugar y dominio real del contenido o competencia evaluada.

La retroalimentación inmediata constituye una de las principales fortalezas de las herramientas interactivas. El estudiante puede conocer resultados, revisar errores y volver a intentarlo dentro de una misma actividad. Esta dinámica favorece el aprendizaje cuando la respuesta ofrecida explica la naturaleza del error y orienta una acción posterior, en lugar de limitarse a mostrar aciertos o fallos (Vera, 2022). Una retroalimentación exclusivamente automática puede resultar insuficiente para aprendizajes complejos, por lo que conviene combinarla con comentarios docentes, preguntas orientadoras y oportunidades de revisión.

La gamificación también puede contribuir a la autorregulación cuando hace visibles metas, avances y dificultades. Los niveles, barras de progreso y secuencias de retos permiten que el estudiante observe su desempeño y ajuste estrategias. Martínez et al. (2025) señalan que estas dinámicas pueden potenciar el compromiso con el aprendizaje. Su efectividad, sin embargo, depende de que los incentivos estén vinculados con logros relevantes. Premiar únicamente rapidez, frecuencia de acceso o acumulación de puntos puede favorecer conductas superficiales y producir una imagen distorsionada del progreso académico.

La competencia entre estudiantes debe gestionarse con cautela. Los rankings y comparaciones públicas pueden motivar a algunos participantes, pero también generar ansiedad, frustración o desinterés en quienes acumulan resultados bajos. La evaluación debe preservar oportunidades de mejora y evitar que la clasificación se convierta en el objetivo principal. Las herramientas interactivas son más formativas cuando permiten avanzar a ritmos diversos, repetir actividades y recibir apoyos diferenciados. La personalización favorece trayectorias ajustadas al desempeño, aunque requiere criterios pedagógicos claros para evitar recorridos excesivamente simplificados (Aparicio y Aparicio, 2024).

Tabla 4

Diferencias entre herramientas interactivas, gamificación y evaluación mediante juegos

Modalidad	Elemento central	Evidencia que genera	Principal aporte	Riesgo evaluativo
Herramienta interactiva	Respuesta y participación digital	Aciertos, errores, intentos y tiempos	Retroalimentación rápida y seguimiento	Reducir el aprendizaje a respuestas cerradas
Gamificación	Puntos, niveles, insignias y retos	Progreso, constancia y cumplimiento	Motivación y autorregulación	Priorizar recompensas sobre comprensión
Evaluación mediante juegos	Resolución de situaciones lúdicas	Decisiones, estrategias y adaptación	Observación de desempeños complejos	Confundir habilidad de juego con aprendizaje
Simulación digital	Representación de escenarios	Procedimientos, elecciones y consecuencias	Aplicación contextualizada del conocimiento	Dependencia de familiaridad tecnológica
Actividad competitiva	Comparación de resultados	Rapidez, precisión y posición relativa	Participación inmediata	Ansiedad, exclusión o superficialidad
Actividad colaborativa	Resolución conjunta de retos	Interacción, argumentación y acuerdos	Desarrollo de competencias transversales	Dificultad para distinguir aportes individuales

Nota. La tabla diferencia el tipo de evidencia producido por cada modalidad y los principales riesgos que deben considerarse en su diseño evaluativo.

La comparación muestra que estas estrategias no son equivalentes ni deben aplicarse de manera indistinta. Las herramientas interactivas resultan útiles para comprobaciones rápidas, mientras la gamificación estructura la motivación y la evaluación mediante juegos permite observar actuaciones complejas. Su selección debe depender del propósito, del tipo de aprendizaje y de la evidencia requerida. Cuando se combinan sin planificación, pueden generar abundantes datos, pero poca información relevante sobre comprensión. El diseño pedagógico debe preceder siempre a la elección de la dinámica tecnológica.

Las plataformas interactivas también producen datos que pueden integrarse en procesos de analítica del aprendizaje. Carrasco et al. (2025) muestran que herramientas como Educaplay y Kahoot pueden utilizarse para examinar diferencias en el desempeño y apoyar procesos de enseñanza. Estos registros permiten identificar preguntas difíciles, patrones de error y niveles de participación. No obstante, la interpretación debe evitar que rapidez, frecuencia o puntuación global se conviertan en indicadores exclusivos. Los resultados requieren contraste

con otras evidencias, especialmente cuando se evalúan competencias complejas o procesos de razonamiento.

La accesibilidad condiciona el valor de las actividades gamificadas y de los juegos educativos. Un diseño basado en estímulos visuales rápidos, navegación compleja o tiempos estrictos puede excluir a estudiantes con distintas necesidades. Las barreras tecnológicas y de diseño reducen la participación incluso cuando el contenido es pertinente (Scanlon et al., 2021). Las directrices del CAST (2024) respaldan la diversificación de formas de interacción y expresión. Esto implica ofrecer alternativas, ajustar tiempos y evitar que la destreza tecnológica determine injustificadamente el resultado académico.

En conjunto, las herramientas interactivas, la gamificación y la evaluación mediante juegos pueden enriquecer la valoración del aprendizaje al incrementar participación, ofrecer retroalimentación y producir evidencias variadas. Su potencial depende de la correspondencia entre dinámica, criterio y competencia evaluada. El componente lúdico no debe ocultar la necesidad de validez, equidad y profundidad pedagógica. Cuando la motivación se articula con tareas significativas, estas estrategias fortalecen el proceso evaluativo; cuando predominan puntos y rankings, pueden reducirlo a una experiencia atractiva, pero académicamente limitada.

2.6 Evaluación audiovisual, simuladores y laboratorios virtuales

La evaluación audiovisual amplía las formas de demostrar el aprendizaje mediante videos, presentaciones narradas, grabaciones de procedimientos y producciones multimedia. Estos formatos permiten observar comunicación, argumentación, secuencias de actuación y capacidad para integrar distintos recursos expresivos. Su valor no depende del acabado técnico, sino de la relación entre la evidencia producida y los criterios establecidos. La diversificación de medios favorece una representación más amplia del desempeño, siempre que la valoración no confunda habilidades de edición con dominio conceptual o calidad académica del contenido presentado.

Los productos audiovisuales resultan especialmente útiles cuando el aprendizaje exige explicar procesos, justificar decisiones o mostrar procedimientos difíciles de representar mediante respuestas escritas. Las directrices del CAST (2024) respaldan el uso de múltiples formas de expresión, lo cual permite que los estudiantes comuniquen sus aprendizajes mediante formatos diversos. No obstante, esta flexibilidad requiere criterios comunes que preserven la exigencia académica. La evaluación debe centrarse en la precisión, coherencia y pertinencia de la evidencia, evitando que la calidad visual sustituya el análisis del conocimiento demostrado.

Los simuladores virtuales permiten representar escenarios, problemas y situaciones de práctica dentro de entornos controlados. Edgar et al. (2022) muestran que la educación basada en simulación virtual puede generar una valoración positiva e influir en el aprendizaje estudiantil. Su utilidad evaluativa radica en observar decisiones, respuestas y secuencias de actuación frente a condiciones que pueden modificarse. A diferencia de una prueba convencional, el simulador produce evidencia sobre el desempeño en acción, aunque su interpretación exige distinguir entre comprensión disciplinar, familiaridad tecnológica y capacidad para operar dentro del entorno.

Los laboratorios virtuales comparten esta lógica, pero se orientan con mayor énfasis a la experimentación, observación y manipulación de variables. En la educación médica, la simulación virtual ha sido utilizada para recrear prácticas y situaciones formativas que requieren seguridad, repetición y acceso controlado (Wu et al., 2022). Estos entornos permiten ensayar procedimientos sin los riesgos o costos asociados con ciertas experiencias presenciales. Sin embargo, no todas las habilidades prácticas pueden validarse exclusivamente mediante interacción digital, por lo que su uso debe complementarse con otras evidencias cuando corresponda.

La retroalimentación inmediata constituye una de las principales fortalezas de simuladores y laboratorios virtuales. El sistema puede señalar errores, mostrar consecuencias, ofrecer pistas y permitir nuevos intentos durante la actividad. Esta respuesta favorece la revisión cuando orienta al estudiante hacia una comprensión más precisa del procedimiento o de la decisión adoptada (Vera, 2022). Si la retroalimentación se limita a indicar éxito o fracaso, su aporte formativo disminuye. La calidad depende de que el estudiante pueda relacionar el resultado con criterios, causas y alternativas de mejora.

Tabla 5
Características evaluativas de los recursos audiovisuales, simuladores y laboratorios virtuales

Recurso	Evidencia principal	Aplicación evaluativa	Aporte pedagógico	Riesgo de interpretación
Video explicativo	Argumentación, comunicación y secuencia conceptual	Presentaciones, demostraciones y defensas	Permite observar comprensión y claridad expositiva	Sobrevalorar edición y calidad visual

Recurso	Evidencia principal	Aplicación evaluativa	Aporte pedagógico	Riesgo de interpretación
Registro audiovisual de procedimiento	Ejecución, orden y toma de decisiones	Prácticas, proyectos y demostraciones	Documenta el desempeño paso a paso	Confundir repetición ensayada con dominio
Simulador virtual	Decisiones, estrategias y respuestas	Escenarios profesionales o situaciones complejas	Facilita práctica segura y repetible	Medir familiaridad con el sistema
Laboratorio virtual	Manipulación de variables y análisis de resultados	Experimentación científica y técnica	Amplía acceso y posibilidad de repetir ensayos	Sustituir indebidamente la práctica real
Simulación colaborativa	Coordinación, comunicación y resolución conjunta	Trabajo en equipo y situaciones integradoras	Permite valorar competencias transversales	Dificultad para diferenciar aportes individuales

Nota. La tabla organiza los tipos de evidencia, aportes y riesgos asociados con la evaluación mediante recursos audiovisuales, simuladores y laboratorios virtuales, con base en las fuentes utilizadas.

La comparación muestra que estos recursos producen evidencias distintas y, por tanto, requieren criterios diferenciados. Un video permite valorar comunicación y argumentación, mientras un simulador se orienta hacia decisiones y respuestas ante situaciones. Los laboratorios virtuales, por su parte, facilitan la observación de procedimientos y relaciones entre variables. Su selección debe responder al tipo de aprendizaje esperado y no a la novedad tecnológica. La combinación de varios recursos puede ofrecer una comprensión más completa del desempeño, siempre que las evidencias no se dupliquen innecesariamente.

La validez de estas evaluaciones depende de que el entorno represente adecuadamente la competencia que se desea valorar. Zhai et al. (2021) sostienen que las evaluaciones apoyadas en tecnologías avanzadas requieren inferencias justificadas entre respuestas, puntuaciones e interpretaciones. Un simulador puede registrar múltiples acciones, pero no todas poseen el mismo valor pedagógico. Es necesario definir qué comportamientos evidencian dominio, cuáles responden al ensayo y error y qué resultados requieren revisión humana. La abundancia de datos no elimina la necesidad de criterios explícitos y análisis contextualizado.

La accesibilidad condiciona la participación en evaluaciones audiovisuales y entornos simulados. Los recursos pueden generar barreras cuando utilizan interfaces complejas, tiempos rígidos, estímulos exclusivamente visuales o formatos incompatibles con tecnologías de apoyo. Estas dificultades pueden limitar el desempeño de estudiantes con discapacidad, aunque posean los conocimientos evaluados (Scanlon et al., 2021). Las alternativas de navegación, subtítulos,

transcripciones y diferentes formas de respuesta permiten ampliar la participación. La accesibilidad debe incorporarse desde el diseño y no aplicarse únicamente como corrección posterior.

En conjunto, la evaluación audiovisual, los simuladores y los laboratorios virtuales amplían la observación del aprendizaje al incorporar comunicación, procedimientos, decisiones y experimentación. Su potencial reside en producir evidencias que las pruebas tradicionales no siempre captan, pero su uso exige criterios válidos, accesibilidad, retroalimentación y supervisión docente. Estas herramientas no sustituyen automáticamente la práctica presencial ni garantizan una valoración más precisa. Su aporte depende de cómo se integran con los objetivos formativos, las condiciones de aplicación y la interpretación crítica de los resultados.

2.7 Criterios para la selección e implementación de herramientas evaluativas

La selección de herramientas evaluativas digitales debe partir de los propósitos de aprendizaje y no de la novedad tecnológica. Una plataforma, aplicación o sistema automatizado resulta pertinente cuando permite recoger evidencias válidas, interpretables y útiles para orientar decisiones pedagógicas. La evaluación en línea comprende múltiples formatos, pero su calidad depende de la coherencia entre objetivos, tareas, criterios y condiciones de aplicación (Heil y Ifenthaler, 2023). Elegir una herramienta por popularidad, facilidad de uso o disponibilidad institucional puede conducir a prácticas eficientes en lo operativo, aunque limitadas en su valor educativo.

El primer criterio corresponde a la alineación pedagógica. La herramienta debe permitir evaluar el tipo de conocimiento, habilidad o competencia que se pretende desarrollar. Sillat et al. (2021) muestran que la evaluación de competencias digitales requiere métodos diversos, lo que impide depender de un único instrumento. Un cuestionario puede ser adecuado para comprobar conocimientos específicos, mientras un portafolio, una simulación o una rúbrica resultan más pertinentes para valorar procesos y desempeños complejos. La elección debe responder a la naturaleza de la evidencia requerida y no a funciones predeterminadas.

La validez constituye otro criterio central, pues toda herramienta debe generar información suficientemente representativa del aprendizaje. Gu (2021) sostiene que las interpretaciones evaluativas necesitan argumentos que vinculen evidencias, criterios y decisiones. En entornos digitales, esta exigencia implica revisar si la actividad mide realmente el constructo previsto o si introduce factores ajenos, como rapidez de navegación, familiaridad con la plataforma o

calidad del dispositivo. Una herramienta puede producir resultados precisos en términos técnicos y, al mismo tiempo, ofrecer conclusiones pedagógicas débiles o injustificadas.

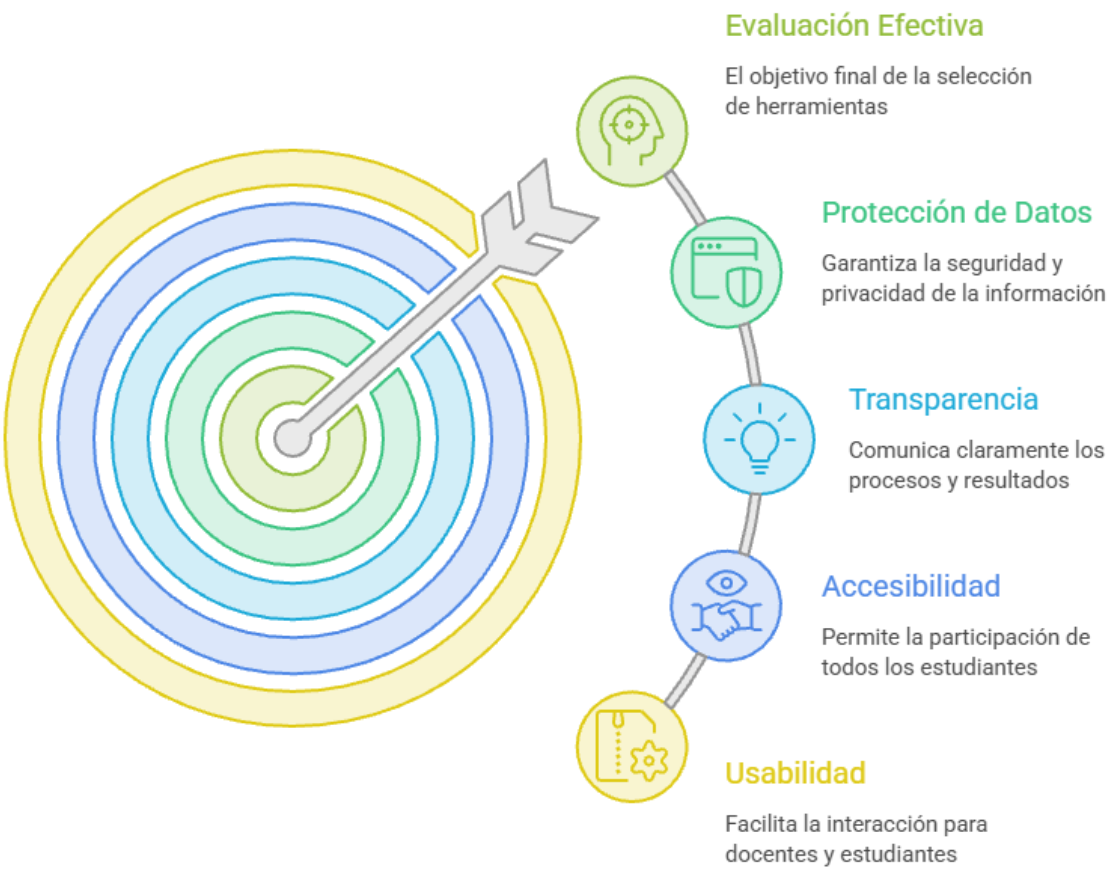
La usabilidad influye directamente en la calidad de la evidencia. Interfaces confusas, instrucciones poco claras o procesos de entrega complejos pueden generar errores que no corresponden al aprendizaje evaluado. Las plataformas deben permitir que docentes y estudiantes comprendan fácilmente cómo acceder, responder, revisar y recibir retroalimentación. La adopción tecnológica se encuentra condicionada por experiencias previas, capacidades institucionales y prioridades educativas (Tsai et al., 2021). Por ello, una herramienta técnicamente avanzada puede resultar poco viable cuando requiere competencias, tiempo o soporte que la institución no está en condiciones de garantizar.

La accesibilidad debe considerarse desde el momento de la selección. Los recursos que dependen de un único formato, presentan barreras de navegación o no son compatibles con tecnologías de apoyo restringen las oportunidades de participación. Scanlon et al. (2021) identifican errores frecuentes de accesibilidad web que afectan a personas con discapacidad. Las directrices del CAST (2024) proponen múltiples formas de participación, representación y expresión. Aplicadas a la evaluación, estas orientaciones permiten diversificar las vías de respuesta sin alterar los criterios esenciales ni reducir las exigencias académicas previstas.

La protección de datos constituye un criterio indispensable cuando las herramientas recopilan información sobre rendimiento, comportamiento, identidad o participación. Korir et al. (2023) muestran que las preocupaciones estudiantiles se relacionan con la recopilación, el uso y la compartición de datos en sistemas de analítica. En Ecuador, estas prácticas deben ajustarse a la Ley Orgánica de Protección de Datos Personales (2021) y a su Reglamento General (2023). La institución debe conocer qué información se almacena, quién accede a ella, dónde se procesa y durante cuánto tiempo permanece disponible.

La transparencia también influye en la legitimidad de la herramienta. Los estudiantes deben conocer qué se evalúa, qué datos se utilizan y cómo se generan las puntuaciones o recomendaciones. Khosravi et al. (2022) destacan la importancia de la inteligencia artificial explicable en educación, especialmente cuando los sistemas automatizados influyen en decisiones académicas. Una herramienta opaca dificulta cuestionar resultados y puede reducir la confianza. La implementación responsable exige comunicar criterios, limitaciones y mecanismos de revisión humana, evitando que la automatización convierta una decisión pedagógica en un resultado aparentemente incuestionable.

Figura 3
Criterios para la selección de herramientas evaluativas



Nota. La figura presenta los criterios de efectividad, protección de datos, transparencia, accesibilidad y usabilidad para seleccionar herramientas evaluativas.

La interoperabilidad representa un criterio técnico con consecuencias pedagógicas. Las herramientas deben integrarse, cuando sea posible, con sistemas de gestión del aprendizaje, repositorios y mecanismos institucionales de registro. La fragmentación entre aplicaciones obliga a duplicar datos, dificulta el seguimiento y aumenta el riesgo de pérdidas o inconsistencias. Liao et al. (2022) señalan que las evaluaciones digitales administradas continuamente requieren vigilancia de calidad. Una integración adecuada facilita trazabilidad y análisis longitudinal, pero debe evitar que la comodidad operativa justifique la recopilación excesiva de información sin propósito educativo claro.

La calidad de la retroalimentación debe evaluarse antes de adoptar una herramienta. No todas las respuestas automáticas ofrecen orientación útil, incluso cuando se entregan de forma inmediata. Banihashem et al. (2022) muestran que la analítica puede fortalecer la retroalimentación cuando produce información pertinente y accionable. Por tanto, conviene revisar si el sistema explica errores, diferencia niveles de desempeño y permite nuevas

oportunidades de mejora. Una herramienta que solo comunica puntuaciones puede agilizar la calificación, pero contribuye poco al aprendizaje cuando no establece vínculos claros entre resultado, criterio y acción posterior.

El costo de implementación no se limita al precio de una licencia. También incluye infraestructura, conectividad, capacitación, soporte, actualización, mantenimiento y tiempo de adaptación. Las desigualdades de preparación digital entre estudiantes e instituciones pueden afectar el acceso y el aprovechamiento de las herramientas (Van de Werfhorst et al., 2022). Mathrani et al. (2022) añaden que la brecha digital comprende factores sociales, educativos y organizativos. La sostenibilidad exige valorar recursos disponibles, continuidad del servicio y alternativas frente a fallas, evitando depender de soluciones difíciles de mantener.

