CONSEJO DE EVALUACIÓN EDUCATIVA DE LA UNAM

Guía de Uso de Inteligencia Artificial Generativa en Evaluación Educativa en licenciatura

Consejo de Evaluación Educativa • Red colaborativa de licenciatura • UNAM

SECRETARÍA GENERAL
Universidad Nacional Autónoma de México

CEIDE
Coordinación de Evaluación, Innovación y Desarrollo Educativos

Primera edición: abril de 2026
D.R. © 2026 UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
Guía de uso de inteligencia artificial generativa en la evaluación educativa en Licenciatura por Coordinación de Evaluación, Innovación y Desarrollo Educativos - Centro de Formación y Profesionalización Docente de la UNAM, se distribuye bajo una Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional.

La presente obra está bajo una licencia de CC BY-NC-SA 4.0 internacional
https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es_ES

Esta licencia permite:

• **Compartir** (copiar y redistribuir el material en cualquier medio o formato).
• **Adaptar** (remezclar, transformar y crear a partir del material).

Bajo los siguientes términos:

Reconocimiento. Debe reconocer adecuadamente la autoría, proporcionar un enlace a la licencia e indicar si se han realizado cambios. Puede hacerlo de cualquier manera razonable, pero no de una manera que sugiera que tiene el apoyo del licenciador o lo recibe por el uso que hace.

No comercial. Usted no puede hacer uso del material para una finalidad comercial.

Compartir igual. Si remezcla, transforma o crea a partir del material, deberá difundir sus contribuciones bajo la misma licencia que el original.

Hecho en México.

Cómo citar:

Universidad Nacional Autónoma de México (2026). *Guía de uso de inteligencia artificial generativa en la evaluación educativa en Licenciatura.* Consejo de Evaluación Educativa; Red Colaborativa de Licenciatura; Coordinación de Evaluación, Innovación y Desarrollo Educativos (CEIDE).

Índice

Introducción 5

1. Principios pedagógicos y éticos 6

1.1 Fundamentos de evaluación del/para/como aprendizaje 9

1.2 Principios Pedagógicos para la Evaluación con IAGen 11

1.3 Consideraciones éticas del uso de IAGen en la evaluación

2. Usos Potenciales de la IAGen para la evaluación educativa 15

2.1 Usos potenciales en el ciclo de la evaluación 15

2.2 Usos de IAGen por tipo de evaluación 19

2.3 Otras aplicaciones útiles de IAGen para la evaluación 22

3. Recomendaciones para el profesorado 24

3.1 Cómo diseñar evaluaciones aprovechando la IAGen 24

3.2 Cómo revisar y validar las respuestas generadas por IAGen 26

3.3 Prevención del uso indebido o acrítico por parte del estudiantado 26

3.4 Casos prácticos y sugerencias de implementación 27

4. Recomendaciones para el estudiantado 32

4.1 Cómo usar IAGen para estudiar, practicar y reflexionar 32

4.2 Cómo citar, verificar y no depender exclusivamente de la IAGen 33

4.3 Recomendaciones para citar el uso de IAGen 33

4.4 Actividades que promueven la metacognición con ayuda de la IA 36

5. Ejemplos prácticos 40

5.1 Plantillas de actividades evaluativas con IA

5.2 Rúbricas que integran el uso de IA 42

5.3 *Prompting* educativo: cómo formular instrucciones claras y efectivas para la IA 45

6. Advertencias y limitaciones 47

6.1 Alucinaciones, sesgos, problemas de privacidad y acceso desigual 47

6.2 Necesidad de supervisión y juicio docente 48

7. Referencias y recursos 50

Introducción

La revolución digital ha generado profundas transformaciones en múltiples ámbitos de la sociedad contemporánea y ha impactado significativamente diversas actividades humanas; el ámbito educativo no es la excepción. A lo largo de la historia, la educación ha estado marcada por la incorporación de tecnologías que modifican la forma en que se desarrolla el proceso de enseñanza-aprendizaje. El primer cambio paradigmático ocurrió con la llegada de las computadoras personales, seguido por la expansión de la conectividad a internet. Aunque cada una de estas innovaciones representó inicialmente un reto para las instituciones educativas, el profesorado y el estudiantado, en retrospectiva podemos reconocer las lecciones que dejaron: la integración de las computadoras fortaleció la capacidad de adaptación y tolerancia al cambio; Internet abrió el acceso universal a la información y favoreció la gestión del conocimiento. Posteriormente, la Web 2.0, mediante diversos programas y plataformas, impulsó el trabajo colaborativo y transformó el rol del usuario, quien pasó de ser un consumidor pasivo de información a un creador activo de contenidos.

Desde el constructivismo social se subraya la importancia de analizar cómo las herramientas tecnológicas median las interacciones sociales en el aula y cómo el profesorado las incorpora en sus prácticas pedagógicas para promover el aprendizaje colaborativo y la construcción del conocimiento.

En la actualidad, el nuevo reto al que se enfrenta la educación es la llegada e incorporación de la Inteligencia Artificial Generativa (IAGen). Su uso cada vez más extendido entre docentes y estudiantes permite afirmar que esta tecnología está transformando la educación superior y, en consecuencia, los procesos de evaluación.

La IA, concebida como un área de la ciencia que desarrolla sistemas, algoritmos y procesos capaces de simular ciertos aspectos de la inteligencia humana, ofrece herramientas que facilitan la optimización del aprendizaje y la enseñanza, así como su adaptación a las necesidades individuales del estudiantado. Es importante reconocer que su uso es cada vez más frecuente en diversos ámbitos de la educación, entre ellos la evaluación.