La implementación requiere una fase de prueba, seguimiento y ajuste. Antes de utilizar una herramienta en decisiones de alta consecuencia, conviene realizar aplicaciones piloto, revisar instrucciones, analizar resultados y recoger percepciones de los usuarios. Stewart et al. (2023) muestran que las creencias del profesorado influyen en la comprensión y adopción de sistemas basados en datos. La formación no debe limitarse a funciones técnicas; necesita abordar criterios pedagógicos, ética, accesibilidad y análisis crítico. La observación continua permite corregir problemas antes de que se conviertan en prácticas institucionalizadas.

En conjunto, la selección e implementación de herramientas evaluativas exige articular alineación pedagógica, validez, usabilidad, accesibilidad, privacidad, transparencia, integración, retroalimentación y sostenibilidad. Ningún criterio debe analizarse de forma aislada, porque una herramienta eficiente puede resultar injusta, opaca o pedagógicamente insuficiente. La tecnología adquiere valor cuando fortalece la producción e interpretación de evidencias sin desplazar el juicio profesional. Una decisión responsable no busca la solución más sofisticada, sino aquella que responde mejor a los objetivos, condiciones y necesidades del contexto educativo.

Capítulo

03

Evaluación formativa y retroalimentación en entornos virtuales



3.1 Fundamentos de la evaluación formativa digital

La evaluación formativa digital se sustenta en la recopilación continua de evidencias para comprender el aprendizaje y orientar decisiones durante el proceso educativo. Su finalidad no consiste en acumular calificaciones, sino en identificar avances, dificultades y necesidades de apoyo mientras todavía es posible intervenir. Huauya et al. (2023) relacionan esta modalidad evaluativa con el desarrollo de competencias, debido a que favorece ajustes oportunos en la enseñanza. La tecnología amplía las posibilidades de seguimiento, pero el sentido formativo depende de cómo se interpretan y utilizan los resultados obtenidos.

Uno de sus fundamentos centrales es la alineación entre objetivos, actividades, criterios y evidencias digitales. Las tareas deben producir información relevante sobre el aprendizaje esperado y no limitarse a registrar participación o cumplimiento. La validez formativa exige justificar la relación entre el desempeño observado y las decisiones pedagógicas derivadas de él (Gu, 2021). Por ello, el uso de plataformas, cuestionarios o herramientas interactivas debe responder a propósitos definidos. Una actividad técnicamente atractiva puede carecer de valor evaluativo cuando no permite interpretar progresos ni orientar mejoras específicas.

La continuidad diferencia a la evaluación formativa digital de las prácticas centradas exclusivamente en exámenes finales. Los entornos virtuales permiten recoger datos mediante entregas sucesivas, revisiones, intervenciones, intentos y producciones desarrolladas durante el curso. Liao et al. (2022) advierten que la administración continua de evaluaciones digitales requiere mecanismos para mantener su calidad. La frecuencia de los registros no garantiza una comprensión más profunda del aprendizaje; resulta necesario seleccionar evidencias pertinentes, revisar los instrumentos y evitar que la cantidad de datos sustituya el análisis pedagógico de los procesos.

La retroalimentación constituye el mecanismo mediante el cual la evidencia se transforma en orientación para mejorar. Su valor depende de que sea comprensible, oportuna, específica y relacionada con criterios previamente conocidos. La retroalimentación efectiva debe señalar la distancia entre el desempeño alcanzado y el esperado, además de proponer acciones viables para avanzar (Vera, 2022). Los entornos digitales facilitan comentarios inmediatos, grabaciones, anotaciones y respuestas automatizadas. Sin embargo, la rapidez pierde sentido cuando el mensaje se limita a informar errores sin explicar sus causas ni ofrecer oportunidades de revisión.

La participación del estudiante constituye otro fundamento esencial. La evaluación formativa digital debe facilitar que el alumno comprenda los criterios, examine sus producciones y tome decisiones sobre su propio aprendizaje. Romero et al. (2025) destacan que las rúbricas pueden hacer visibles las expectativas y apoyar procesos de revisión. También pueden emplearse listas de cotejo, portafolios y registros de progreso para fortalecer la autoevaluación. Estas herramientas favorecen la autorregulación únicamente cuando se utilizan de manera reflexiva y no como procedimientos mecánicos destinados a verificar tareas completadas.

Tabla 6

Fundamentos operativos de la evaluación formativa digital

Fundamento	Función dentro del proceso	Recurso digital aplicable	Condición para su efectividad
Alineación pedagógica	Relacionar objetivos, tareas y criterios	Plataformas, rúbricas y cuestionarios	Correspondencia con el aprendizaje esperado
Recogida continua de evidencias	Observar progresos y dificultades	Registros, portafolios y entregas sucesivas	Selección de datos relevantes
Retroalimentación	Orientar acciones de mejora	Comentarios, anotaciones y respuestas automáticas	Claridad, oportunidad y posibilidad de revisión
Participación estudiantil	Favorecer autoevaluación y autorregulación	Rúbricas, listas y paneles de progreso	Comprensión de criterios y acompañamiento
Interpretación contextual	Convertir datos en decisiones pedagógicas	Analítica del aprendizaje	Supervisión docente y contraste de evidencias
Inclusión	Ampliar formas de participación y expresión	Recursos multimodales y formatos alternativos	Accesibilidad desde el diseño
Responsabilidad ética	Proteger información y limitar usos indebidos	Sistemas de gestión y analítica	Transparencia y control de acceso

Nota. La tabla organiza los fundamentos que permiten distinguir una práctica formativa digital de una simple recopilación tecnológica de resultados.

La tabla evidencia que la evaluación formativa digital exige integrar dimensiones pedagógicas, tecnológicas y éticas. Ningún recurso funciona de manera aislada: la retroalimentación necesita criterios, la analítica requiere interpretación y la participación estudiantil depende de orientaciones comprensibles. Banihashem et al. (2022) muestran que la analítica puede fortalecer las prácticas de retroalimentación cuando produce información útil sobre el progreso. No obstante, los indicadores generados por las plataformas deben contrastarse con

producciones y observaciones, porque la actividad registrada no equivale necesariamente a aprendizaje demostrado.

La personalización constituye una posibilidad relevante de la evaluación formativa digital. Los sistemas adaptativos pueden ajustar actividades, niveles de dificultad y apoyos según las respuestas del estudiante. Esta capacidad favorece trayectorias diferenciadas cuando los criterios de adaptación son transparentes y pedagógicamente justificados (Aparicio y Aparicio, 2024). Sin embargo, la personalización no debe encerrar al estudiante en rutas predeterminadas ni reducir las oportunidades de afrontar tareas complejas. Su finalidad consiste en ofrecer apoyos oportunos, no en sustituir el acompañamiento docente ni automatizar decisiones sin revisión humana.

La inclusión también forma parte de sus fundamentos, ya que una evaluación solo puede orientar el aprendizaje cuando todos los estudiantes tienen oportunidades reales de participar. Las directrices del CAST (2024) proponen múltiples formas de implicación, representación y expresión, principios aplicables al diseño de evidencias digitales. La accesibilidad permite diversificar formatos, tiempos y medios de respuesta sin modificar injustificadamente los criterios académicos. Cuando una plataforma presenta barreras de navegación o depende de recursos inaccesibles, los resultados pueden reflejar dificultades técnicas antes que diferencias auténticas en el aprendizaje.

La dimensión ética completa este enfoque al exigir transparencia sobre los datos recopilados y las decisiones que se toman a partir de ellos. Las preocupaciones estudiantiles incluyen la recopilación, utilización y compartición de información generada en procesos de analítica (Korir et al., 2023). La evaluación formativa digital debe limitarse a datos pertinentes, proteger su acceso y explicar su finalidad. Su calidad depende, en definitiva, de articular evidencia, retroalimentación, participación, accesibilidad y supervisión docente, de modo que la tecnología fortalezca la mejora continua sin desplazar el juicio pedagógico.

3.2 Diagnóstico y seguimiento continuo del aprendizaje

El diagnóstico y el seguimiento continuo del aprendizaje permiten comprender el punto de partida del estudiante y observar sus transformaciones a lo largo del proceso formativo. En entornos virtuales, ambas funciones se apoyan en evidencias producidas mediante cuestionarios, tareas, registros de interacción y actividades secuenciadas. La evaluación formativa adquiere valor cuando utiliza esa información para ajustar la enseñanza y fortalecer

el desarrollo de competencias (Huauya et al., 2023). Por ello, diagnosticar no significa clasificar, sino identificar condiciones que orienten decisiones pedagógicas pertinentes.

La evaluación diagnóstica debe explorar conocimientos previos, dificultades, estrategias y condiciones de participación antes de iniciar una secuencia de aprendizaje. Esta información permite adaptar contenidos, apoyos y niveles de complejidad sin convertir los resultados iniciales en juicios definitivos. Gu (2021) sostiene que las inferencias evaluativas deben relacionar evidencias e interpretaciones mediante argumentos válidos. En consecuencia, un diagnóstico digital requiere instrumentos coherentes con el aprendizaje esperado y cautela frente a factores técnicos que puedan afectar el desempeño sin representar realmente el nivel académico.

Los cuestionarios en línea facilitan diagnósticos rápidos mediante preguntas cerradas, respuestas breves y retroalimentación automática. Su principal aporte consiste en ofrecer información inicial sobre conceptos, errores frecuentes y niveles de comprensión. Sin embargo, estos instrumentos resultan insuficientes cuando se pretende valorar razonamiento, argumentación o aplicación contextualizada. La evaluación en línea debe integrar formatos diversos para evitar que la interpretación se reduzca a puntuaciones obtenidas en una sola prueba (Heil y Ifenthaler, 2023). La triangulación de evidencias fortalece la precisión del diagnóstico.

El seguimiento continuo amplía la mirada hacia los cambios que se producen durante el aprendizaje. Las plataformas permiten registrar entregas, revisiones, tiempos, intentos y participación en diferentes actividades. Liao et al. (2022) advierten que la administración continua de evaluaciones digitales requiere mecanismos para mantener la calidad de los instrumentos y de las interpretaciones. La frecuencia de los registros no garantiza un seguimiento útil; los datos deben seleccionarse de acuerdo con objetivos claros y analizarse dentro de las condiciones concretas en que fueron generados.

La analítica del aprendizaje ofrece recursos para organizar grandes volúmenes de información y reconocer patrones de progreso o dificultad. Muñoz et al. (2021) explican que estas tecnologías permiten examinar datos procedentes de la actividad académica. No obstante, los indicadores de acceso, permanencia o frecuencia de interacción no equivalen directamente a comprensión. Un seguimiento responsable debe contrastar registros técnicos con productos, respuestas y observaciones cualitativas. La interpretación pedagógica evita convertir

comportamientos digitales aislados en conclusiones definitivas sobre capacidad, compromiso o logro educativo.

La retroalimentación convierte el seguimiento en una intervención formativa. Cuando los resultados se comunican de manera clara y oportuna, el estudiante puede reconocer brechas y ajustar sus estrategias. La analítica puede fortalecer estas prácticas al proporcionar información sobre avances y dificultades recurrentes (Banihashem et al., 2022). Sin embargo, los reportes automáticos deben traducirse en orientaciones comprensibles y acciones posibles. Mostrar gráficos o porcentajes sin explicación puede informar sobre el rendimiento, pero no necesariamente ayuda a comprender qué debe cambiar ni cómo realizar esa mejora.

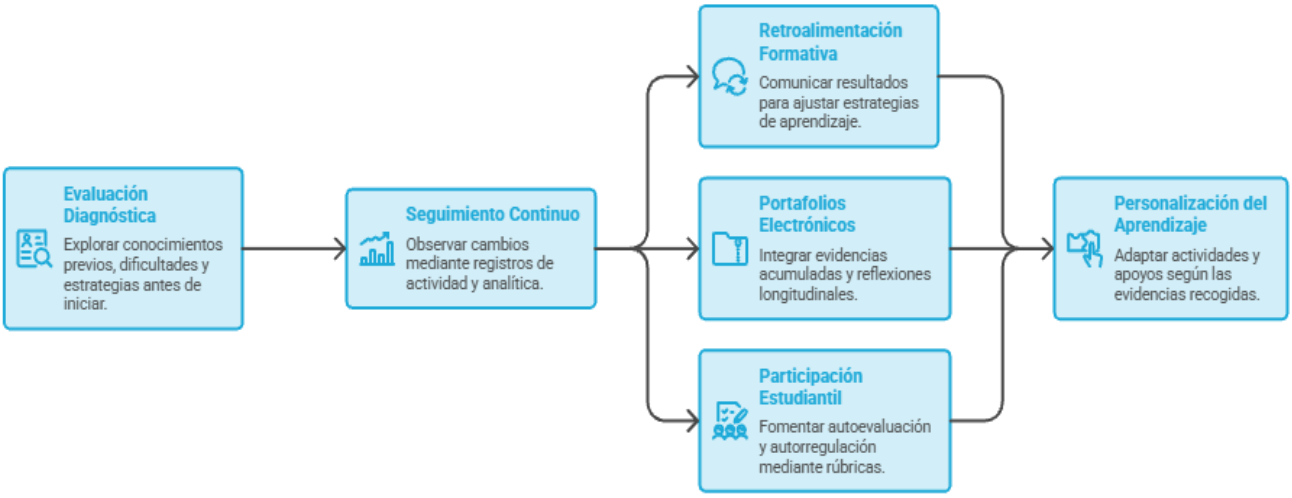
Los portafolios electrónicos permiten observar el desarrollo del aprendizaje mediante evidencias acumuladas y versiones sucesivas. Zhang y Tur (2024) destacan que estos recursos integran productos, reflexiones y muestras de progreso dentro de una estructura longitudinal. Su uso favorece un seguimiento más profundo que el ofrecido por registros de actividad aislados. Para cumplir esta función, el portafolio debe incluir criterios claros, selección intencional de evidencias y espacios de reflexión. La acumulación indiscriminada de archivos dificulta identificar cambios significativos y reduce su utilidad evaluativa.

La participación del estudiante fortalece el seguimiento cuando incluye autoevaluación, revisión de criterios y análisis del propio progreso. Las rúbricas digitales pueden hacer visibles los niveles de desempeño y facilitar comparaciones entre producciones sucesivas (Romero et al., 2025). Este proceso favorece la autorregulación porque el estudiante identifica avances y decide qué aspectos necesita revisar. No obstante, la autoevaluación requiere acompañamiento y ejemplos de calidad. Sin orientación, puede convertirse en una práctica superficial o en una repetición mecánica de categorías definidas por el docente.

La personalización constituye una consecuencia posible del diagnóstico y del seguimiento continuo. Los sistemas adaptativos pueden modificar actividades, apoyos y niveles de dificultad según las evidencias recogidas. Aparicio y Aparicio (2024) relacionan estas tecnologías con propuestas de aprendizaje más ajustadas a las necesidades del estudiante. Su implementación debe evitar decisiones rígidas o trayectorias cerradas basadas en resultados iniciales. El diagnóstico debe abrir oportunidades de apoyo, no limitar expectativas. Toda adaptación automatizada requiere revisión docente y posibilidad de modificar las rutas sugeridas.

Figura 4

Proceso de diagnóstico y seguimiento del aprendizaje



Nota. La figura representa las etapas del diagnóstico, seguimiento, retroalimentación y personalización del aprendizaje.

La equidad condiciona la validez del seguimiento digital. Las diferencias de conectividad, dispositivos, familiaridad tecnológica y apoyo institucional pueden influir en los registros disponibles. La preparación digital desigual afecta las oportunidades de participación y puede distorsionar la interpretación del desempeño (Van de Werfhorst et al., 2022). Mathrani et al. (2022) muestran que la brecha digital incluye dimensiones sociales y educativas. Por ello, el seguimiento debe considerar ausencias de datos, interrupciones técnicas y formas alternativas de demostrar aprendizaje antes de atribuir resultados bajos a falta de compromiso.

El diagnóstico y el seguimiento continuo cumplen una función formativa cuando producen información útil, contextualizada y vinculada con decisiones concretas. Su eficacia depende de combinar instrumentos, registros, producciones y retroalimentación dentro de una secuencia coherente. La tecnología permite observar trayectorias con mayor detalle, pero también puede generar interpretaciones reduccionistas si se confunde actividad con aprendizaje. La supervisión docente, la participación del estudiante y la atención a la equidad resultan indispensables para convertir los datos en apoyos reales y no en mecanismos permanentes de clasificación.

3.3 Retroalimentación efectiva, oportuna y personalizada

La retroalimentación constituye uno de los componentes más relevantes de la evaluación formativa porque permite transformar los resultados en orientaciones para mejorar el aprendizaje. Su efectividad depende de que la información proporcionada sea clara, específica

y vinculada con criterios conocidos. Vera (2022) sostiene que la retroalimentación favorece el aprendizaje cuando ayuda al estudiante a reconocer errores, comprender sus causas y tomar decisiones de mejora. Por ello, no debe limitarse a comunicar una calificación ni a señalar deficiencias sin explicar cómo puede fortalecerse el desempeño.

Una retroalimentación efectiva debe establecer una relación comprensible entre el aprendizaje esperado, la evidencia producida y las acciones necesarias para avanzar. Los comentarios generales, como "debe mejorar" o "trabajo incompleto", aportan poca información para orientar una revisión. La utilidad formativa aumenta cuando se identifican aspectos concretos del desempeño y se proponen pasos alcanzables para corregirlos (García, 2024). Este tipo de orientación permite que el estudiante interprete el resultado, organice sus esfuerzos y comprenda que la evaluación forma parte del proceso de aprendizaje.

La oportunidad representa una condición esencial, porque la retroalimentación pierde valor cuando se entrega después de que la actividad ha concluido o cuando ya no existe posibilidad de revisar el trabajo. Banihashem et al. (2022) destacan que la analítica del aprendizaje puede fortalecer las prácticas de retroalimentación al facilitar información sobre avances y dificultades. En los entornos virtuales, las plataformas permiten reducir el tiempo entre la producción de la evidencia y la respuesta. Sin embargo, rapidez no significa calidad si el mensaje es superficial o descontextualizado.

La personalización implica adaptar la retroalimentación a las necesidades, dificultades y trayectorias del estudiante. No consiste en redactar mensajes completamente diferentes para cada persona, sino en utilizar evidencias específicas para ofrecer orientaciones pertinentes. Los sistemas adaptativos pueden ajustar contenidos, actividades y apoyos según el desempeño registrado (Aparicio y Aparicio, 2024). Esta capacidad resulta útil cuando evita respuestas uniformes ante errores distintos. No obstante, toda personalización automatizada debe ser revisada por el docente para impedir interpretaciones rígidas o recomendaciones inadecuadas.

La retroalimentación automatizada ha ampliado las posibilidades de responder de manera inmediata a grandes grupos de estudiantes. De Schipper et al. (2021) analizan sistemas capaces de ofrecer comentarios personalizados dentro de evaluaciones sumativas. Estas herramientas pueden identificar patrones de error y recomendar recursos o actividades complementarias. Su eficacia depende de la calidad de los criterios y de la precisión con que se relacionan las respuestas con las sugerencias. Cuando el sistema solo informa si una respuesta es correcta, la retroalimentación conserva un carácter limitado y escasamente formativo.

En tareas complejas, la retroalimentación automatizada puede apoyar procesos de revisión, aunque no debe sustituir completamente el juicio pedagógico. Los sistemas basados en aprendizaje automático han mostrado capacidad para generar comentarios sobre argumentaciones científicas y producciones escritas (Lee et al., 2021; Wilson et al., 2021). Estas aplicaciones facilitan la identificación de aspectos recurrentes y permiten realizar varios intentos. Sin embargo, su utilidad disminuye cuando no comprenden matices, contexto o intenciones comunicativas, por lo que requieren complementarse con intervención humana.

La retroalimentación sobre escritura académica constituye un ámbito especialmente sensible, debido a que la calidad de un texto depende de contenido, estructura, coherencia y adecuación discursiva. Ossa y Willatt (2023) examinan el uso de inteligencia artificial generativa para retroalimentar procesos de escritura en la formación docente. Estas herramientas pueden proponer cambios y señalar problemas, pero sus recomendaciones deben analizarse críticamente. El estudiante necesita comprender por qué una modificación mejora el texto y no limitarse a aceptar correcciones automáticas sin reflexión ni apropiación del aprendizaje.

La participación activa del estudiante determina si la retroalimentación produce efectos reales. Recibir comentarios no garantiza que estos sean comprendidos, aceptados o utilizados. La evaluación formativa debe generar oportunidades para revisar, responder y demostrar mejoras a partir de la orientación recibida (Huauya et al., 2023). Por esta razón, los entornos virtuales deben incluir ciclos de entrega, retroalimentación y nueva producción. Cuando la calificación se registra como resultado definitivo, el comentario pierde parte de su función y se convierte en una explicación tardía del desempeño.

Las rúbricas digitales pueden fortalecer la retroalimentación al mostrar criterios y niveles de desempeño de forma estructurada. Romero et al. (2025) destacan su presencia en prácticas de evaluación formativa, especialmente por su capacidad para explicitar expectativas. Sin embargo, seleccionar un nivel dentro de una rúbrica no siempre ofrece suficiente información para mejorar. La orientación debe complementarse con comentarios que expliquen la elección, identifiquen prioridades y señalen acciones concretas. Así, la rúbrica deja de ser un mecanismo de puntuación y se convierte en una guía para la revisión.

La personalización también debe considerar la accesibilidad, el lenguaje y las condiciones de participación. Un mensaje técnicamente correcto puede resultar poco útil si utiliza expresiones complejas, formatos incompatibles o indicaciones difíciles de aplicar. Las directrices del CAST (2024) proponen diversificar las formas de representación y comunicación, lo que permite

ofrecer comentarios escritos, orales o audiovisuales. Esta variedad favorece la comprensión sin modificar los criterios académicos. La retroalimentación inclusiva elimina barreras comunicativas y permite que cada estudiante acceda a orientaciones en condiciones más equitativas.

La retroalimentación efectiva, oportuna y personalizada requiere articular criterios, evidencias, comunicación y oportunidades de revisión. Las tecnologías digitales pueden acelerar la respuesta, identificar patrones y adaptar mensajes, pero no garantizan por sí mismas una mejora del aprendizaje. Su valor depende de que los comentarios sean comprensibles, pertinentes y aplicables dentro de una secuencia formativa. La supervisión docente continúa siendo indispensable para contextualizar resultados, corregir recomendaciones y asegurar que la retroalimentación fortalezca autonomía, reflexión y capacidad de mejorar progresivamente.

3.4 Autoevaluación y desarrollo de la autorregulación

La autoevaluación constituye una práctica mediante la cual el estudiante analiza su propio desempeño con base en criterios previamente definidos. En entornos virtuales, esta función puede apoyarse en rúbricas, listas de cotejo, portafolios y paneles de progreso. Su finalidad no consiste en sustituir la valoración docente, sino en fortalecer la capacidad para reconocer avances, identificar dificultades y tomar decisiones sobre el aprendizaje. La evaluación formativa favorece este proceso cuando convierte la evidencia en una oportunidad de reflexión y ajuste continuo (Huauya et al., 2023).

La autorregulación se relaciona con la capacidad de planificar, supervisar y modificar estrategias durante el aprendizaje. La autoevaluación contribuye a este desarrollo porque obliga a comparar el desempeño alcanzado con metas y criterios explícitos. Martínez et al. (2025) vinculan la gamificación con el fortalecimiento de la autorregulación, especialmente cuando los estudiantes pueden observar su progreso y responder a retos graduales. Sin embargo, la visibilidad de resultados solo resulta formativa cuando orienta decisiones conscientes y no se reduce a acumular puntos, niveles o recompensas externas.

Los criterios comprensibles constituyen una condición indispensable para que el estudiante pueda valorar su trabajo con cierta precisión. Las rúbricas digitales permiten hacer visibles dimensiones, niveles y expectativas de calidad, facilitando la revisión antes de una entrega. Su utilidad dentro de la evaluación formativa ha sido destacada por Romero et al. (2025). No obstante, una rúbrica excesivamente técnica o rígida puede dificultar la reflexión. Para

favorecer la autorregulación, sus descriptores deben permitir identificar diferencias reales entre niveles y traducirse en acciones concretas de mejora.

Las listas de cotejo también pueden apoyar la autoevaluación cuando se utilizan para verificar pasos, requisitos o componentes esenciales de una tarea. Su estructura simple facilita la revisión autónoma, aunque ofrece poca información sobre la calidad del desempeño. Por ello, conviene combinarlas con escalas o comentarios reflexivos. La autoevaluación pierde profundidad cuando se limita a marcar elementos cumplidos sin analizar por qué fueron alcanzados ni qué aspectos requieren fortalecimiento. El instrumento debe provocar interpretación y no únicamente control mecánico de actividades terminadas.

Los portafolios electrónicos ofrecen condiciones favorables para desarrollar autorregulación, porque permiten conservar versiones, comentarios y evidencias acumuladas. Zhang y Tur (2024) muestran que estos recursos facilitan la reflexión sobre procesos y progresos a lo largo del tiempo. Al seleccionar muestras, justificar su inclusión y comparar producciones, el estudiante construye una visión más amplia de su trayectoria. Esta práctica fortalece la conciencia sobre estrategias eficaces, errores recurrentes y cambios alcanzados, siempre que el portafolio se organice con objetivos claros y no como simple depósito de archivos.

La retroalimentación desempeña una función decisiva en la calidad de la autoevaluación. El estudiante necesita contrastar su propio juicio con observaciones externas que le permitan ajustar interpretaciones poco precisas. Vera (2022) sostiene que la retroalimentación resulta efectiva cuando ofrece información comprensible y orientada a la mejora. En este proceso, el docente no impone únicamente una valoración, sino que ayuda a identificar evidencias, justificar criterios y reconocer prioridades. La interacción entre retroalimentación y autoevaluación favorece juicios progresivamente más consistentes sobre el propio aprendizaje.

Tabla 7
Relación entre autoevaluación digital y desarrollo de la autorregulación

Proceso de autorregulación	Acción de autoevaluación	Herramienta digital aplicable	Resultado esperado	Riesgo de uso superficial
Planificación	Revisar metas y criterios antes de iniciar	Rúbrica o lista de cotejo	Organización de estrategias y tiempos	Cumplimiento mecánico sin comprensión
Supervisión	Comparar avances con el desempeño esperado	Panel de progreso o registro de actividad	Detección temprana de dificultades	Confundir actividad con aprendizaje

Proceso de autorregulación	Acción de autoevaluación	Herramienta digital aplicable	Resultado esperado	Riesgo de uso superficial
Reflexión	Analizar fortalezas, errores y decisiones	Portafolio electrónico o diario digital	Comprensión de la propia trayectoria	Descripción sin análisis
Ajuste	Modificar estrategias y recursos	Retroalimentación digital y nuevos intentos	Mejora progresiva del desempeño	Repetición sin cambio de estrategia
Verificación	Contrastar el juicio propio con criterios externos	Rúbrica compartida o coevaluación	Mayor precisión evaluativa	Dependencia de la opinión de otros
Proyección	Establecer metas posteriores	Plan de mejora digital	Continuidad del aprendizaje autónomo	Formulación de metas poco realistas

Nota. La tabla relaciona las fases de la autorregulación con acciones concretas de autoevaluación y herramientas que pueden apoyarlas en entornos virtuales.

La tabla muestra que la autoevaluación puede intervenir en distintas fases del aprendizaje y no únicamente al finalizar una actividad. Su aporte se fortalece cuando acompaña la planificación, el seguimiento, la reflexión y el ajuste de estrategias. Los registros digitales hacen visible parte de este proceso, pero deben interpretarse con cautela. La actividad almacenada en una plataforma no equivale necesariamente a comprensión, por lo que los indicadores requieren contraste con evidencias cualitativas y juicios contextualizados (Muñoz et al., 2021).

La coevaluación puede complementar la autoevaluación al ofrecer perspectivas diferentes sobre una misma producción. Santana et al. (2025) relacionan la evaluación entre pares con el desarrollo de competencias transversales en proyectos integradores. Comparar criterios, justificar observaciones y recibir comentarios de compañeros permite revisar el propio juicio. No obstante, esta práctica necesita orientación para evitar valoraciones basadas en afinidades, impresiones generales o reciprocidad. La autorregulación se fortalece cuando el estudiante aprende a aceptar, cuestionar y utilizar críticamente la información recibida.

Las herramientas digitales pueden personalizar apoyos y sugerir actividades según el desempeño registrado. Los sistemas adaptativos ofrecen rutas diferenciadas que pueden contribuir a la toma de decisiones autónomas (Aparicio y Aparicio, 2024). Sin embargo, la autorregulación no debe reducirse a seguir recomendaciones automáticas. El estudiante necesita comprender por qué una actividad resulta pertinente, evaluar su utilidad y decidir cómo incorporarla. La supervisión docente continúa siendo necesaria para evitar trayectorias rígidas, dependencia tecnológica o interpretaciones automatizadas que limiten oportunidades de aprendizaje más complejas.

En conjunto, la autoevaluación fortalece la autorregulación cuando permite planificar, supervisar, interpretar y ajustar el aprendizaje con base en evidencias. Las herramientas digitales amplían las oportunidades para visualizar progresos, conservar producciones y recibir orientación, pero no garantizan reflexión por sí mismas. Su valor depende de criterios claros, retroalimentación, acompañamiento y oportunidades reales de revisión. Una práctica bien diseñada favorece autonomía y responsabilidad; una aplicación mecánica puede convertir la autoevaluación en un trámite sin influencia efectiva sobre las decisiones del estudiante.

3.5 Coevaluación y aprendizaje colaborativo en espacios virtuales

La coevaluación es una práctica mediante la cual los estudiantes valoran el trabajo de sus compañeros con base en criterios previamente establecidos. En espacios virtuales, puede desarrollarse mediante foros, documentos compartidos, rúbricas digitales y plataformas de revisión entre pares. Su propósito no consiste solo en distribuir la tarea evaluativa, sino en promover análisis, argumentación y comprensión de los estándares de calidad. La evaluación entre pares puede contribuir al desarrollo de competencias transversales cuando se integra en proyectos y actividades colaborativas (Santana et al., 2025).

El aprendizaje colaborativo se fortalece cuando la evaluación acompaña los procesos de interacción y no se limita a calificar un producto final. Ruiz y del Moral (2022) relacionan los entornos digitales con metodologías activas y competencias socioemocionales, lo que permite comprender la colaboración como una experiencia académica y relacional. En este marco, la coevaluación ayuda a reconocer aportes, contrastar perspectivas y discutir decisiones. No obstante, su valor depende de que existan objetivos compartidos, responsabilidades definidas y oportunidades para revisar el trabajo a partir de los comentarios recibidos.

La claridad de los criterios constituye una condición básica para evitar valoraciones basadas en preferencias personales o impresiones generales. Las rúbricas digitales permiten organizar dimensiones y niveles de desempeño, facilitando que los estudiantes justifiquen sus juicios mediante evidencias observables. Su uso puede fortalecer la evaluación formativa cuando las expectativas son comprensibles y se vinculan con posibilidades de mejora (Romero et al., 2025). Una rúbrica ambigua, en cambio, puede aumentar las diferencias entre evaluadores y producir comentarios poco consistentes, aun cuando la plataforma registre todas las puntuaciones.

La calidad de la coevaluación también depende de la preparación de los participantes. Antes de valorar a sus compañeros, los estudiantes necesitan analizar ejemplos, discutir criterios y

practicar la formulación de comentarios. La retroalimentación resulta útil cuando identifica aspectos concretos y propone acciones viables para mejorar (Vera, 2022). Sin esta preparación, los mensajes pueden reducirse a elogios generales, correcciones superficiales o juicios excesivamente severos. La formación previa permite comprender que evaluar implica argumentar y no simplemente expresar una opinión sobre el trabajo de otra persona.

Los espacios virtuales amplían la posibilidad de realizar coevaluaciones asincrónicas. Este formato ofrece tiempo para revisar producciones, consultar criterios y redactar observaciones más elaboradas. A diferencia de la interacción inmediata, la asincronía puede favorecer una participación reflexiva y documentada. Sin embargo, también puede generar retrasos, respuestas breves o falta de continuidad cuando no se establecen plazos y responsabilidades. La evaluación en línea requiere diseños claros que articulen tiempos, instrumentos y formas de participación coherentes con los propósitos pedagógicos (Heil y Ifenthaler, 2023).

La coevaluación sincrónica, desarrollada mediante videoconferencias o sesiones en tiempo real, permite discutir decisiones y aclarar interpretaciones de forma inmediata. Esta modalidad resulta pertinente para presentaciones, defensas orales y revisión conjunta de proyectos. Su principal fortaleza radica en el intercambio directo, aunque puede favorecer la intervención de estudiantes con mayor seguridad comunicativa. Las condiciones de conectividad y participación también influyen en la calidad de la interacción, por lo que las diferencias de preparación digital deben considerarse al interpretar el desempeño (Van de Werfhorst et al., 2022).

La colaboración virtual produce evidencias sobre comunicación, coordinación, cumplimiento de responsabilidades y construcción conjunta de conocimiento. No obstante, estas evidencias deben diferenciarse del simple registro de actividad. La frecuencia de mensajes o accesos no demuestra necesariamente un aporte académico significativo. La analítica del aprendizaje puede organizar información sobre interacciones, pero requiere interpretación contextual y contraste con productos concretos (Muñoz et al., 2021). Evaluar la colaboración exige examinar la calidad de las contribuciones, su relación con los objetivos y su influencia en el resultado colectivo.

La valoración de aportes individuales representa uno de los principales desafíos del aprendizaje colaborativo. Un producto grupal puede ocultar distribuciones desiguales del trabajo o dificultades para reconocer quién desarrolló determinadas tareas. La coevaluación permite obtener información sobre responsabilidades y contribuciones, aunque puede verse afectada

por amistad, conflicto o reciprocidad. Santana et al. (2025) muestran que la evaluación entre pares adquiere mayor valor cuando se integra con competencias transversales y proyectos estructurados. Conviene combinarla con autoevaluación, evidencias del proceso y valoración docente.

La retroalimentación entre pares puede fortalecer el aprendizaje de quien recibe comentarios y también de quien los formula. Al revisar una producción ajena, el estudiante debe aplicar criterios, identificar fortalezas y explicar posibles mejoras. Esta actividad favorece la comprensión de la calidad y puede trasladarse posteriormente al propio trabajo. La evaluación formativa promueve desarrollo de competencias cuando genera reflexión y ajuste durante el proceso (Huauya et al., 2023). Para que esto ocurra, los comentarios deben ser analizados y utilizados, no simplemente entregados como requisito de participación.

Figura 5
¿Qué modalidad de coevaluación virtual se ajusta mejor a los objetivos de mi actividad pedagógica?



Nota. La figura compara las características, ventajas y limitaciones de la coevaluación virtual asincrónica y sincrónica.

Las herramientas digitales facilitan la asignación de revisores, el anonimato, el registro de comentarios y la comparación de valoraciones. El anonimato puede reducir ciertas presiones sociales, aunque también puede disminuir la responsabilidad si no existen normas claras. Las plataformas deben garantizar que la identidad, los trabajos y las puntuaciones se gestionen de

forma adecuada. Las preocupaciones estudiantiles sobre recopilación y compartición de datos aumentan cuando las interacciones quedan almacenadas y pueden utilizarse con fines no explicados (Korir et al., 2023).

La inclusión exige que la coevaluación contemple diferentes formas de participación, comunicación y acceso. Algunos estudiantes pueden requerir más tiempo, formatos alternativos o instrucciones adicionales para revisar y comentar producciones. Las directrices del CAST (2024) promueven múltiples formas de implicación, representación y expresión, principios aplicables al trabajo colaborativo. Esta flexibilidad no debe eliminar los criterios comunes, sino garantizar que las barreras tecnológicas o comunicativas no impidan demostrar capacidades. Un diseño accesible amplía la participación y mejora la diversidad de perspectivas dentro del grupo.

En conjunto, la coevaluación fortalece el aprendizaje colaborativo cuando se apoya en criterios claros, preparación previa, retroalimentación argumentada y oportunidades de revisión. Los espacios virtuales amplían la interacción y conservan evidencias del proceso, pero también pueden generar superficialidad, desigualdad de aportes y problemas de privacidad. Su calidad depende de combinar valoración entre pares, autoevaluación y juicio docente dentro de un diseño coherente. Más que transferir la calificación al estudiante, la coevaluación busca desarrollar capacidad para interpretar, justificar y mejorar el trabajo académico compartido.

3.6 Evaluación auténtica mediante proyectos y resolución de problemas

La evaluación auténtica mediante proyectos y resolución de problemas busca valorar la capacidad del estudiante para aplicar conocimientos en situaciones contextualizadas, tomar decisiones y producir respuestas con sentido. A diferencia de las pruebas centradas en reproducción de contenidos, este enfoque observa cómo se integran saberes, habilidades y criterios durante una actuación compleja. García (2024) relaciona la evaluación auténtica con la mejora del aprendizaje, porque permite examinar desempeños próximos a situaciones reales. Su valor depende de que las tareas planteen desafíos relevantes y coherentes con los objetivos formativos.

Los proyectos permiten recopilar evidencias a lo largo de diferentes fases, desde la planificación hasta la presentación del producto final. Esta secuencia facilita valorar búsqueda de información, organización, toma de decisiones, revisión y comunicación de resultados. La evaluación formativa se fortalece cuando las evidencias obtenidas durante el proceso se utilizan para orientar ajustes y no únicamente para asignar una calificación final (Huauya et al., 2023).

Por ello, el proyecto debe incluir momentos de seguimiento, criterios explícitos y oportunidades para modificar el trabajo antes de su cierre.

La resolución de problemas exige interpretar una situación, identificar variables, seleccionar estrategias y justificar una respuesta. Estas tareas permiten observar razonamiento y transferencia, siempre que el problema no se reduzca a aplicar un procedimiento memorizado. Gu (2021) sostiene que la validez evaluativa depende de la relación entre evidencias, interpretaciones y decisiones. En este marco, el docente debe definir qué actuaciones demuestran comprensión, qué errores revelan dificultades relevantes y cómo se distinguirán respuestas parcialmente correctas de soluciones fundamentadas y coherentes.

Los entornos virtuales amplían las posibilidades de desarrollar proyectos mediante documentos compartidos, foros, portafolios, simulaciones y producciones multimedia. La diversidad de formatos permite reunir evidencias sobre procesos y resultados que no siempre son visibles en una prueba convencional (Zhang y Tur, 2024). Sin embargo, la tecnología no convierte automáticamente una tarea en auténtica. Un proyecto digital puede seguir siendo superficial si se limita a recopilar información o reproducir contenidos sin análisis, toma de decisiones ni conexión con problemas significativos para el aprendizaje.

Las rúbricas resultan especialmente pertinentes para valorar proyectos y problemas complejos, porque permiten organizar criterios como calidad de la solución, argumentación, aplicación de conocimientos, colaboración y comunicación. Romero et al. (2025) destacan su utilidad dentro de la evaluación formativa, dado que hacen visibles las expectativas. No obstante, sus descriptores deben evitar formulaciones vagas o excesivamente fragmentadas. Una rúbrica útil no divide el desempeño en elementos desconectados, sino que ayuda a interpretar la coherencia entre el proceso seguido, la evidencia presentada y la calidad del resultado alcanzado.

Tabla 8
Dimensiones evaluativas de los proyectos y la resolución de problemas

Dimensión	Evidencia esperada	Instrumento pertinente	Aporte formativo	Riesgo evaluativo
Comprensión del problema	Identificación de condiciones, necesidades y restricciones	Guía de análisis o rúbrica	Evita respuestas descontextualizadas	Valorar solo la solución final
Planificación	Metas, etapas, recursos y responsabilidades	Cronograma o lista de cotejo	Facilita seguimiento y organización	Confundir planificación formal con calidad

Dimensión	Evidencia esperada	Instrumento pertinente	Aporte formativo	Riesgo evaluativo
Aplicación de conocimientos	Uso pertinente de conceptos y procedimientos	Rúbrica analítica	Permite observar transferencia	Premiar acumulación de información
Toma de decisiones	Selección y justificación de alternativas	Registro reflexivo o defensa oral	Evidencia razonamiento	Aceptar decisiones sin fundamento
Revisión y mejora	Cambios entre versiones y respuesta a comentarios	Portafolio electrónico	Favorece autorregulación	Calificar únicamente el producto final
Comunicación de resultados	Presentación clara y sustentada	Escala de valoración	Integra dominio y expresión	Sobrevalorar diseño o fluidez
Colaboración	Distribución de aportes y construcción conjunta	Coevaluación y registro del proceso	Fortalece competencias transversales	Ocultar desigualdad de contribuciones

Nota. La tabla organiza las principales dimensiones que deben considerarse al evaluar proyectos y resolución de problemas, diferenciando evidencias, instrumentos y riesgos de interpretación.

La tabla muestra que la autenticidad no depende únicamente del contexto de la tarea, sino de la calidad de las evidencias recogidas en cada fase. Evaluar solo el producto final impide comprender cómo se planificó, qué decisiones se tomaron y qué cambios surgieron durante el proceso. Los portafolios electrónicos permiten documentar versiones, comentarios y reflexiones, fortaleciendo una valoración longitudinal (Zhang y Tur, 2024). Esta trazabilidad facilita distinguir entre un resultado elaborado de manera progresiva y una producción presentada sin evidencia suficiente de desarrollo.

La retroalimentación ocupa un lugar central porque los proyectos y problemas suelen requerir múltiples revisiones. Vera (2022) señala que la orientación resulta efectiva cuando identifica aspectos específicos y propone acciones de mejora. En este tipo de evaluación, los comentarios deben referirse tanto al producto como a las estrategias empleadas. Una observación centrada solo en errores formales puede desviar la atención del razonamiento. La retroalimentación debe ayudar a reconsiderar decisiones, fortalecer argumentos y revisar la pertinencia de la solución frente a las condiciones iniciales.

La coevaluación puede aportar información relevante sobre la calidad del trabajo colaborativo y la distribución de responsabilidades. Los proyectos integradores favorecen competencias transversales cuando incorporan evaluación entre pares y criterios compartidos (Santana et al., 2025). Sin embargo, las valoraciones estudiantiles pueden verse afectadas por afinidades,

conflictos o desconocimiento de los estándares. Para reducir estos riesgos, conviene combinar coevaluación, autoevaluación, registros de participación y juicio docente, evitando que la calificación grupal oculte diferencias significativas entre las contribuciones individuales.

Las simulaciones y escenarios virtuales fortalecen la resolución de problemas al permitir experimentar con decisiones y observar consecuencias dentro de condiciones controladas. La simulación virtual puede generar experiencias formativas valiosas cuando representa situaciones complejas y ofrece oportunidades de repetición (Edgar et al., 2022). No obstante, la interpretación del desempeño debe distinguir entre conocimiento disciplinar y familiaridad con la interfaz. La validez exige que las acciones registradas por el sistema correspondan realmente con las competencias evaluadas y que los resultados puedan revisarse mediante criterios comprensibles.

La evaluación auténtica mediante proyectos y resolución de problemas permite integrar conocimiento, razonamiento, colaboración, comunicación y autorregulación. Su aplicación exige tareas significativas, criterios claros, seguimiento continuo y evidencias suficientes del proceso. Las tecnologías digitales amplían las posibilidades de documentación, interacción y retroalimentación, pero también pueden privilegiar aspectos visuales o técnicos ajenos al propósito académico. Una implementación rigurosa combina rúbricas, portafolios, coevaluación y observación docente para construir interpretaciones más completas, evitando que la autenticidad se reduzca a presentar actividades extensas o productos llamativos.

3.7 Integración de la evaluación formativa en la planificación docente

La integración de la evaluación formativa en la planificación docente exige concebirla como parte del proceso de enseñanza y no como una actividad añadida al final de una unidad. Desde el inicio, deben definirse los aprendizajes esperados, las evidencias que permitirán reconocerlos y los criterios para interpretar su calidad. La evaluación formativa adquiere sentido cuando orienta ajustes durante el desarrollo de las actividades y no únicamente cuando registra resultados alcanzados (Huauya et al., 2023). Esta articulación convierte la planificación en un proceso flexible y sustentado en evidencias.

El punto de partida consiste en establecer objetivos claros y observables que puedan relacionarse con tareas, instrumentos y decisiones pedagógicas. Gu (2021) sostiene que la validez de la evaluación formativa depende de la coherencia entre las evidencias recogidas y las interpretaciones que se realizan. Por ello, la planificación debe anticipar qué desempeños demostrarán comprensión y qué información será insuficiente para sostener una conclusión.

Cuando los objetivos son ambiguos, los instrumentos tienden a recoger datos dispersos y la retroalimentación pierde precisión, porque no existe una referencia clara para valorar el progreso.

La selección de actividades debe responder a la naturaleza del aprendizaje esperado y permitir evidenciar procesos, no solo productos finales. Una planificación equilibrada puede combinar cuestionarios, tareas abiertas, proyectos, portafolios, foros y simulaciones según las competencias implicadas. La evaluación en línea admite múltiples métodos, pero su pertinencia depende de que cada recurso cumpla una función definida dentro de la secuencia didáctica (Heil y Ifenthaler, 2023). Incorporar muchas herramientas sin articulación puede aumentar la carga de trabajo sin mejorar la comprensión del aprendizaje.

Los criterios de evaluación deben comunicarse antes de desarrollar las actividades para que orienten la actuación del estudiante. Las rúbricas digitales favorecen esta transparencia al presentar dimensiones y niveles de desempeño de manera organizada. Su uso dentro de la evaluación formativa permite hacer visibles las expectativas y apoyar procesos de revisión (Romero et al., 2025). No obstante, los criterios deben analizarse con los estudiantes y vincularse con ejemplos. Una rúbrica entregada sin explicación puede convertirse en un documento formal que se consulta únicamente al recibir la calificación.

La planificación también debe incluir momentos específicos para recoger evidencias y analizarlas. El seguimiento continuo requiere decidir qué información será observada, con qué frecuencia y para qué se utilizará. Liao et al. (2022) advierten que la evaluación digital administrada de manera permanente necesita procedimientos para mantener su calidad. Registrar todos los accesos, tiempos o interacciones no equivale a evaluar de forma formativa. El docente debe seleccionar indicadores pertinentes y contrastarlos con producciones concretas para evitar interpretaciones basadas únicamente en la actividad tecnológica.

La retroalimentación debe planificarse como una fase vinculada con nuevas oportunidades de actuación. Vera (2022) destaca que su utilidad depende de que sea clara, oportuna y orientada hacia la mejora. En consecuencia, la secuencia didáctica necesita reservar tiempo para revisar comentarios, corregir errores y presentar nuevas versiones. Si la retroalimentación se entrega después de cerrar la actividad, su función se reduce a explicar una calificación. Integrarla en la planificación implica prever ciclos de producción, orientación y ajuste, no solo fechas de entrega y registro de resultados.

La analítica del aprendizaje puede apoyar la toma de decisiones cuando se integra con propósitos definidos. Banihashem et al. (2022) muestran que los datos generados por plataformas pueden fortalecer la retroalimentación y el seguimiento. Sin embargo, la planificación debe establecer qué patrones justificarán una intervención y qué acciones se adoptarán frente a ellos. Un reporte automático carece de valor si no conduce a cambios en actividades, apoyos o explicaciones. La información digital debe transformarse en decisiones pedagógicas contextualizadas y no permanecer como un registro descriptivo del comportamiento estudiantil.

La diferenciación de apoyos constituye otra dimensión de la planificación formativa. Los resultados obtenidos durante las actividades pueden revelar ritmos, dificultades y necesidades distintas. Los sistemas adaptativos permiten ajustar contenidos y niveles de complejidad según el desempeño registrado (Aparicio y Aparicio, 2024). No obstante, las rutas personalizadas deben conservar flexibilidad y supervisión docente. La planificación no puede fijar trayectorias rígidas a partir de un resultado inicial, porque el diagnóstico debe orientar apoyos y ampliar oportunidades, no limitar expectativas ni anticipar conclusiones definitivas sobre la capacidad del estudiante.

La participación del estudiante debe integrarse mediante autoevaluación, coevaluación y análisis de criterios. Estas prácticas fortalecen la autorregulación cuando permiten interpretar evidencias y tomar decisiones sobre el propio trabajo. La evaluación entre pares puede contribuir al desarrollo de competencias transversales en proyectos colaborativos (Santana et al., 2025). Para que esta participación sea formativa, la planificación debe incluir preparación previa, instrumentos claros y espacios para utilizar los comentarios recibidos. De lo contrario, puede convertirse en una actividad aislada sin influencia real sobre el aprendizaje.

La accesibilidad y la equidad también deben preverse desde el diseño de la planificación. Las actividades digitales pueden generar barreras cuando dependen de un único formato, tiempos rígidos o recursos tecnológicos no disponibles para todos. Las directrices del CAST (2024) promueven múltiples formas de participación, representación y expresión, principios aplicables a la evaluación. La diversificación de evidencias no significa reducir criterios, sino ofrecer vías razonables para demostrarlos. Una planificación inclusiva anticipa dificultades y evita que las adaptaciones se realicen únicamente después de que aparezca la exclusión.

La integración de la evaluación formativa en la planificación docente requiere alinear objetivos, actividades, criterios, evidencias, retroalimentación y decisiones de mejora. Su eficacia no

depende de aumentar el número de instrumentos, sino de definir la función de cada uno dentro de una secuencia coherente. Las tecnologías digitales facilitan seguimiento, personalización y comunicación, pero necesitan supervisión pedagógica constante. Cuando la planificación incorpora ciclos de observación, interpretación y ajuste, la evaluación deja de operar como cierre administrativo y se convierte en un mecanismo permanente para mejorar la enseñanza y el aprendizaje.

Capítulo

04

Inteligencia artificial, analítica de datos y automatización evaluativa



4.1 Fundamentos de la inteligencia artificial aplicada a la evaluación

La inteligencia artificial aplicada a la evaluación educativa comprende sistemas capaces de procesar datos, reconocer patrones, generar predicciones y apoyar decisiones relacionadas con el aprendizaje. Su utilización incluye puntuación automatizada, retroalimentación, clasificación de respuestas y adaptación de actividades. Wang et al. (2024) identifican la evaluación como uno de los campos de mayor expansión dentro de la inteligencia artificial educativa. Sin embargo, estas tecnologías no evalúan de manera autónoma en sentido pedagógico, pues operan a partir de criterios, datos y modelos previamente definidos por personas e instituciones.

El aprendizaje automático constituye uno de los fundamentos técnicos más relevantes de estos sistemas. Mediante algoritmos entrenados con ejemplos, es posible identificar regularidades y asignar puntuaciones a nuevas respuestas. Esta capacidad se ha utilizado para valorar producciones multidimensionales en ciencias y argumentaciones complejas (Maestrales et al., 2021). No obstante, el rendimiento del modelo depende de la calidad y representatividad de los datos de entrenamiento. Cuando los ejemplos son limitados, inconsistentes o poco diversos, las predicciones pueden resultar técnicamente estables, pero pedagógicamente inadecuadas.

La automatización evaluativa requiere definir con precisión el constructo que se desea valorar. Haudek y Zhai (2024) muestran que las características del constructo influyen directamente en el desempeño de los modelos utilizados para puntuar argumentaciones científicas. Esto significa que la inteligencia artificial no interpreta cualquier respuesta con la misma profundidad ni precisión. Los resultados dependen de cómo se operacionalizan habilidades como razonamiento, coherencia o aplicación conceptual. Una definición ambigua del aprendizaje puede producir puntuaciones exactas en apariencia, aunque poco representativas de la competencia evaluada.

La validez constituye un principio indispensable para interpretar resultados generados por inteligencia artificial. Las puntuaciones automatizadas deben conectarse con inferencias justificadas sobre el desempeño y no aceptarse únicamente por su nivel de precisión estadística. Zhai et al. (2021) proponen una red inferencial para examinar la validez de evaluaciones científicas basadas en aprendizaje automático. Esta perspectiva obliga a revisar la relación entre respuestas, modelos, puntuaciones e interpretaciones. La tecnología aporta capacidad de procesamiento, pero la legitimidad educativa depende de argumentos y evidencias pedagógicas consistentes.

La retroalimentación automatizada representa otra aplicación central de la inteligencia artificial. Estos sistemas pueden detectar errores, reconocer patrones y generar orientaciones de manera inmediata. Lee et al. (2021) muestran que el aprendizaje automático puede apoyar la revisión de argumentos científicos mediante comentarios vinculados con el desempeño. También se han desarrollado sistemas capaces de transformar la evaluación automatizada de ensayos en retroalimentación más informativa (Gombert et al., 2024). Su utilidad depende de que los mensajes sean comprensibles, específicos y relacionados con oportunidades reales de mejora.

La personalización se fundamenta en la capacidad de adaptar contenidos, actividades y apoyos según los datos producidos por cada estudiante. Los sistemas de aprendizaje adaptativo pueden modificar niveles de dificultad, secuencias y recursos en función del desempeño registrado (Aparicio y Aparicio, 2024). Esta posibilidad favorece intervenciones diferenciadas, pero también puede generar trayectorias rígidas cuando las recomendaciones se aceptan sin revisión. La personalización debe ampliar oportunidades y no restringirlas mediante predicciones que anticipen de manera permanente lo que el estudiante puede o no realizar.

La explicabilidad resulta esencial cuando la inteligencia artificial participa en decisiones académicas. Khosravi et al. (2022) sostienen que los sistemas educativos explicables deben ofrecer información comprensible sobre los factores que influyen en sus resultados. Sin esta transparencia, docentes y estudiantes pueden quedar subordinados a puntuaciones difíciles de revisar. Explicar no significa revelar cada detalle técnico del algoritmo, sino comunicar los criterios utilizados, las limitaciones existentes y las razones relevantes que sustentan una recomendación. La revisión humana debe mantenerse disponible frente a resultados controvertidos.

El sesgo algorítmico constituye uno de los principales riesgos de la evaluación automatizada. Los modelos pueden reproducir desigualdades presentes en los datos, favorecer determinados estilos de respuesta o generar diferencias injustificadas entre grupos. Baker y Hawn (2022) muestran que el sesgo en educación puede originarse en distintas fases, desde la selección de datos hasta la implementación del sistema. Por ello, no basta con evaluar la precisión general. Es necesario analizar resultados diferenciados, revisar errores recurrentes y comprobar si ciertos estudiantes reciben sistemáticamente puntuaciones o recomendaciones desfavorables.

La dimensión ética abarca privacidad, consentimiento, transparencia y responsabilidad institucional. Los sistemas evaluativos recopilan grandes volúmenes de información sobre

respuestas, interacciones y trayectorias académicas. Las preocupaciones estudiantiles aumentan cuando no se conoce cómo se utilizan o comparten estos datos (Korir et al., 2023). La UNESCO (2021) establece que la inteligencia artificial debe desarrollarse con respeto a los derechos humanos, la equidad y la supervisión humana. En consecuencia, la eficiencia tecnológica no puede justificar prácticas opacas ni una recopilación desproporcionada de información educativa.

La inteligencia artificial generativa amplía este campo al producir comentarios, preguntas, ejemplos y propuestas de corrección. Su uso puede apoyar la retroalimentación de escritura académica y ofrecer respuestas rápidas durante el proceso formativo (Ossa y Willatt, 2023). Sin embargo, también introduce problemas de autoría, precisión y dependencia tecnológica. La UNESCO (2023) recomienda una utilización educativa guiada por criterios humanos y reconocimiento de sus limitaciones. Estos sistemas pueden generar información plausible pero incorrecta, por lo que sus aportes deben revisarse antes de integrarse en decisiones evaluativas.

Los fundamentos de la inteligencia artificial aplicada a la evaluación combinan procesamiento automatizado, validez, explicabilidad, equidad y supervisión humana. Su potencial reside en analizar evidencias, personalizar apoyos y ofrecer retroalimentación con mayor rapidez, pero ninguna de estas ventajas elimina la responsabilidad docente. La calidad depende de la definición del constructo, los datos empleados y los mecanismos de revisión. La inteligencia artificial debe funcionar como apoyo para interpretar el aprendizaje, no como autoridad independiente capaz de sustituir criterios pedagógicos, éticos e institucionales.

4.2 Analítica del aprendizaje y seguimiento de trayectorias estudiantiles

La analítica del aprendizaje comprende la recopilación, organización e interpretación de datos generados durante la actividad educativa con el propósito de comprender y mejorar los procesos formativos. En evaluación, su principal aporte consiste en ampliar la observación más allá de resultados aislados mediante registros sobre participación, entregas, intentos, tiempos y progresión. Muñoz et al. (2021) señalan que estas tecnologías permiten examinar información procedente de entornos digitales. Sin embargo, su valor depende de convertir los datos en evidencias pedagógicas y no en simples descripciones de actividad.

El seguimiento de trayectorias estudiantiles exige analizar cambios producidos a lo largo del tiempo y no únicamente comparar calificaciones finales. Las plataformas pueden integrar resultados sucesivos, revisiones, retroalimentaciones y patrones de interacción para identificar avances o interrupciones. La adopción de la analítica depende de prioridades institucionales,

experiencias acumuladas y capacidades para interpretar información educativa (Tsai et al., 2021). Por ello, una trayectoria no debe reducirse a una secuencia automática de indicadores, sino comprenderse dentro de las condiciones académicas, tecnológicas y personales que influyen en el aprendizaje.

Las fuentes de datos utilizadas por la analítica son diversas e incluyen cuestionarios, tareas, foros, portafolios, simulaciones y registros de acceso. González y Lugo (2021) muestran que esta información puede fortalecer la práctica docente cuando ayuda a reconocer patrones de participación y desempeño. No obstante, cada fuente representa dimensiones distintas del aprendizaje. Un tiempo prolongado de conexión puede responder a dificultad técnica, mientras una baja frecuencia de acceso no implica necesariamente escaso compromiso. La interpretación requiere contrastar registros digitales con producciones, criterios y observaciones cualitativas.

La identificación temprana de dificultades constituye una de las aplicaciones más relevantes de la analítica. Los sistemas pueden reconocer disminuciones en la participación, acumulación de tareas pendientes o errores recurrentes y generar alertas para orientar intervenciones. Estas señales deben entenderse como indicios y no como diagnósticos definitivos, porque los modelos pueden omitir factores contextuales. Los enfoques de analítica centrados en las personas procuran equilibrar el procesamiento automatizado con necesidades, experiencias y decisiones humanas (Alfredo et al., 2024), evitando que una predicción determine por sí sola la trayectoria del estudiante.

La personalización de apoyos se relaciona con la capacidad de utilizar datos para recomendar actividades, recursos o niveles de dificultad. Los sistemas adaptativos pueden responder a patrones de desempeño y ofrecer recorridos diferenciados (Aparicio y Aparicio, 2024). Esta función puede fortalecer el aprendizaje cuando amplía oportunidades y atiende necesidades específicas. Sin embargo, una recomendación automatizada puede consolidar expectativas bajas o encerrar al estudiante en tareas simplificadas. La trayectoria debe conservar posibilidades de cambio, revisión docente y acceso a experiencias de mayor complejidad conforme aparezcan nuevas evidencias.

Tabla 9

Dimensiones de la analítica del aprendizaje para el seguimiento de trayectorias estudiantiles

Dimensión analítica	Datos considerados	Decisión educativa posible	Principal aporte	Riesgo de interpretación
Participación	Accesos, intervenciones y entregas	Contacto y acompañamiento temprano	Detecta disminución de actividad	Confundir conexión con aprendizaje
Desempeño	Puntuaciones, errores e intentos	Refuerzo de contenidos y ajuste de tareas	Identifica dificultades recurrentes	Reducir la trayectoria a calificaciones
Progresión	Comparación entre evidencias sucesivas	Modificación de apoyos y metas	Visibiliza cambios a lo largo del tiempo	Ignorar variaciones del contexto
Retroalimentación	Comentarios recibidos y revisiones	Planificación de nuevas orientaciones	Muestra respuesta ante sugerencias	Asumir que recibir comentarios implica utilizarlos
Colaboración	Interacciones y aportes grupales	Reorganización de roles o actividades	Permite observar dinámicas colectivas	Valorar cantidad sobre calidad
Predicción	Combinación de patrones históricos	Activación de alertas preventivas	Facilita intervención anticipada	Etiquetar o limitar expectativas
Personalización	Resultados y necesidades detectadas	Recomendación de recursos y rutas	Ajusta apoyos al desempeño	Crear trayectorias rígidas

Nota. La tabla relaciona fuentes de datos, decisiones educativas y riesgos que deben considerarse al interpretar trayectorias mediante analítica del aprendizaje.

La tabla evidencia que los datos solo adquieren valor cuando se relacionan con decisiones pedagógicas concretas y revisables. Banihashem et al. (2022) muestran que la analítica puede fortalecer la retroalimentación al ofrecer información sobre progreso y dificultades. Sin embargo, una alerta o recomendación carece de sentido si no conduce a orientaciones, apoyos o cambios en la enseñanza. El seguimiento debe producir intervenciones proporcionales y evitar que los indicadores se conviertan en mecanismos permanentes de clasificación, vigilancia o comparación entre estudiantes.

La calidad del seguimiento depende también de la validez y estabilidad de los datos. Liao et al. (2022) señalan que las evaluaciones digitales administradas continuamente requieren supervisión para mantener la consistencia de sus resultados. Cambios en plataformas, actividades o criterios pueden alterar los registros y dificultar comparaciones longitudinales. Antes de interpretar una trayectoria, es necesario comprobar que las evidencias representan

aprendizajes equivalentes y que las variaciones no proceden exclusivamente de modificaciones técnicas. La continuidad analítica exige revisión periódica de instrumentos, indicadores y supuestos.

La privacidad constituye una condición indispensable, porque el seguimiento longitudinal implica recopilar y conservar información sobre comportamiento y rendimiento. Las preocupaciones estudiantiles abarcan la recopilación, el uso y la compartición de datos empleados por sistemas de analítica (Korir et al., 2023). Además, estas percepciones varían entre contextos culturales, lo que influye en la confianza y aceptación de las prácticas institucionales (Viberg et al., 2024). La gestión responsable exige limitar los datos, comunicar su finalidad, controlar los accesos y establecer periodos claros de conservación.

La analítica del aprendizaje puede enriquecer el seguimiento de trayectorias cuando combina datos continuos, interpretación contextual y supervisión docente. Su potencial reside en reconocer progresos, anticipar dificultades y ajustar apoyos, pero no en reemplazar el juicio pedagógico mediante predicciones automáticas. Los indicadores deben contrastarse con evidencias diversas y revisarse antes de producir consecuencias académicas. Una aplicación responsable mantiene abiertas las oportunidades de aprendizaje, protege la privacidad y utiliza la información para acompañar al estudiante, evitando convertir su trayectoria en una clasificación fija determinada por algoritmos.

4.3 Sistemas automatizados de calificación y corrección

Los sistemas automatizados de calificación y corrección utilizan algoritmos para analizar respuestas, asignar puntuaciones y generar resultados con menor intervención manual. Su aplicación se ha extendido desde cuestionarios cerrados hasta tareas de escritura, argumentación científica y evaluaciones multidimensionales. Maestrales et al. (2021) muestran que el aprendizaje automático puede emplearse para puntuar respuestas complejas en ciencias. Aun así, automatizar la corrección no equivale a garantizar una evaluación válida, pues la calidad depende de los criterios, datos y modelos utilizados.

En las preguntas de respuesta cerrada, la automatización suele operar mediante reglas previamente definidas que comparan la respuesta del estudiante con una opción correcta. Este procedimiento ofrece rapidez, consistencia operativa y capacidad para procesar grandes grupos. Sin embargo, su alcance es limitado cuando el aprendizaje exige justificar, argumentar o transferir conocimientos. La eficiencia técnica puede favorecer el uso excesivo de reactivos

fáciles de corregir, aunque poco adecuados para representar competencias complejas (Heil y Ifenthaler, 2023). El instrumento debe diseñarse desde el propósito pedagógico.

La corrección de respuestas abiertas requiere procedimientos más complejos, porque el sistema debe reconocer variaciones lingüísticas, relaciones conceptuales y niveles de calidad. Haudek y Zhai (2024) demuestran que las características del constructo evaluado influyen directamente en el desempeño de los modelos de puntuación. Esto implica que no todas las argumentaciones pueden analizarse con la misma precisión. Cuando el constructo es amplio, ambiguo o depende del contexto, la automatización puede simplificar indebidamente la respuesta y asignar puntuaciones que no representan su verdadero valor académico.

Los modelos de aprendizaje automático se entrenan con conjuntos de respuestas previamente calificadas por evaluadores humanos. A partir de esos ejemplos, identifican patrones asociados con determinados niveles de desempeño. La calidad del modelo depende de la consistencia de las calificaciones originales y de la diversidad de las respuestas utilizadas. Si los datos de entrenamiento contienen errores, sesgos o criterios inestables, el sistema puede reproducirlos a gran escala. El sesgo algorítmico educativo puede originarse tanto en los datos como en el diseño y la implementación (Baker y Hawn, 2022).

La validez de las puntuaciones automatizadas debe examinarse mediante una relación justificable entre respuesta, resultado e interpretación. Zhai et al. (2021) proponen analizar estas evaluaciones mediante redes de inferencias que permitan verificar si las puntuaciones sostienen las conclusiones académicas esperadas. Una alta coincidencia entre el algoritmo y los evaluadores humanos no garantiza por sí sola validez. También debe comprobarse que ambos valoren realmente el constructo previsto y no características superficiales, como extensión, vocabulario frecuente o estructura repetitiva de la respuesta.

La consistencia constituye una ventaja relevante de estos sistemas, porque aplican los mismos criterios a grandes volúmenes de respuestas. A diferencia de la corrección humana, no experimentan fatiga ni cambios de atención durante la jornada. No obstante, esta uniformidad puede convertirse en una limitación cuando el modelo interpreta de forma incorrecta una respuesta no prevista. La calidad de una evaluación digital administrada continuamente requiere supervisión, revisión de ítems y análisis de patrones anómalos (Liao et al., 2022). La estabilidad técnica necesita controles pedagógicos permanentes.

La corrección automatizada también puede vincularse con retroalimentación inmediata. Lee et al. (2021) muestran que los sistemas basados en aprendizaje automático pueden apoyar la

revisión de argumentaciones científicas mediante comentarios asociados con los datos producidos por el estudiante. De manera similar, Wilson et al. (2021) analizaron sistemas de puntuación y retroalimentación automatizadas en escritura escolar. Estas aplicaciones permiten nuevos intentos y revisión progresiva, pero su utilidad disminuye cuando los mensajes son genéricos, repetitivos o incapaces de explicar por qué una respuesta necesita mejorar.

En la evaluación de ensayos, los sistemas han evolucionado desde la asignación de puntuaciones globales hacia comentarios más específicos sobre contenido y estructura. Gombert et al. (2024) muestran que la evaluación automatizada puede transformarse en retroalimentación informativa cuando el sistema identifica dimensiones relevantes del texto. Aun así, la escritura académica incluye matices, intenciones y decisiones discursivas difíciles de reducir a patrones. La corrección automática debe funcionar como apoyo para la revisión, no como autoridad exclusiva sobre la calidad de una producción compleja.

La explicabilidad resulta indispensable cuando una puntuación automatizada produce consecuencias académicas. Khosravi et al. (2022) señalan que los sistemas educativos explicables deben ofrecer información comprensible sobre los factores que influyen en sus decisiones. El estudiante necesita saber qué criterios fueron considerados, por qué obtuvo determinado resultado y cómo puede solicitar revisión. Sin mecanismos de explicación, la automatización puede presentar sus resultados como objetivos e incuestionables, aunque dependan de supuestos técnicos, datos históricos y decisiones institucionales que requieren examen crítico.

La supervisión humana debe mantenerse especialmente en evaluaciones de alta consecuencia, como acreditación, promoción o certificación. Los resultados automatizados pueden apoyar el trabajo docente, identificar patrones y reducir tareas repetitivas, pero no deberían determinar por sí solos decisiones irreversibles. La UNESCO (2021) establece la importancia de preservar responsabilidad, equidad y control humano en el uso de inteligencia artificial. Esto implica definir cuándo la revisión docente es obligatoria, cómo se resuelven discrepancias y qué procedimientos permiten corregir errores detectados después de la calificación.

Figura 6

¿Deberíamos usar sistemas automatizados de calificación y corrección?



Nota. La figura presenta criterios para utilizar sistemas automatizados de calificación según el tipo de respuesta y el nivel de supervisión requerido.

La privacidad también forma parte de la implementación, porque estos sistemas procesan respuestas, patrones lingüísticos y datos de desempeño. Las preocupaciones estudiantiles aumentan cuando no existe claridad sobre la recopilación, almacenamiento y compartición de la información (Korir et al., 2023). En Ecuador, su tratamiento debe ajustarse a la Ley Orgánica de Protección de Datos Personales (2021) y a su Reglamento General (2023). La automatización no justifica recopilar datos innecesarios ni conservarlos durante periodos indefinidos sin una finalidad educativa claramente establecida.

Los sistemas automatizados de calificación y corrección pueden aumentar rapidez, consistencia y capacidad de respuesta, especialmente en contextos con grandes volúmenes de evidencias. Su uso responsable exige precisión del constructo, datos representativos, validación, explicabilidad, revisión humana y protección de la información. La automatización ofrece mayor valor cuando complementa el juicio docente y libera tiempo para tareas interpretativas y formativas. Cuando sustituye la deliberación pedagógica o se presenta como neutral, puede consolidar errores, sesgos y decisiones difíciles de cuestionar.

4.4 Retroalimentación adaptativa mediante sistemas inteligentes

La retroalimentación adaptativa mediante sistemas inteligentes utiliza datos del desempeño para ajustar mensajes, apoyos y actividades según las necesidades detectadas. A diferencia de una respuesta automática uniforme, busca reconocer patrones de error, nivel de dominio y progresión del estudiante. Aparicio y Aparicio (2024) relacionan los sistemas adaptativos con experiencias educativas más personalizadas. Sin embargo, la adaptación solo resulta

pedagógicamente válida cuando se vincula con criterios claros y evita convertir una predicción algorítmica en una clasificación permanente de las capacidades individuales.

Estos sistemas pueden analizar respuestas, intentos, tiempos y secuencias de resolución para determinar qué orientación ofrecer. La retroalimentación puede consistir en pistas, ejemplos, preguntas guiadas, recomendaciones de recursos o cambios en la dificultad. Esta variedad permite responder de manera diferenciada a errores que, aunque produzcan el mismo resultado, tienen causas distintas. La analítica del aprendizaje fortalece la retroalimentación cuando transforma datos en información útil para orientar mejoras (Banihashem et al., 2022). Sin interpretación pedagógica, los registros solo describen actividad.

La oportunidad representa una de las principales fortalezas de la retroalimentación adaptativa. El estudiante recibe orientación durante la tarea, cuando todavía puede revisar su estrategia y corregir el procedimiento. De Schipper et al. (2021) muestran que los sistemas de recomendación pueden generar retroalimentación personalizada incluso en contextos de evaluación sumativa. Esta posibilidad reduce la distancia entre error y corrección, aunque debe evitarse que la ayuda revele directamente la respuesta. Una intervención excesiva puede facilitar el acierto inmediato sin producir comprensión ni transferencia.

La calidad de los mensajes depende de la precisión con que el sistema identifica la dificultad. Lee et al. (2021) evidencian que el aprendizaje automático puede apoyar la revisión de argumentaciones científicas mediante comentarios asociados con el desempeño observado. Cuando el diagnóstico es correcto, la retroalimentación puede orientar cambios específicos y promover nuevos intentos. Si el modelo interpreta de forma equivocada la respuesta, puede recomendar acciones innecesarias o confusas. Por ello, resulta indispensable revisar la correspondencia entre patrones detectados, criterios y orientación proporcionada.

Los sistemas inteligentes también pueden adaptar la complejidad de las actividades. Ante respuestas correctas y consistentes, el sistema puede proponer desafíos mayores; frente a dificultades reiteradas, puede ofrecer apoyos progresivos. Esta lógica favorece trayectorias diferenciadas y evita mantener a todos los estudiantes en una misma secuencia (Aparicio y Aparicio, 2024). No obstante, la adaptación debe conservar oportunidades de avance. Reducir permanentemente la dificultad a partir de errores iniciales puede limitar el aprendizaje y consolidar expectativas bajas mediante decisiones automatizadas.

La retroalimentación sobre producciones escritas constituye uno de los campos con mayor desarrollo. Gombert et al. (2024) muestran que la evaluación automatizada de ensayos puede

evolucionar hacia comentarios más informativos sobre dimensiones específicas del texto. La inteligencia artificial generativa también puede sugerir mejoras en estructura, claridad y argumentación (Ossa y Willatt, 2023). Estas recomendaciones deben utilizarse críticamente, porque un sistema puede proponer cambios lingüísticamente plausibles que alteren el sentido, simplifiquen ideas o no respondan a los criterios académicos establecidos.

La explicabilidad resulta esencial para que el estudiante comprenda por qué recibe determinada orientación. Khosravi et al. (2022) sostienen que los sistemas educativos explicables deben comunicar los factores relevantes que intervienen en sus resultados. Una recomendación adaptativa pierde legitimidad cuando aparece como una instrucción sin fundamento visible. El usuario necesita conocer qué evidencia activó el mensaje, qué criterio se relaciona con la dificultad y qué alternativas existen. Esta transparencia favorece confianza, autonomía y posibilidad de cuestionar sugerencias inadecuadas.

La intervención docente continúa siendo necesaria para interpretar casos complejos y corregir limitaciones del sistema. La retroalimentación automatizada puede atender errores frecuentes y liberar tiempo para orientaciones más profundas, pero no comprende siempre el contexto, la intención o las circunstancias del estudiante. Alfredo et al. (2024) destacan la relevancia de enfoques centrados en las personas dentro de la analítica y la inteligencia artificial educativa. Esta orientación exige que el sistema complemente la relación pedagógica y no reemplace la conversación, el acompañamiento ni el juicio profesional.

El sesgo algorítmico puede afectar la personalización cuando el modelo funciona mejor con determinados estilos lingüísticos, formas de respuesta o grupos de estudiantes. Baker y Hawn (2022) advierten que los sesgos educativos pueden originarse en los datos, el diseño y la implementación. Una retroalimentación aparentemente individualizada puede reproducir recomendaciones desiguales si el sistema interpreta con menor precisión ciertas respuestas. La evaluación debe incluir pruebas diferenciadas, revisión de errores y mecanismos de corrección para evitar que la adaptación amplíe desigualdades existentes.

La privacidad también condiciona la legitimidad de estos sistemas, porque la personalización requiere recopilar y analizar información detallada sobre desempeño y comportamiento. Las preocupaciones estudiantiles se relacionan con el uso, almacenamiento y compartición de datos generados mediante analítica (Korir et al., 2023). En Ecuador, estas prácticas deben ajustarse a la Ley Orgánica de Protección de Datos Personales (2021) y a su Reglamento General (2023).

La adaptación no justifica recopilar información innecesaria ni mantener perfiles indefinidos sin una finalidad pedagógica explícita.

La retroalimentación adaptativa puede fortalecer el aprendizaje cuando ofrece orientación oportuna, específica y vinculada con nuevas oportunidades de actuación. Su potencial reside en diferenciar apoyos, ajustar dificultades y responder a patrones que serían difíciles de identificar manualmente. No obstante, su calidad depende de la validez del diagnóstico, la explicabilidad, la supervisión docente y la protección de datos. Los sistemas inteligentes deben ampliar la capacidad de intervención pedagógica sin transformar recomendaciones probabilísticas en decisiones rígidas ni sustituir la relación educativa.

4.5 Evaluación personalizada y aprendizaje adaptativo

La evaluación personalizada busca ajustar tareas, criterios de apoyo y formas de retroalimentación a las características del estudiante, mientras el aprendizaje adaptativo modifica secuencias y niveles de dificultad según su desempeño. Aparicio y Aparicio (2024) vinculan estos sistemas con experiencias educativas más flexibles y diferenciadas. Su propósito no consiste en aplicar exigencias arbitrariamente distintas, sino en ofrecer caminos diversos para alcanzar objetivos comunes. La personalización adquiere valor cuando responde a evidencias válidas y no cuando se sustenta en perfiles rígidos o predicciones difíciles de revisar.

Los sistemas adaptativos funcionan mediante la recopilación de datos sobre respuestas, errores, intentos, tiempos y progresión. A partir de estos registros, pueden recomendar actividades, modificar la complejidad o proporcionar apoyos específicos. La analítica del aprendizaje permite reconocer patrones y orientar decisiones sobre el proceso formativo (Muñoz et al., 2021). Sin embargo, los datos deben interpretarse con cautela, porque una baja frecuencia de interacción o una respuesta incorrecta pueden estar relacionadas con factores técnicos, contextuales o metodológicos y no únicamente con falta de comprensión.

La personalización evaluativa exige identificar con claridad qué dimensión del aprendizaje necesita ajustarse. No toda diferencia entre estudiantes justifica modificar la tarea, ni toda adaptación mejora la validez. Gu (2021) sostiene que las inferencias evaluativas deben vincular evidencias, interpretaciones y decisiones mediante argumentos consistentes. Por tanto, el ajuste puede aplicarse al tipo de apoyo, al tiempo, al formato o a la secuencia, pero debe conservar los criterios esenciales del aprendizaje. Adaptar no significa reducir expectativas ni sustituir objetivos por actividades más simples.

La retroalimentación constituye uno de los mecanismos principales de la evaluación personalizada. Los sistemas pueden ofrecer comentarios diferenciados según el tipo de error, el nivel de dominio o la estrategia empleada. Banihashem et al. (2022) muestran que la analítica puede fortalecer la retroalimentación cuando produce información pertinente y oportuna. Esta posibilidad resulta valiosa si el mensaje explica la dificultad y propone una acción concreta. Cuando la personalización se limita a variar frases automáticas, su efecto pedagógico es reducido y puede generar una apariencia de atención individual sin orientación real.

El aprendizaje adaptativo también permite modificar la dificultad de las actividades conforme aparecen nuevas evidencias. Una secuencia puede avanzar hacia tareas más complejas cuando el desempeño es consistente u ofrecer apoyos adicionales frente a errores recurrentes. Estas rutas diferenciadas favorecen una mayor correspondencia entre desafío y nivel de preparación (Aparicio y Aparicio, 2024). No obstante, una adaptación excesivamente conservadora puede mantener al estudiante en actividades simples y limitar su exposición a problemas complejos. Toda ruta debe conservar posibilidades de avance y revisión docente.

La inteligencia artificial amplía la capacidad de personalización al procesar grandes volúmenes de información y reconocer patrones difíciles de observar manualmente. Wang et al. (2024) identifican la adaptación y la evaluación como áreas relevantes dentro de la inteligencia artificial educativa. Estos sistemas pueden combinar resultados históricos, respuestas actuales y registros de interacción para recomendar acciones. Sin embargo, la precisión algorítmica no garantiza pertinencia pedagógica. Las decisiones deben analizarse dentro del contexto, porque un modelo puede identificar correlaciones útiles sin comprender las causas que explican el desempeño.

Tabla 10
Dimensiones de la evaluación personalizada y el aprendizaje adaptativo

Dimensión	Evidencia utilizada	Adaptación posible	Aporte pedagógico	Riesgo principal
Nivel de dominio	Respuestas, errores e intentos	Ajuste de dificultad	Mantiene un desafío progresivo	Simplificar permanentemente las tareas
Ritmo de aprendizaje	Tiempo y secuencia de avance	Flexibilización de plazos y recorridos	Respeta diferencias de progresión	Confundir lentitud con bajo rendimiento
Tipo de dificultad	Patrones de error	Pistas, ejemplos o recursos específicos	Ofrece apoyo focalizado	Diagnosticar incorrectamente la causa

Dimensión	Evidencia utilizada	Adaptación posible	Aporte pedagógico	Riesgo principal
Forma de participación	Preferencias y condiciones de acceso	Diversificación de formatos	Amplía oportunidades de expresión	Personalizar con base en datos insuficientes
Retroalimentación	Desempeño frente a criterios	Comentarios diferenciados	Orienta acciones concretas de mejora	Generar mensajes automáticos superficiales
Trayectoria	Evidencias acumuladas	Recomendación de nuevas actividades	Facilita seguimiento longitudinal	Convertir predicciones en etiquetas
Intervención docente	Alertas y reportes analíticos	Acompañamiento dirigido	Prioriza apoyos oportunos	Delegar decisiones al sistema

Nota. La tabla organiza las principales dimensiones de personalización y muestra cómo una misma adaptación puede generar beneficios o riesgos según la calidad de las evidencias utilizadas.

La tabla evidencia que la personalización no debe reducirse al ajuste automático de dificultad. También puede involucrar tiempos, formatos, recursos, retroalimentación y formas de participación. Alfredo et al. (2024) destacan la importancia de enfoques centrados en las personas dentro de la analítica y la inteligencia artificial educativa. Esta perspectiva exige considerar necesidades, experiencias y contextos, no solo indicadores cuantitativos. La adaptación resulta más sólida cuando combina datos digitales, observación docente y participación del estudiante en la toma de decisiones sobre su trayectoria.

La explicabilidad constituye una condición para que las decisiones adaptativas sean comprensibles y cuestionables. Khosravi et al. (2022) sostienen que los sistemas educativos deben ofrecer información sobre los factores que influyen en sus recomendaciones. El estudiante necesita conocer por qué recibe una actividad distinta, qué evidencia motivó el ajuste y cómo puede avanzar hacia otros niveles. Sin esta transparencia, la personalización puede percibirse como una clasificación arbitraria. La revisión humana debe estar disponible cuando una recomendación no represente adecuadamente el aprendizaje o limite oportunidades.

Los sesgos algorítmicos pueden afectar la personalización cuando los modelos funcionan mejor con determinados grupos, estilos de respuesta o trayectorias. Baker y Hawn (2022) advierten que el sesgo puede originarse en los datos, el diseño o la implementación de los sistemas educativos. Una adaptación aparentemente individual puede reproducir desigualdades si ciertos estudiantes reciben de forma reiterada tareas menos exigentes o alertas negativas. La

supervisión debe analizar resultados diferenciados, revisar errores y evitar que las predicciones se conviertan en expectativas permanentes sobre la capacidad académica.

La evaluación personalizada y el aprendizaje adaptativo pueden fortalecer la equidad cuando ofrecen apoyos oportunos y diferentes vías para alcanzar objetivos compartidos. Su aplicación exige evidencias válidas, transparencia, revisión docente y posibilidades de modificar las rutas sugeridas. La tecnología debe ampliar las oportunidades de aprendizaje, no consolidar perfiles ni restringir desafíos mediante decisiones automatizadas. Una personalización responsable combina flexibilidad con criterios comunes, reconoce el carácter cambiante del desempeño y mantiene al estudiante como participante activo en la interpretación y orientación de su propia trayectoria.

4.6 Predicción del desempeño y detección temprana de dificultades

La predicción del desempeño educativo utiliza datos históricos y actuales para estimar posibles resultados futuros, identificar patrones de riesgo y orientar intervenciones antes de que las dificultades se consoliden. En los entornos digitales, estas estimaciones pueden construirse a partir de calificaciones, intentos, tiempos, entregas, participación y trayectorias previas. La analítica del aprendizaje permite organizar estos registros y convertirlos en información potencialmente útil para el seguimiento (Muñoz et al., 2021). Sin embargo, una predicción no equivale a un diagnóstico definitivo ni debe interpretarse como una certeza sobre el estudiante.

Los modelos predictivos buscan reconocer combinaciones de variables asociadas con bajo rendimiento, abandono de actividades o retrasos sostenidos. Tsai et al. (2021) muestran que la adopción de analítica depende de factores institucionales, experiencias previas y prioridades educativas. Por ello, la utilidad de una predicción no reside únicamente en su precisión técnica, sino en la capacidad de la institución para interpretarla y actuar. Un sistema que genera alertas sin protocolos de intervención puede acumular información sin producir mejoras reales en la enseñanza o en el acompañamiento.

La detección temprana de dificultades permite intervenir cuando todavía existen oportunidades de recuperación. Entre las señales más frecuentes se encuentran la disminución de entregas, los errores reiterados, la baja participación y la ausencia de progresión. No obstante, estos indicadores deben analizarse con cautela porque pueden estar influidos por conectividad, organización del curso o condiciones personales. La preparación digital desigual de estudiantes e instituciones condiciona las oportunidades de participación y puede distorsionar los registros disponibles (Van de Werfhorst et al., 2022).

La selección de variables constituye una decisión crítica. Un modelo puede utilizar datos fáciles de recopilar, como tiempo de conexión o número de accesos, aunque estos no representen comprensión ni compromiso académico. González y Lugo (2021) señalan que la analítica puede fortalecer la práctica docente cuando los datos se interpretan dentro del contexto educativo. La calidad predictiva disminuye cuando se confunde actividad con aprendizaje. Por esta razón, los indicadores técnicos deben contrastarse con producciones, respuestas, retroalimentación recibida y observaciones cualitativas del docente.

La predicción del desempeño también depende de la calidad y estabilidad de los datos utilizados. Liao et al. (2022) advierten que las evaluaciones digitales administradas continuamente requieren supervisión para mantener su consistencia. Si cambian las plataformas, los criterios o las actividades, las comparaciones longitudinales pueden perder validez. Un modelo entrenado con información procedente de un curso o población específica puede no funcionar de la misma manera en otro contexto. La actualización y revisión periódica son necesarias para evitar decisiones sustentadas en patrones desfasados.

La inteligencia artificial amplía la capacidad de reconocer relaciones complejas entre variables y producir estimaciones a gran escala. Wang et al. (2024) identifican la predicción y la personalización entre las aplicaciones relevantes de la inteligencia artificial en educación. Estos modelos pueden combinar múltiples fuentes de información y detectar señales difíciles de observar manualmente. Aun así, su capacidad estadística no implica comprensión causal. Una correlación entre determinadas conductas y bajo rendimiento no explica por sí sola qué origina la dificultad ni qué intervención será más adecuada.

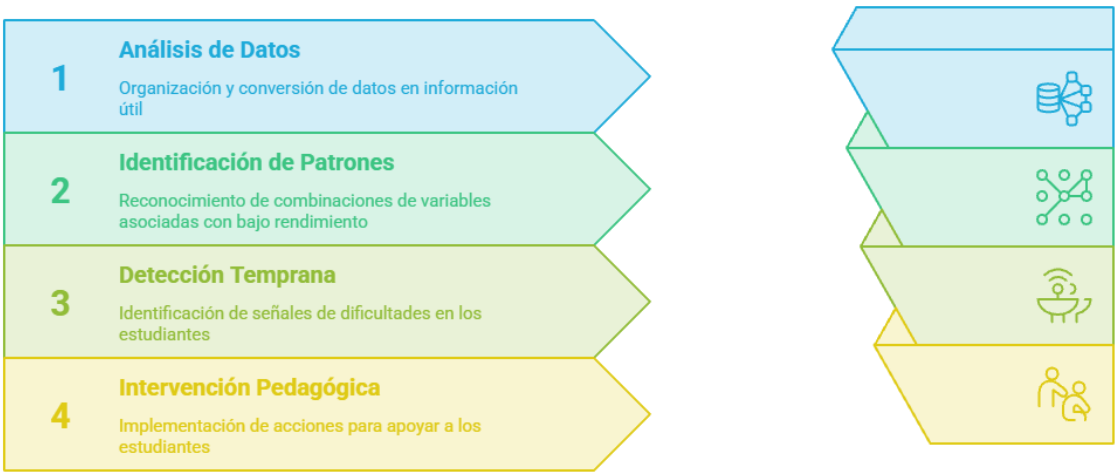
El sesgo algorítmico representa un riesgo central cuando los datos históricos reflejan desigualdades o condiciones específicas de determinados grupos. Baker y Hawn (2022) señalan que los sesgos pueden originarse en la selección de datos, el diseño del modelo y su implementación. Un sistema puede identificar con mayor frecuencia a ciertos estudiantes como “de riesgo” debido a patrones asociados con su contexto y no con su capacidad real. Por ello, resulta necesario analizar errores diferenciados y revisar si las alertas se distribuyen de manera injustificada.

La explicabilidad es indispensable para comprender por qué el sistema genera una alerta o una predicción. Khosravi et al. (2022) sostienen que la inteligencia artificial explicable permite comunicar los factores relevantes que intervienen en una decisión. Docentes y estudiantes necesitan conocer qué indicadores influyeron en el resultado y qué margen de revisión existe.

Sin esta información, una etiqueta predictiva puede parecer objetiva e incuestionable, aunque dependa de supuestos, ponderaciones y datos incompletos. La transparencia facilita la corrección de interpretaciones erróneas.

La intervención pedagógica constituye la finalidad principal de la detección temprana. Una alerta debe activar acciones concretas, como tutoría, retroalimentación adicional, ajuste de actividades o revisión de barreras de acceso. Banihashem et al. (2022) muestran que la analítica puede fortalecer la retroalimentación cuando se convierte en información pertinente para orientar mejoras. Si la predicción solo se utiliza para clasificar o sancionar, pierde su sentido formativo. El valor reside en ampliar apoyos y no en anticipar el fracaso como resultado inevitable.

Figura 7
Proceso de predicción del desempeño educativo



Nota. La figura muestra las etapas para analizar datos, identificar patrones, detectar dificultades e implementar intervenciones pedagógicas oportunas.

La personalización puede apoyarse en estas predicciones para recomendar recursos, rutas o niveles de dificultad. Los sistemas adaptativos ofrecen posibilidades de ajuste según el desempeño observado (Aparicio y Aparicio, 2024). Sin embargo, una predicción inicial no debe encerrar al estudiante en trayectorias simplificadas. Las recomendaciones necesitan revisión constante y posibilidad de cambio conforme aparecen nuevas evidencias. La adaptación responsable amplía oportunidades y evita que una estimación probabilística se convierta en una expectativa permanente sobre lo que el estudiante puede lograr.

La privacidad también condiciona la legitimidad de estos sistemas, porque la predicción requiere recopilar y combinar información detallada sobre comportamiento y rendimiento. Las preocupaciones estudiantiles incluyen el uso, almacenamiento y compartición de datos

generados mediante analítica (Korir et al., 2023). En Ecuador, estas prácticas deben ajustarse a la Ley Orgánica de Protección de Datos Personales (2021) y a su Reglamento General (2023). La institución debe limitar la recopilación, controlar accesos y explicar con claridad la finalidad de los modelos predictivos.

La predicción del desempeño y la detección temprana de dificultades pueden fortalecer el acompañamiento cuando se apoyan en datos pertinentes, interpretaciones prudentes y acciones formativas. Su eficacia no depende únicamente de anticipar resultados, sino de intervenir de manera oportuna y revisable. Los modelos deben funcionar como sistemas de alerta y no como mecanismos de etiquetado. La supervisión docente, la explicabilidad, el análisis de sesgos y la protección de datos son condiciones indispensables para utilizar estas tecnologías sin reducir la trayectoria estudiantil a una probabilidad.

4.7 Alcances, limitaciones y supervisión humana de la evaluación automatizada

La evaluación automatizada puede ampliar la capacidad institucional para procesar grandes volúmenes de respuestas, reducir tiempos de corrección y ofrecer resultados con mayor rapidez. También facilita la identificación de patrones, la generación de retroalimentación y el seguimiento de desempeños sucesivos. Wang et al. (2024) sitúan estas aplicaciones entre los desarrollos más relevantes de la inteligencia artificial educativa. No obstante, su alcance debe interpretarse como apoyo a la evaluación y no como sustitución completa del juicio pedagógico, especialmente cuando se valoran aprendizajes complejos o decisiones de alta consecuencia.

Entre sus principales ventajas se encuentra la consistencia operativa, pues un sistema aplica reglas similares a todas las respuestas y evita variaciones asociadas con fatiga o cambios de atención. Esta uniformidad resulta útil en cuestionarios, tareas estructuradas y grandes grupos. Sin embargo, la consistencia no garantiza validez, porque un modelo puede aplicar de manera estable criterios poco adecuados. La calidad de las inferencias depende de la correspondencia entre constructo, evidencia y puntuación (Zhai et al., 2021). La precisión técnica debe analizarse junto con su pertinencia educativa.

La automatización presenta mayores limitaciones cuando las respuestas incluyen ambigüedad, creatividad, matices lingüísticos o razonamientos no previstos durante el entrenamiento. Haudek y Zhai (2024) muestran que las características del constructo influyen en el desempeño de los modelos de puntuación. Por ello, una herramienta puede funcionar con relativa precisión en tareas específicas y presentar dificultades frente a producciones abiertas. La evaluación de

argumentaciones, ensayos o soluciones originales requiere interpretar intenciones, contexto y coherencia global, dimensiones que no siempre pueden reducirse a patrones estadísticos.

Otra limitación procede de la calidad de los datos utilizados para entrenar y ajustar los sistemas. Si las respuestas de referencia contienen errores, criterios inconsistentes o baja representación de determinados grupos, el modelo puede reproducir esas deficiencias. El sesgo algorítmico puede originarse en los datos, las variables seleccionadas o la manera de implementar la herramienta (Baker y Hawn, 2022). En consecuencia, una evaluación aparentemente neutral puede generar diferencias sistemáticas en las puntuaciones, recomendaciones o alertas producidas para distintos estudiantes.

La explicabilidad constituye una condición esencial para que los resultados automatizados puedan comprenderse, cuestionarse y revisarse. Khosravi et al. (2022) sostienen que los sistemas educativos explicables deben comunicar los factores relevantes que intervienen en sus decisiones. Cuando una puntuación aparece sin fundamento visible, el estudiante enfrenta dificultades para reconocer errores o solicitar una revisión informada. La transparencia debe incluir criterios, limitaciones y procedimientos de reclamación, sin exigir necesariamente que todos los usuarios comprendan los detalles técnicos internos del algoritmo.

Tabla 11
Alcances, limitaciones y funciones de la supervisión humana en la evaluación automatizada

Dimensión	Alcance de la automatización	Limitación principal	Función de la supervisión humana
Corrección de respuestas cerradas	Rapidez y consistencia en grandes grupos	Dependencia de claves y reglas predeterminadas	Revisar calidad de ítems y excepciones
Puntuación de respuestas abiertas	Reconocimiento de patrones y categorías	Dificultad ante matices y respuestas originales	Examinar casos ambiguos y validar criterios
Retroalimentación	Entrega inmediata y personalizada	Mensajes genéricos o diagnósticos erróneos	Contextualizar y ampliar orientaciones
Predicción del desempeño	Detección temprana de patrones de riesgo	Posibilidad de etiquetado y falsas alertas	Interpretar señales y decidir intervenciones
Personalización	Ajuste de rutas y niveles de dificultad	Trayectorias rígidas o simplificadas	Garantizar oportunidades de avance
Gestión de datos	Integración y análisis continuo de registros	Riesgos de privacidad y uso indebido	Limitar accesos y controlar finalidades

Dimensión	Alcance de la automatización	Limitación principal	Función de la supervisión humana
Decisiones de alta consecuencia	Apoyo para promoción o certificación	Error con efectos académicos relevantes	Mantener decisión final y derecho a revisión

Nota. La tabla diferencia las capacidades técnicas de la evaluación automatizada y las responsabilidades que deben conservar las personas responsables del proceso educativo.

La tabla muestra que la supervisión humana no debe reducirse a revisar errores después de que el sistema ha tomado una decisión. Su función comienza en el diseño, continúa durante la aplicación y se mantiene en la interpretación de resultados. Alfredo et al. (2024) destacan la necesidad de enfoques centrados en las personas dentro de la analítica y la inteligencia artificial educativa. Esto implica examinar si la herramienta responde a necesidades reales, si sus resultados son comprensibles y si las decisiones derivadas conservan posibilidad de revisión.

La intervención humana resulta especialmente necesaria en evaluaciones de alta consecuencia, como promoción, certificación o acreditación. La UNESCO (2021) sostiene que la inteligencia artificial debe utilizarse bajo principios de responsabilidad, equidad y control humano. Una puntuación automatizada puede aportar información relevante, pero no debería determinar por sí sola una decisión irreversible. Deben existir mecanismos para revisar discrepancias, considerar evidencias adicionales y corregir errores. La responsabilidad institucional no puede trasladarse al algoritmo ni diluirse bajo la apariencia de objetividad tecnológica.

La privacidad y la protección de datos también limitan el alcance legítimo de la automatización. Los sistemas pueden recopilar respuestas, tiempos, patrones lingüísticos y trayectorias detalladas, aunque no toda información disponible resulta necesaria. Las preocupaciones estudiantiles aumentan cuando se desconoce cómo se utilizan o comparten los datos (Korir et al., 2023). En Ecuador, estas prácticas deben ajustarse a la Ley Orgánica de Protección de Datos Personales (2021) y a su Reglamento General (2023), garantizando finalidad, proporcionalidad, seguridad y control de acceso.

La evaluación automatizada ofrece beneficios importantes en eficiencia, consistencia y capacidad de procesamiento, pero presenta límites vinculados con validez, sesgo, explicabilidad, privacidad y comprensión contextual. Su uso responsable exige supervisión humana activa, no meramente formal. Los sistemas deben complementar el juicio docente, facilitar tareas repetitivas y aportar información para decisiones revisables. Cuando la automatización se convierte en autoridad exclusiva, aumenta el riesgo de consolidar errores y

desigualdades. La calidad depende de mantener la responsabilidad pedagógica y ética en manos de personas e instituciones.

Capítulo

05

Retos éticos y desafíos de la evaluación digital contemporánea



5.1 Privacidad, protección de datos y consentimiento informado

La evaluación digital depende de la recopilación y el tratamiento de información académica, conductual y, en determinados casos, identificativa del estudiante. Las plataformas registran respuestas, calificaciones, tiempos, accesos, interacciones y trayectorias, lo que amplía la capacidad de seguimiento, pero también incrementa los riesgos asociados con la privacidad. Korir et al. (2023) muestran que las preocupaciones estudiantiles abarcan la recopilación, utilización y compartición de datos para analítica del aprendizaje. Por ello, el uso evaluativo de información requiere finalidades explícitas, límites definidos y responsabilidades institucionales verificables.

La privacidad no implica ocultar toda información educativa, sino preservar la capacidad de las personas para conocer y controlar cómo se utilizan sus datos. En contextos evaluativos, esta garantía adquiere especial relevancia porque los registros pueden influir en calificaciones, recomendaciones, alertas y decisiones académicas. Las diferencias culturales afectan la manera en que los estudiantes perciben los riesgos derivados de la analítica, lo que impide asumir una aceptación uniforme de estas prácticas (Viberg et al., 2024). La confianza depende de transparencia, proporcionalidad y posibilidades efectivas de cuestionamiento.

La protección de datos comprende principios y medidas destinadas a evitar accesos indebidos, tratamientos incompatibles y conservación injustificada de información. En Ecuador, la Ley Orgánica de Protección de Datos Personales (2021) establece el marco aplicable al tratamiento de datos personales, mientras su Reglamento General (2023) desarrolla condiciones para su implementación. Las instituciones educativas deben identificar qué información recopilan, con qué fundamento, durante cuánto tiempo y quién puede consultarla. La existencia de una plataforma o herramienta contratada no elimina la responsabilidad institucional sobre el tratamiento realizado.

El principio de finalidad exige que los datos se utilicen únicamente para propósitos educativos determinados y comunicados. Un registro obtenido para retroalimentar una actividad no debería emplearse posteriormente para clasificar estudiantes, investigar conductas o entrenar sistemas sin una justificación compatible. La UNESCO (2021) sostiene que el uso de inteligencia artificial debe respetar derechos, dignidad y control humano. Aplicado a la evaluación, este criterio obliga a diferenciar entre información necesaria para valorar el aprendizaje y datos recopilados porque la tecnología permite obtenerlos fácilmente.

La minimización de datos complementa este principio al exigir que solo se recoja la información estrictamente necesaria. Las plataformas pueden registrar múltiples comportamientos, aunque no todos contribuyen a una interpretación válida del aprendizaje. La analítica educativa corre el riesgo de convertir cada interacción en un dato evaluativo sin comprobar su pertinencia pedagógica (Muñoz et al., 2021). Registrar ubicación, patrones de navegación o características del dispositivo puede resultar desproporcionado cuando el objetivo consiste únicamente en valorar una tarea. La disponibilidad técnica no constituye por sí misma una razón legítima para recopilar información.

El consentimiento informado supone que la persona conozca de forma comprensible qué datos serán tratados, con qué finalidad, quiénes accederán y qué consecuencias pueden derivarse. No debe reducirse a aceptar términos extensos o marcar una casilla antes de ingresar a una plataforma. Korir et al. (2023) evidencian que las preocupaciones estudiantiles se relacionan con distintas etapas del tratamiento de información. Para que el consentimiento sea significativo, la explicación debe utilizar lenguaje claro, evitar ambigüedades y distinguir los usos indispensables de aquellos opcionales o complementarios.

En el ámbito educativo, el consentimiento presenta límites específicos debido a la relación institucional y a la posible ausencia de alternativas reales. Cuando una herramienta es obligatoria para rendir una evaluación, la negativa puede implicar exclusión académica, lo que debilita la libertad de la decisión. Por esta razón, la protección no puede descansar exclusivamente en la aceptación individual. La institución debe justificar la necesidad del tratamiento, ofrecer alternativas razonables y aplicar garantías reforzadas, especialmente cuando participan niños y adolescentes protegidos por el Código de la Niñez y Adolescencia (2003).

Los sistemas de autenticación y supervisión remota plantean riesgos particulares porque pueden recopilar imágenes, voz, documentos de identidad, patrones de comportamiento o información sobre el entorno doméstico. Laamanen et al. (2021) muestran que la aceptación de mecanismos de autenticación electrónica varía entre estudiantes, especialmente cuando existen necesidades educativas especiales o discapacidades. Una medida diseñada para preservar integridad puede resultar invasiva o excluyente si no se evalúan su necesidad y proporcionalidad. La seguridad académica debe equilibrarse con privacidad, accesibilidad y dignidad.

La inteligencia artificial amplía la sensibilidad de los datos porque permite combinar registros y producir inferencias sobre rendimiento, riesgo o comportamiento. Khosravi et al. (2022) destacan que la explicabilidad es necesaria para comprender cómo se generan resultados automatizados. El estudiante debería conocer qué información interviene en una puntuación, recomendación o alerta y disponer de mecanismos de revisión. Las inferencias algorítmicas no dejan de ser datos relevantes por haber sido producidas automáticamente, especialmente cuando condicionan oportunidades, apoyos o decisiones dentro de la trayectoria educativa.

La seguridad de la información exige medidas organizativas y técnicas para impedir pérdida, alteración, filtración o acceso no autorizado. Estas responsabilidades abarcan cuentas, permisos, almacenamiento, copias de seguridad y relación con proveedores tecnológicos. El Reglamento General de la Ley Orgánica de Protección de Datos Personales (2023) refuerza la necesidad de aplicar garantías acordes con los riesgos del tratamiento. La protección no puede limitarse a contraseñas individuales; requiere políticas institucionales, delimitación de funciones, capacitación y procedimientos para responder ante incidentes que comprometan información evaluativa.

La conservación de datos debe limitarse al tiempo necesario para cumplir la finalidad que justificó su recopilación. Mantener indefinidamente respuestas, grabaciones, perfiles o registros de comportamiento aumenta riesgos y facilita usos posteriores no previstos. Stewart et al. (2023) muestran que la dataficación de la educación genera tensiones vinculadas con las creencias y comprensiones del profesorado. Una gobernanza responsable debe establecer plazos de eliminación, condiciones de archivo y procedimientos para ejercer derechos sobre la información, evitando que el historial digital se convierta en una etiqueta permanente.

La privacidad, la protección de datos y el consentimiento informado constituyen condiciones de legitimidad para la evaluación digital, no obstáculos externos a la innovación. Su cumplimiento exige finalidad definida, minimización, transparencia, seguridad, plazos de conservación y mecanismos de revisión. La tecnología evaluativa solo resulta justificable cuando el beneficio pedagógico guarda proporción con la información recopilada y los riesgos asumidos. Las instituciones deben conservar responsabilidad sobre las herramientas empleadas y garantizar que el seguimiento del aprendizaje no se transforme en vigilancia opaca, invasiva o difícil de cuestionar.

5.2 Sesgos algorítmicos, equidad e inclusión educativa

Los sesgos algorítmicos aparecen cuando un sistema automatizado produce resultados sistemáticamente desfavorables, imprecisos o desiguales para determinados estudiantes o grupos. En evaluación digital, pueden influir en puntuaciones, recomendaciones, alertas y decisiones sobre trayectorias académicas. Baker y Hawn (2022) explican que el sesgo educativo no procede únicamente del algoritmo, pues también puede originarse en los datos, las variables seleccionadas y las condiciones de implementación. Su análisis exige examinar todo el proceso evaluativo, desde la definición del aprendizaje hasta la utilización institucional de los resultados.

Los datos de entrenamiento constituyen una fuente crítica de desigualdad. Cuando las respuestas utilizadas para desarrollar un modelo representan de manera insuficiente ciertos contextos lingüísticos, culturales o educativos, el sistema puede interpretar con menor precisión producciones legítimas. Una puntuación técnicamente consistente puede favorecer estilos de respuesta predominantes y penalizar otras formas de expresar conocimientos (Baker y Hawn, 2022). El problema no se resuelve únicamente aumentando la cantidad de datos; se requiere revisar su diversidad, calidad, procedencia y relación con la población donde será aplicada la herramienta.

La definición del constructo también puede introducir sesgos cuando reduce aprendizajes complejos a características fáciles de procesar. Haudek y Zhai (2024) muestran que las propiedades del constructo evaluado influyen en el desempeño de los modelos de puntuación automatizada. Si la argumentación, creatividad o comprensión se representan mediante indicadores superficiales, el sistema puede privilegiar extensión, vocabulario o estructura, sin captar la calidad del razonamiento. La equidad exige comprobar que los criterios automatizados valoren aspectos académicamente relevantes y no características asociadas indirectamente con determinados grupos.

La equidad evaluativa no significa producir resultados idénticos, sino garantizar oportunidades comparables para demostrar los aprendizajes previstos. Las diferencias de conectividad, disponibilidad de dispositivos y preparación digital pueden afectar el desempeño sin reflejar dominio académico. La brecha digital comprende condiciones sociales, familiares e institucionales que influyen en la educación en línea (Mathrani et al., 2022). Además, la preparación desigual de estudiantes y centros educativos modifica las posibilidades de participación y aprovechamiento tecnológico (Van de Werfhorst et al., 2022), lo que debe considerarse al interpretar resultados.

La inclusión requiere que las herramientas sean accesibles para estudiantes con diferentes necesidades y formas de interacción. Scanlon et al. (2021) identifican errores de accesibilidad web que generan barreras para personas con discapacidad, incluso cuando los contenidos son adecuados. Las directrices del CAST (2024) proponen múltiples formas de implicación, representación y expresión para ampliar la participación. En evaluación, estos principios permiten diversificar formatos y vías de respuesta sin modificar los criterios esenciales. Una plataforma inaccesible puede convertir una limitación tecnológica en una desventaja académica injustificada.

Tabla 12

Fuentes de sesgo y criterios de equidad en la evaluación digital automatizada

Fuente de sesgo	Manifestación evaluativa	Grupo o dimensión afectada	Criterio de revisión	Medida de mitigación
Datos poco representativos	Mayor error en determinadas respuestas	Diversidad lingüística, cultural o educativa	Comparar precisión entre grupos	Ampliar y equilibrar los datos
Constructo simplificado	Puntuación basada en rasgos superficiales	Razonamiento y producción compleja	Revisar correspondencia con objetivos	Incorporar evidencias complementarias
Diseño inaccesible	Dificultad para navegar o responder	Estudiantes con discapacidad	Evaluar accesibilidad y compatibilidad	Ofrecer formatos y medios alternativos
Brecha digital	Resultados afectados por acceso desigual	Estudiantes con menor conectividad	Analizar condiciones de aplicación	Flexibilizar tiempos y modalidades
Predicción algorítmica	Alertas o rutas diferenciadas recurrentes	Trayectorias consideradas de riesgo	Examinar falsas alertas y distribución	Mantener revisión docente
Implementación institucional	Uso de resultados fuera de contexto	Estudiantes con datos incompletos	Revisar finalidad y consecuencias	Limitar decisiones automatizadas
Falta de transparencia	Imposibilidad de cuestionar resultados	Toda la comunidad educativa	Explicar criterios y variables	Establecer mecanismos de reclamación

Nota. La tabla relaciona las principales fuentes de sesgo con criterios y medidas orientadas a fortalecer la equidad y la inclusión en la evaluación automatizada.

La tabla evidencia que el sesgo puede aparecer en distintas etapas y no únicamente durante la ejecución del algoritmo. La explicabilidad permite reconocer qué variables influyeron en una puntuación, recomendación o alerta y facilita revisar resultados cuestionables. Khosravi et al. (2022) sostienen que los sistemas educativos explicables deben proporcionar información

comprensible a sus usuarios. Sin transparencia, las diferencias producidas por un modelo pueden presentarse como resultados objetivos, aunque respondan a decisiones técnicas, supuestos pedagógicos o datos que requieren revisión crítica.

La detección del sesgo debe incluir análisis diferenciados del desempeño del sistema. No basta con informar una precisión global elevada si los errores se concentran en determinados estudiantes. Baker y Hawn (2022) recomiendan examinar las múltiples formas en que el sesgo puede afectar las aplicaciones educativas. Esta revisión debe comparar puntuaciones automatizadas con valoraciones humanas, analizar falsos positivos y negativos y observar las consecuencias de las decisiones. La equidad se evalúa tanto en el funcionamiento técnico como en los efectos pedagógicos e institucionales producidos.

La supervisión humana constituye una garantía necesaria, especialmente cuando los resultados afectan promoción, apoyos o acceso a oportunidades. Los enfoques centrados en las personas buscan articular inteligencia artificial, analítica y necesidades educativas sin desplazar la agencia humana (Alfredo et al., 2024). El docente debe revisar casos atípicos, considerar evidencias adicionales y corregir decisiones que no representen adecuadamente el aprendizaje. Esta intervención no debe limitarse a validar automáticamente los resultados, sino incluir capacidad real para cuestionarlos, modificarlos y explicar las razones de la decisión final.

La reducción de los sesgos algorítmicos exige combinar datos representativos, constructos válidos, accesibilidad, transparencia y supervisión humana. La UNESCO (2021) plantea que la inteligencia artificial debe respetar la equidad, los derechos humanos y la diversidad. Una evaluación inclusiva no se alcanza solo mediante ajustes técnicos, pues también requiere revisar las condiciones sociales e institucionales donde funciona la herramienta. Los sistemas automatizados adquieren legitimidad cuando amplían oportunidades, permiten impugnar resultados y evitan que diferencias tecnológicas, culturales o funcionales se conviertan en desigualdades educativas permanentes.

5.3 Integridad académica, plagio y uso responsable de la inteligencia artificial

La integridad académica comprende principios de honestidad, responsabilidad y transparencia aplicados a la producción y evaluación del aprendizaje. En entornos digitales, su protección se vuelve más compleja debido al acceso inmediato a información, herramientas de edición, plataformas colaborativas y sistemas de inteligencia artificial generativa. Holden et al. (2021) muestran que la evaluación en línea enfrenta desafíos relacionados con autoría, autenticidad y supervisión. Estos problemas no deben interpretarse únicamente como conductas individuales,

pues también dependen del diseño de las tareas, las reglas institucionales y el acompañamiento pedagógico.

El plagio consiste en presentar ideas, expresiones o producciones ajenas como propias sin el reconocimiento correspondiente. En entornos virtuales, copiar, reutilizar textos o compartir respuestas puede realizarse con mayor facilidad, aunque la disponibilidad tecnológica no explica por sí sola estas prácticas. Las percepciones sobre integridad varían entre docentes y estudiantes, especialmente cuando las normas no son suficientemente claras (Alessio y Messinger, 2021). La prevención exige enseñar criterios de citación, autoría y colaboración permitida, en lugar de limitarse a comunicar sanciones después de detectar una infracción.

La evaluación digital puede favorecer conductas deshonestas cuando se apoya en preguntas previsibles, respuestas fácilmente localizables o tareas idénticas para todos los participantes. Holden et al. (2021) advierten que el diseño evaluativo influye en los riesgos de integridad presentes en contextos en línea. Las actividades contextualizadas, los procesos de revisión, las defensas orales y la justificación de decisiones reducen la utilidad de copiar respuestas terminadas. Esta estrategia no elimina toda posibilidad de fraude, pero desplaza la atención desde la vigilancia hacia la producción de evidencias más auténticas y verificables.

La aparición de inteligencia artificial generativa modifica la discusión porque permite producir textos, explicaciones, resúmenes y respuestas con apariencia académica. Su uso no constituye automáticamente una infracción; depende del propósito de la actividad, las reglas comunicadas y el grado de intervención permitido. Tan y Maravilla (2024) sostienen que la inteligencia artificial generativa no tiene que debilitar la educación si se integra con criterios de responsabilidad. La cuestión central consiste en distinguir entre apoyo legítimo, asistencia no declarada y sustitución del trabajo intelectual esperado.

El uso responsable exige transparencia sobre la intervención de la inteligencia artificial. Cuando una herramienta se emplea para generar ideas, revisar estructura o proponer mejoras, el estudiante debe conservar la responsabilidad sobre la precisión, coherencia y originalidad del resultado. La UNESCO (2023) recomienda utilizar la inteligencia artificial generativa con supervisión humana y reconocimiento de sus limitaciones. Aceptar una respuesta automática sin verificarla puede introducir errores, referencias inexistentes o argumentos inconsistentes. La responsabilidad académica no se transfiere al sistema, incluso cuando su participación ha sido autorizada.

La retroalimentación generada por inteligencia artificial ofrece posibilidades formativas, especialmente en procesos de escritura. Ossa y Willatt (2023) analizan su utilización para orientar producciones académicas en la formación inicial docente. Este apoyo puede facilitar revisiones frecuentes y ofrecer sugerencias inmediatas, pero también fomentar dependencia cuando el estudiante acepta cambios sin comprenderlos. Para preservar el aprendizaje, las actividades deben requerir explicación de decisiones, comparación entre versiones y reflexión sobre las recomendaciones recibidas. La herramienta debe apoyar la revisión, no reemplazar el razonamiento ni la voz del autor.

La detección automática de plagio y de textos generados por inteligencia artificial presenta límites técnicos y éticos. Un resultado probabilístico no constituye una prueba concluyente de deshonestidad, especialmente cuando puede producir falsos positivos. La evaluación automatizada necesita explicabilidad, revisión y posibilidad de cuestionar sus resultados (Khosravi et al., 2022). Utilizar una puntuación de similitud o detección como fundamento exclusivo para sancionar puede generar decisiones injustas. Estos indicadores deben iniciar una revisión pedagógica y documental, no sustituir el análisis de las evidencias ni el diálogo con el estudiante.

Las medidas de supervisión remota tampoco resuelven por sí solas los problemas de integridad. Sistemas de autenticación, cámaras y monitoreo pueden aumentar el control, pero también afectar privacidad, accesibilidad y confianza. La aceptación de mecanismos de autenticación electrónica varía, especialmente entre estudiantes con necesidades educativas especiales o discapacidades (Laamanen et al., 2021). Una política proporcional debe valorar el nivel de riesgo, la naturaleza de la evaluación y las alternativas disponibles. La integridad no puede protegerse mediante prácticas invasivas que generen nuevas formas de exclusión.

Figura 8
Desafíos de la integridad académica en la era digital



Nota. La figura sintetiza los principales desafíos para preservar la integridad académica en los entornos digitales.

Las instituciones necesitan establecer normas específicas sobre usos permitidos, restringidos y prohibidos de la inteligencia artificial. Expresiones generales como “no usar IA” resultan insuficientes cuando las herramientas se integran en procesadores, buscadores y plataformas educativas. Temper et al. (2025) proponen un marco para regular su utilización en instituciones de educación superior, destacando la necesidad de gobernanza y responsabilidades definidas. Las políticas deben diferenciar actividades, explicar formas de reconocimiento y establecer procedimientos de revisión, evitando respuestas uniformes ante situaciones académicamente distintas.

La formación ética debe incluir capacidades para verificar información, reconocer límites, atribuir apoyos y asumir responsabilidad por el producto entregado. La UNESCO (2021) vincula la ética de la inteligencia artificial con transparencia, equidad y supervisión humana. En evaluación, estos principios exigen que estudiantes y docentes comprendan cómo las herramientas pueden influir en la autoría y en la interpretación del desempeño. La alfabetización no debe restringirse al manejo técnico; requiere juicio crítico para decidir cuándo utilizar una herramienta, cómo revisar sus resultados y qué usos comprometen el propósito formativo.

La integridad académica en contextos digitales no se protege únicamente mediante detectores, sanciones o vigilancia. Requiere tareas auténticas, criterios transparentes, formación sobre autoría y reglas diferenciadas para el uso de inteligencia artificial. La tecnología puede apoyar la escritura, la retroalimentación y el aprendizaje, pero su intervención debe declararse cuando corresponda y mantenerse bajo responsabilidad humana. Un enfoque educativo sólido combina prevención, acompañamiento y revisión justa, evitando tanto la aceptación acrítica de la inteligencia artificial como su tratamiento automático como evidencia de fraude.

5.4 Brecha digital y desigualdades en el acceso a la evaluación

La brecha digital en la evaluación educativa comprende las desigualdades relacionadas con acceso a dispositivos, conectividad, competencias tecnológicas y condiciones para participar en actividades en línea. Estas diferencias pueden alterar el desempeño y producir resultados que reflejen limitaciones de acceso antes que aprendizajes reales. Mathrani et al. (2022) explican que la brecha digital en países en desarrollo incluye dimensiones tecnológicas, sociales e institucionales. Por ello, la evaluación digital requiere analizar no solo los instrumentos empleados, sino también las condiciones materiales desde las cuales cada estudiante responde.

El acceso a internet constituye una de las expresiones más visibles de esta desigualdad. Las conexiones inestables, los planes de datos limitados y la dependencia de redes compartidas pueden interrumpir exámenes, retrasar entregas o impedir la participación en actividades sincrónicas. La preparación digital desigual de estudiantes y centros educativos condiciona las oportunidades de aprendizaje en línea (Van de Werfhorst et al., 2022). Cuando estas diferencias no se consideran, una falla técnica puede interpretarse como falta de preparación, escaso compromiso o incumplimiento académico, generando decisiones evaluativas injustas.

La disponibilidad de dispositivos también influye en la calidad de la participación. No todas las actividades pueden realizarse con igual facilidad desde un teléfono móvil, una tableta o una computadora. Algunas plataformas requieren pantallas amplias, programas específicos o capacidad de procesamiento que no se encuentra disponible para todos. Mathrani et al. (2022) muestran que el acceso digital debe analizarse dentro de contextos familiares y educativos más amplios. Compartir un dispositivo con otros integrantes del hogar puede limitar tiempos de estudio y producir retrasos que no expresan menor capacidad académica.

Las competencias digitales constituyen otra dimensión de la brecha. El acceso físico no garantiza que el estudiante pueda navegar, cargar archivos, utilizar herramientas o resolver

problemas técnicos durante una evaluación. Sillat et al. (2021) señalan que la competencia digital comprende diversas capacidades y requiere métodos específicos de valoración. En consecuencia, una prueba puede terminar midiendo familiaridad tecnológica junto con el conocimiento disciplinar. Para evitar esta distorsión, se necesitan instrucciones claras, prácticas previas y soporte técnico que permitan diferenciar el dominio del contenido de la destreza instrumental.

Las condiciones del entorno doméstico afectan especialmente a las evaluaciones remotas. El ruido, la falta de privacidad, las responsabilidades familiares y los espacios compartidos pueden limitar la concentración o impedir la participación en horarios rígidos. La brecha digital no se reduce al equipamiento, porque también comprende las posibilidades reales de utilizarlo en condiciones apropiadas (Mathrani et al., 2022). Una evaluación equitativa debe contemplar flexibilidad temporal, alternativas asincrónicas y procedimientos frente a interrupciones, sin convertir circunstancias contextuales en desventajas permanentes para el estudiante.

La accesibilidad agrega una dimensión específica vinculada con estudiantes con discapacidad o necesidades educativas diversas. Los errores de navegación, la ausencia de subtítulos, los tiempos estrictos y la incompatibilidad con tecnologías de apoyo pueden restringir la participación (Scanlon et al., 2021). Las directrices del CAST (2024) promueven múltiples formas de implicación, representación y expresión. Aplicadas a la evaluación, permiten ofrecer formatos alternativos y distintas vías de respuesta, manteniendo los criterios esenciales. La inclusión no consiste en reducir exigencias, sino en eliminar barreras irrelevantes para demostrar aprendizajes.

Las desigualdades también pueden profundizarse cuando las instituciones adoptan herramientas sin considerar infraestructura, formación y soporte. Tsai et al. (2021) muestran que la adopción de analítica del aprendizaje depende de experiencias, prioridades y capacidades organizativas. Una institución con recursos limitados puede implementar plataformas sin acompañamiento suficiente, generando usos fragmentados y dificultades recurrentes. La innovación tecnológica requiere inversión sostenida, protocolos de contingencia y formación docente. De lo contrario, la evaluación digital puede trasladar a estudiantes y profesores la responsabilidad de resolver problemas estructurales.

Las pruebas sincrónicas de tiempo limitado suelen amplificar estas desigualdades. Una desconexión breve, una plataforma lenta o un dispositivo compartido puede tener consecuencias desproporcionadas cuando no existen posibilidades de reingreso o recuperación.

La evaluación en línea requiere diseños compatibles con las condiciones reales de participación (Heil y Ifenthaler, 2023). La flexibilidad no implica eliminar controles, sino establecer mecanismos proporcionales: ventanas amplias de acceso, guardado automático, intentos justificados y alternativas equivalentes ante fallas. Estas medidas protegen la validez del resultado y la continuidad académica.

La desigualdad en el acceso también afecta la cantidad y calidad de datos disponibles para analítica del aprendizaje. Un estudiante con conectividad intermitente puede generar menos registros y aparecer como poco participativo, aunque desarrolle actividades fuera de la plataforma. Los indicadores digitales necesitan interpretación contextual y contraste con otras evidencias (Muñoz et al., 2021). Utilizar accesos o tiempos de conexión como medidas directas de compromiso puede perjudicar a quienes enfrentan mayores limitaciones tecnológicas. La ausencia de datos no debe convertirse automáticamente en evidencia de ausencia de aprendizaje.

Las políticas de equidad deben combinar recursos institucionales, diseño accesible y alternativas evaluativas. Proporcionar dispositivos o conectividad puede reducir barreras materiales, pero no resuelve por sí solo diferencias de competencias, accesibilidad y acompañamiento. Van de Werfhorst et al. (2022) destacan que la preparación digital involucra tanto a estudiantes como a instituciones. Por ello, las respuestas deben incluir formación, soporte, flexibilidad y revisión de instrumentos. La evaluación necesita múltiples evidencias para evitar que una única condición tecnológica determine el resultado académico o limite oportunidades de progresión.

La brecha digital constituye un problema de justicia evaluativa porque afecta quién puede participar, en qué condiciones y con qué posibilidades de demostrar el aprendizaje. Las tecnologías pueden ampliar acceso y flexibilidad, pero también reproducir desigualdades cuando se implementan sin considerar recursos, competencias, accesibilidad y contexto. Una evaluación digital equitativa debe anticipar barreras, ofrecer alternativas y revisar los resultados afectados por condiciones técnicas. La calidad no depende solo de la plataforma utilizada, sino de garantizar que las diferencias de acceso no se conviertan en diferencias artificiales de rendimiento.

5.5 Transparencia, explicabilidad y responsabilidad en las decisiones evaluativas

La transparencia en la evaluación digital exige que estudiantes y docentes conozcan los propósitos, criterios, datos y procedimientos que intervienen en una decisión académica. Esta

obligación adquiere mayor relevancia cuando sistemas automatizados asignan puntuaciones, recomiendan actividades o generan alertas sobre el desempeño. Khosravi et al. (2022) sostienen que la inteligencia artificial explicable debe ofrecer información comprensible sobre los factores que influyen en sus resultados. Sin transparencia, una decisión técnica puede presentarse como objetiva, aunque dependa de supuestos pedagógicos y elecciones institucionales discutibles.

La explicabilidad complementa la transparencia al responder por qué un sistema produjo un resultado específico. No basta con informar que se utilizó inteligencia artificial; es necesario identificar los criterios relevantes, las evidencias consideradas y las limitaciones del procedimiento. Las explicaciones deben adaptarse a las necesidades de los usuarios y permitir que una puntuación pueda comprenderse y cuestionarse (Khosravi et al., 2022). Un sistema puede ser técnicamente complejo sin resultar completamente opaco, siempre que comunique de forma clara los elementos necesarios para interpretar sus decisiones.

La transparencia también debe abarcar el diseño del instrumento. Los estudiantes necesitan conocer qué aprendizaje se valora, cómo se ponderan los criterios y qué evidencias resultan pertinentes. Gu (2021) afirma que la validez depende de argumentos que conecten desempeño, interpretación y decisión. Cuando estos vínculos no son visibles, la calificación pierde capacidad orientadora y dificulta la revisión. Los criterios explícitos permiten comprender el resultado, mientras que las reglas ocultas aumentan la dependencia del sistema o del evaluador y reducen la posibilidad de ejercer una defensa informada.

La responsabilidad implica identificar quién responde por los efectos de una decisión automatizada. La institución, el proveedor y el docente pueden intervenir en diferentes fases, pero la existencia de un algoritmo no elimina la obligación humana de justificar resultados. La UNESCO (2021) establece que el uso de inteligencia artificial debe conservar control, responsabilidad y supervisión humana. Por ello, las instituciones no pueden atribuir un error exclusivamente a la herramienta. Deben establecer funciones, procedimientos de revisión y autoridades competentes para corregir decisiones que produzcan consecuencias académicas injustificadas.

La rendición de cuentas exige conservar documentación sobre el funcionamiento del sistema, los datos utilizados y las modificaciones realizadas. Esta trazabilidad permite investigar discrepancias, identificar patrones de error y verificar si los resultados mantienen calidad a lo largo del tiempo. Las evaluaciones digitales administradas continuamente necesitan

mecanismos de supervisión y control de consistencia (Liao et al., 2022). Sin registros suficientes, resulta difícil determinar por qué cambió una puntuación o quién autorizó determinada configuración. La documentación constituye una garantía para usuarios, docentes e instituciones.

Tabla 13
Condiciones de transparencia, explicabilidad y responsabilidad evaluativa

Dimensión	Pregunta central	Información necesaria	Responsabilidad institucional	Riesgo ante su ausencia
Propósito	¿Para qué se utiliza el sistema?	Finalidad y decisiones asociadas	Limitar el uso a objetivos definidos	Uso de datos fuera del contexto
Criterios	¿Qué se evalúa y cómo se pondera?	Constructos, indicadores y niveles	Comunicar reglas antes de la evaluación	Calificaciones difíciles de comprender
Datos	¿Qué información se procesa?	Fuentes, variables y periodos de conservación	Garantizar pertinencia y protección	Recopilación excesiva u opaca
Explicación	¿Por qué se produjo el resultado?	Factores relevantes y limitaciones	Ofrecer razones comprensibles	Dependencia de resultados incuestionables
Supervisión	¿Quién revisa la decisión?	Autoridades y procedimientos	Mantener intervención humana efectiva	Delegación total al sistema
Reclamación	¿Cómo puede cuestionarse el resultado?	Plazos, canales y evidencias admitidas	Garantizar revisión imparcial	Indefensión del estudiante
Trazabilidad	¿Qué ocurrió durante el proceso?	Registros, versiones y modificaciones	Conservar documentación verificable	Imposibilidad de determinar responsabilidades

Nota. La tabla organiza las condiciones necesarias para que las decisiones evaluativas digitales sean comprensibles, revisables y atribuibles a responsables identificables.

La tabla evidencia que la transparencia no consiste únicamente en publicar información técnica. Debe permitir que los afectados comprendan la decisión y actúen frente a ella. Alfredo et al. (2024) destacan la necesidad de enfoques centrados en las personas dentro de la analítica y la inteligencia artificial educativa. Esto implica diseñar explicaciones útiles, canales de consulta y procedimientos de revisión. Una descripción extensa del algoritmo carece de valor si el estudiante continúa sin saber qué criterio incumplió ni cómo puede corregir o impugnar el resultado.

El derecho a revisión humana resulta esencial cuando la decisión produce consecuencias relevantes sobre promoción, certificación o acceso a oportunidades. Las puntuaciones

automatizadas pueden apoyar la deliberación, pero no deberían operar como resultados irreversibles. Los sistemas educativos explicables deben facilitar la interpretación y el cuestionamiento de sus recomendaciones (Khosravi et al., 2022). La revisión requiere una persona con autoridad para examinar otras evidencias, modificar la decisión y justificar el resultado final. Una validación meramente formal no constituye supervisión humana efectiva.

La responsabilidad se relaciona también con el análisis de sesgos y efectos diferenciados. Baker y Hawn (2022) advierten que las desigualdades algorítmicas pueden originarse en datos, modelos e implementación. Una institución responsable debe comparar errores entre grupos, investigar resultados atípicos y corregir patrones injustificados. Informar una precisión promedio no basta cuando determinados estudiantes reciben con mayor frecuencia puntuaciones erróneas o alertas negativas. La rendición de cuentas exige valorar no solo el funcionamiento técnico, sino también las consecuencias educativas generadas por el sistema.

La transparencia, la explicabilidad y la responsabilidad convierten la evaluación automatizada en un proceso susceptible de comprensión, revisión y corrección. Estas condiciones no eliminan todos los errores, pero impiden que permanezcan ocultos bajo la autoridad técnica del algoritmo. La legitimidad depende de criterios visibles, datos pertinentes, explicaciones accesibles, supervisión humana y mecanismos de reclamación. Cuando las instituciones conservan responsabilidad efectiva, la tecnología puede apoyar decisiones evaluativas; cuando la delegan, aumentan la opacidad, la indefensión y el riesgo de consolidar resultados injustos.

5.6 Gobernanza y perspectivas futuras de la evaluación digital

La gobernanza de la evaluación digital comprende las normas, responsabilidades y procedimientos que orientan la selección, aplicación y supervisión de tecnologías evaluativas. Su finalidad consiste en asegurar que plataformas, sistemas automatizados y herramientas de inteligencia artificial respondan a objetivos educativos legítimos. Temper et al. (2025) proponen un marco institucional para regular el uso de inteligencia artificial en la educación superior. Esta necesidad surge porque las decisiones tecnológicas no son meramente operativas: pueden afectar calificaciones, privacidad, equidad y oportunidades académicas de los estudiantes.

Una gobernanza sólida requiere distribuir responsabilidades entre autoridades, docentes, equipos técnicos y proveedores. La contratación de una herramienta externa no transfiere al proveedor toda la responsabilidad sobre sus efectos educativos. La UNESCO (2021) sostiene que los sistemas de inteligencia artificial deben mantenerse bajo control y responsabilidad

humana. En evaluación, esto implica determinar quién valida el instrumento, quién revisa errores, quién autoriza el tratamiento de datos y qué instancia responde ante reclamaciones. La falta de funciones definidas favorece decisiones opacas y dificulta corregir consecuencias injustificadas.

Las políticas institucionales deben diferenciar usos permitidos, condicionados y prohibidos de las tecnologías evaluativas. Una regulación demasiado general puede resultar insuficiente frente a herramientas que puntúan respuestas, predicen dificultades o generan retroalimentación. Temper et al. (2025) destacan la necesidad de establecer marcos específicos para el uso de inteligencia artificial en instituciones de educación superior. Estas disposiciones deben considerar el nivel de consecuencia de cada decisión, exigir mayor supervisión en procesos de acreditación o promoción y permitir aplicaciones más flexibles cuando la tecnología cumple funciones de apoyo formativo.

La evaluación previa de las herramientas constituye una responsabilidad central de la gobernanza. Antes de su implementación, deben analizarse validez, accesibilidad, seguridad, sesgos, explicabilidad e integración con los objetivos institucionales. Los sistemas educativos centrados en las personas requieren equilibrar capacidades técnicas con necesidades y experiencias de los usuarios (Alfredo et al., 2024). Una herramienta eficiente puede resultar inadecuada si produce barreras, recopila datos desproporcionados o genera resultados difíciles de interpretar. La evaluación técnica debe complementarse con revisión pedagógica y ética.

La gestión de datos ocupa un lugar prioritario porque la evaluación digital produce registros detallados sobre desempeño y comportamiento. En Ecuador, la Ley Orgánica de Protección de Datos Personales (2021) y su Reglamento General (2023) establecen obligaciones aplicables a su tratamiento. La gobernanza debe definir finalidades, periodos de conservación, niveles de acceso y procedimientos de eliminación. También necesita delimitar el uso de datos para analítica o entrenamiento de modelos, evitando que la información recopilada con un propósito académico se reutilice sin justificación, conocimiento o garantías suficientes.

La participación de docentes y estudiantes fortalece la legitimidad de las decisiones institucionales. Stewart et al. (2023) muestran que las creencias del profesorado influyen en la comprensión de los procesos de datificación educativa. Las políticas diseñadas exclusivamente desde perspectivas técnicas pueden ignorar problemas de aula, accesibilidad o carga de trabajo. Incluir a los usuarios permite identificar riesgos, evaluar la comprensión de los sistemas y

ajustar procedimientos. La participación no sustituye la responsabilidad institucional, pero aporta evidencia para construir normas aplicables y sensibles a las condiciones reales.

La formación continua será decisiva para el futuro de la evaluación digital. El profesorado necesita capacidades para seleccionar herramientas, interpretar datos, revisar resultados automatizados y comunicar sus limitaciones. Sánchez (2023) identifica diversos usos y posibilidades de la inteligencia artificial como recurso docente, lo que evidencia la necesidad de preparación pedagógica y crítica. La capacitación no debe limitarse al manejo técnico de plataformas; requiere abordar privacidad, sesgo, explicabilidad, integridad y criterios de intervención humana. Sin estas capacidades, las normas pueden existir formalmente sin modificar las prácticas.

Las perspectivas futuras apuntan hacia evaluaciones más continuas, adaptativas y multimodales, capaces de integrar respuestas, interacciones, simulaciones y producciones complejas. Wang et al. (2024) muestran que la inteligencia artificial continúa ampliando sus aplicaciones educativas, especialmente en personalización, seguimiento y retroalimentación. Este desarrollo puede mejorar la oportunidad de las intervenciones, pero también ampliar la vigilancia y dependencia tecnológica. La innovación debe orientarse por el valor pedagógico de las evidencias y no por la posibilidad de recopilar o procesar una cantidad creciente de datos.

La regulación futura combinará principios internacionales, normativa nacional y procedimientos internos. El Reglamento de Inteligencia Artificial de la Unión Europea establece un enfoque basado en riesgos para sistemas que pueden afectar derechos y oportunidades (Parlamento Europeo y Consejo de la Unión Europea, 2024). Aunque su aplicación jurídica no es universal, ofrece criterios relevantes sobre transparencia, supervisión y responsabilidad. En el contexto ecuatoriano, estas orientaciones deben articularse con la Constitución de la República del Ecuador (2008), la normativa de protección de datos y las particularidades del sistema educativo.

La gobernanza futura de la evaluación digital deberá equilibrar innovación, derechos y calidad pedagógica mediante decisiones revisables y responsabilidades identificables. La UNESCO (2023) recomienda que el uso educativo de inteligencia artificial generativa conserve supervisión humana y atención a sus limitaciones. Ningún marco puede anticipar todas las tecnologías emergentes, por lo que las instituciones necesitan procedimientos flexibles de evaluación, monitoreo y actualización. El futuro no dependerá únicamente de sistemas más

avanzados, sino de la capacidad para gobernarlos con transparencia, equidad y criterios educativos consistentes.

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, R. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *Computers and Education: Artificial Intelligence, 6*. https://doi.org/10.1016/j.caeai.2024.100215

Aparicio, O., y Aparicio, W. (2024). Innovación educativa con sistemas de aprendizaje adaptativo impulsados por Inteligencia Artificial. *Revista Internacional De Pedagogía E Innovación Educativa, 4*(2), 343-363. https://doi.org/10.51660/ripie42222

Baker, R., y Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education, 32*(4), 1052-1092. https://doi.org/10.1007/s40593-021-00285-9

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*. https://doi.org/10.1016/j.edurev.2022.100489

Carrasco, Á., Cacoango, W., y Maliza, W. (2025). Analítica del aprendizaje sustentada en Educaplay y Kahoot como medios de enseñanza de la asignatura de Emprendimiento y Gestión. *MQRInvestigar, 9*(1). https://doi.org/10.56048/MQR20225.9.1.2025.e5

CAST. (2024). *Directrices de diseño universal para el aprendizaje de CAST, versión 3.0*. https://udlguidelines.cast.org/

Código de la Niñez y Adolescencia. (2003). Registro Oficial No. 737: https://www.igualdad.gob.ec/wp-content/uploads/downloads/2017/11/codigo_ninezyadolescencia.pdf

Constitución de la República del Ecuador. (2008). Registro Oficial No. 449: https://www.defensa.gob.ec/wp-content/uploads/downloads/2021/02/Constitucion-de-la-Republica-del-Ecuador_act_ene-2021.pdf

de Schipper, E. d., Feskens, R., y Keuning, J. (2021). Personalized and automated feedback in summative assessment using recommender systems. *Frontiers in Education, 6*, 1-11. 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*, 1-13. https://doi.org/10.1186/s12909-022-03912-8

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

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

González, L., y Lugo, C. (2021). Fortalecimiento de la práctica docente con Learning Analytics: estudio de caso. *Praxis & Saber, 11*(25), 227-254. https://doi.org/10.19053/22160159.v11.n25.2020.9075

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

Haudek, K., y Zhai, X. (2024). Examining the effect of assessment construct characteristics on machine learning scoring of scientific argumentation. *International Journal of Artificial Intelligence in Education, 34*, 1482-1509. https://doi.org/10.1007/s40593-023-00385-8

Heil, J., y Ifenthaler, D. (2023). Online Assessment in Higher Education: A Systematic Review. *Online Learning, 27*(1). 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*. https://doi.org/10.3389/feduc.2021.639814

Huauya, L., Solis, B., y Vacas, F. (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

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*, 1-22. https://doi.org/10.1016/j.caeai.2022.100074

Korir, M., Slade, S., Holmes, W., Héliot, Y., y Rienties, B. (2023). Investigating the dimensions of students' privacy concern in the collection, use, and sharing of data for learning analytics. *Computers in Human Behavior Reports, 9*. https://doi.org/10.1016/j.chbr.2022.100262

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

Lee, H.-S., Gweon, G.-H., Lord, T., Paessel, N., Pallant, A., y Pryputniewicz, S. (2021). Machine learning-enabled automated feedback: Supporting students' revision of scientific arguments based on data drawn from simulation. *Journal of Science Education and Technology, 30*(2), 168-192. https://doi.org/10.1007/s10956-020-09889-7

Ley Orgánica de Protección de Datos Personales. (2021). Quinto Suplemento del Registro Oficial No. 459: https://www.finanzaspopulares.gob.ec/wp-content/uploads/2021/07/ley_organica_de_proteccion_de_datos_personales.pdf

Liao, M., Attali, Y., Lockwood, J., y von Davier, A. (2022). Maintaining and monitoring quality of a continuously administered digital assessment. *Frontiers in Education, 7*. https://doi.org/10.3389/feduc.2022.857496

Maestrales, S., Zhai, X., Touitou, I., Baker, Q., Schneider, B., y Krajcik, J. (2021). Using machine learning to score multi-dimensional assessments of chemistry and physics. *Journal of Science Education and Technology, 30*(2), 239-254. https://doi.org/10.1007/s10956-020-09895-9

Martínez, M., Encalada, G., Villamar, V., Zhindón, E., y Matute, M. (2025). Gamificación como Herramienta Pedagógica para Potenciar la Autorregulación en el Aprendizaje.

Estudios Y Perspectivas Revista Científica Y Académica, 5(3). https://doi.org/10.61384/r.c.a..v5i3.1448

Mathrani, A., Sarvesh, T., y Umer, R. (2022). Digital divide framework: Online learning in developing countries during the COVID-19 lockdown. *Globalisation, Societies and Education, 20*(5), 625-640. https://doi.org/10.1080/14767724.2021.1981253

Muñoz, J., Ortega, C., y Velasquez, J. (2021). Tecnología y analítica del aprendizaje: una revisión a la literatura. *Revista científica*(41), 150-168. http://www.scielo.org.co/scielo.php?pid=S0124-22532021000200150&script=sci_arttext

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), 5. https://dialnet.unirioja.es/servlet/articulo?codigo=9235913

Parlamento Europeo y Consejo de la Unión Europea. (2024). *Reglamento (UE) 2024/1689 por el que se establecen normas armonizadas en materia de inteligencia artificial.* Diario Oficial de la Unión Europea: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

Reglamento General de la Ley Orgánica de Protección de Datos Personales. (2023). Decreto Ejecutivo No. 904. Registro Oficial: https://www.cosede.gob.ec/wp-content/uploads/2023/12/REGLAMENTO-GENERAL-A-LA-LEY-ORG%C3%81NICA-DE-PROTECCION-DE-DATOS-PERSONALES_compressed-1.pdf

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

Ruiz, R., y del Moral, M. (2022). Aprendizaje activo y competencias socioemocionales en entornos digitales de educación superior. *UTE Teaching & Technology (Universitas Tarraconensis), 1*(1), 30-49. https://doi.org/10.17345/ute.2021.1.3210

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

Santana, A., Cruz, B., Castillo, M., y Acevedo, J. (2025). Impacto de proyectos integradores en el desarrollo de competencias transversales y evaluación entre pares. *CIENCIA ergo-sum, 32*(1). https://doi.org/10.30878/ces.v31n0a6

Scanlon, E., Taylor, Z., Raible, J., Bates, J., y Chini, J. (2021). Physics webpages create barriers to participation for people with disabilities: Five common web accessibility errors and possible solutions. *International Journal of STEM Education, 8*. https://doi.org/10.1186/s40594-021-00282-3

Sillat, L., Tammets, K., y Laanpere, M. (2021). Digital Competence Assessment Methods in Higher Education: A Systematic Literature Review. *Education Sciences, 11*(8). https://doi.org/10.3390/educsci11080402

Stewart, B., Miklas, E., Szcyrek, S., y Le, T. (2023). Barriers and beliefs: A comparative case study of how university educators understand the datafication of higher education systems. *International Journal of Educational Technology in Higher Education, 20*(1), 1-20. https://doi.org/10.1186/s41239-023-00402-9

Tan, M., y Maravilla, N. (2024). Shaping integrity: Why generative artificial intelligence does not have to undermine education. *Frontiers in Artificial Intelligence, 7*. 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*, 1-13. https://doi.org/10.3389/feduc.2025.1505370

Topuz, A., Saka, E., Fatsa, Ö., y Kurşun, E. (2022). Emerging trends of online assessment systems in the emergency remote teaching period. *Smart Learning Environments, 9*. https://doi.org/10.1186/s40561-022-00199-6

Tsai, Y.-S., Kovanović, V., y Gašević, D. (2021). Connecting the dots: An exploratory study on learning analytics adoption factors, experience, and priorities. *The Internet and Higher Education, 50*. https://doi.org/10.1016/j.iheduc.2021.100794

UNESCO. (2021). *Ética de la inteligencia artificial.* https://www.unesco.org/es/artificial-intelligence/recommendation-ethics

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

Van de Werfhorst, H., Kessenich, E., y Geven, S. (2022). The digital divide in online education: Inequality in digital readiness of students and schools. *Computers and Education Open, 3*. https://doi.org/10.1016/j.caeo.2022.100100

Vera, M. (2022). Retroalimentación como herramienta efectiva para el aprendizaje. *TZHOECOEN*, *14*(2), 21-33. https://doi.org/10.26495/tzh.v14i2.2281

Viberg, O., Kizilcec, R., Jivet, I., Martínez, A., Oh, A., Mutimukwe, C., . . . Scheffel, M. (2024). Cultural differences in students' privacy concerns in learning analytics across Germany, South Korea, Spain, Sweden, and the United States. *Computers in Human Behavior Reports, 14*. https://doi.org/10.1016/j.chbr.2024.100416

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., y Du, Z. (2024). Inteligencia artificial en educación: una revisión sistemática de la literatura. *Expert Systems with Applications, 252*. https://doi.org/10.1016/j.eswa.2024.124167

Wilson, J., Huang, Y., Palermo, C., Beard, G., y MacArthur, C. (2021). Automated feedback and automated scoring in the elementary grades: Usage, attitudes, and associations with writing outcomes in a districtwide implementation of MI Write. *International Journal of Artificial Intelligence in Education, 31*(2), 234-276. https://doi.org/10.1007/s40593-020-00236-w

Wu, Q., Wang, Y., Lu, L., Chen, Y., Long, H., y Wang, J. (2022). Virtual simulation in undergraduate medical education: A scoping review of recent practice. *Frontiers in Medicine*. https://doi.org/10.3389/fmed.2022.855403

Zhai, X., Krajcik, J., y Pellegrino, J. (2021). On the validity of machine learning-based next generation science assessments: A validity inferential network. *Journal of Science Education and Technology, 30*(2), 298-312. https://doi.org/10.1007/s10956-020-09879-9

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). https://doi.org/10.55982/openpraxis.16.3.656

LA TRANSFORMACIÓN DE LA EVALUACIÓN EN ENTORNOS DIGITALES



SOBRE EL LIBRO

La obra analiza cómo las tecnologías digitales han transformado la evaluación educativa, desde los modelos tradicionales hasta enfoques continuos, formativos y sustentados en evidencias diversas. Examina el uso de plataformas virtuales, rúbricas, portafolios, inteligencia artificial, analítica del aprendizaje y sistemas automatizados. También aborda la retroalimentación, la privacidad, los sesgos algorítmicos, la integridad académica y la brecha digital. El libro sostiene que la innovación tecnológica debe mantener coherencia pedagógica, equidad, accesibilidad y supervisión humana para contribuir realmente a mejorar el aprendizaje.

CIDPROS
Centro de innovación y desarrollo profesional

ISBN: 978-9907-9574-1-9