La presente guía tiene como objetivo ofrecer una orientación práctica, ética y pedagógicamente fundamentada para el uso de herramientas de Inteligencia Artificial Generativa (IAGen) en la evaluación del aprendizaje, para el aprendizaje y como aprendizaje en el nivel superior de la UNAM. Su propósito central es mostrar cómo esta tecnología puede convertirse en un aliado y no en una amenaza dentro de los procesos de evaluación educativa. A lo largo del documento se expondrá la manera en que la IA se integra en este entorno, destacando no solo sus beneficios, sino también sus desafíos inherentes y las limitaciones que puede representar.

1. Principios pedagógicos y éticos

El avance tecnológico se manifiesta en numerosos aspectos de la vida cotidiana, y el ámbito educativo no es la excepción. Los ajustes derivados de la integración de la IA afectan distintos elementos del quehacer académico, como los objetivos de aprendizaje, los métodos de enseñanza y, especialmente, los procesos de evaluación, que ofrecen al estudiantado nuevas formas de participar y monitorear su propio progreso. De acuerdo con la Encuesta Global de IA al Profesorado 2025 del Consejo de Educación Digital, el 50% del profesorado considera que la evaluación requiere cambios significativos, y uno de cada dos docentes sostiene que las tareas deben rediseñarse para ser más "resistentes a la IA" (DEC, 2025).

Aunque persiste cierta resistencia hacia la integración de estas tecnologías por las implicaciones que podrían tener en la integridad, la transparencia y la ética académica, es importante reconocer que la IA posee un considerable potencial para fortalecer procesos evaluativos, especialmente en lo relacionado con el desarrollo de habilidades y el logro de resultados de aprendizaje. Martínez et al. (2022) presentan una revisión detallada que muestra cómo la educación superior está experimentando una transformación significativa derivada de la integración de las TIC y la IA. Por su parte, Mora et al. (2024) señalan que la evaluación se beneficia ampliamente del uso de estas tecnologías: herramientas como los exámenes en línea y los sistemas de autoevaluación no solo reducen la carga de trabajo docente, sino que también aumentan la precisión y coherencia de las evaluaciones.

Asimismo, posibilitan una evaluación formativa continua que brinda al estudiantado retroalimentación inmediata, elemento fundamental para su desempeño académico, al ofrecer información sobre los cambios que se producen y aquellos que son necesarios para lograr un aprendizaje significativo tal como lo mencionan Avolio de Cols & Lacolutti, (2006). No obstante, la implementación de las TIC y la IA en la educación superior no está exenta de desafíos relacionados con la precisión, la equidad y la personalización. Estas tecnologías reabren el debate sobre los modelos de evaluación basados en pruebas estandarizadas y sumativas; aunque existen propuestas previas para fortalecer la evaluación formativa, es en la actualidad cuando adquieren una relevancia renovada. El reto para el profesorado consiste en rediseñar las evaluaciones existentes de modo que mantengan su pertinencia en la era de la IA e incorporar nuevas estrategias que utilicen esta tecnología como un recurso de apoyo para enriquecer el aprendizaje significativo del estudiantado.

Ante este panorama, Chan (2023) propone un modelo de política educativa de IAGen sostenible que busca atender estas preocupaciones en la educación superior. Su propuesta se estructura en tres dimensiones:

- **Pedagógica**: orientada al uso ético y responsable de la IAG para mejorar el impacto en los procesos de enseñanza, aprendizaje y evaluación.

- **De gobernanza**: centrada en aspectos relacionados con la privacidad, la seguridad y la responsabilidad, lo que implica definir pautas claras y regulaciones de uso que promuevan la conciencia y la corresponsabilidad de los distintos actores educativos.

- **Operativa**: enfocada en los elementos asociados con la infraestructura y la formación del profesorado, el estudiantado y el personal técnico, con el fin de que comprendan y utilicen estas tecnologías de manera ética y responsable.

La IAGen plantea el reto de promover en estudiantes, profesores e instituciones la capacidad de reflexión y autocrítica necesaria para un uso ético y culturalmente situado de esta tecnología y, de manera particular, en el ámbito de la evaluación del, para y como aprendizaje.

1.1 Fundamentos de evaluación del/para/como aprendizaje

De acuerdo con el *Glosario de Innovación Educativa* (Santos, González-Flores y Sánchez, 2022), el término evaluación se refiere a un proceso sistemático y metódico que emplea diversos medios e instrumentos para recopilar información verificable sobre el alcance, los logros y los resultados de distintos procesos y acciones, con el propósito de juzgar su mérito o valor. Esta información puede ser general o específica, cuantitativa o cualitativa, y debe ser relevante y pertinente para sustentar decisiones, identificar dificultades o limitaciones, mejorar las acciones y resultados, y orientar alternativas para su ajuste.

La evaluación educativa es necesaria para obtener información centrada en el proceso educativo y en las posibilidades de transformación y mejora de las acciones formativas. Mediante la evaluación es posible valorar tanto el proceso de aprendizaje como el de enseñanza, así como sus distintos componentes y niveles: los planes y programas de estudio, las estrategias didácticas, los materiales educativos, los sistemas, políticas y acciones institucionales, además del equipamiento e infraestructura de los centros educativos, entre otros.

Por tal motivo, como señala Moreno (2016), la evaluación constituye una piedra angular que condiciona tanto la enseñanza como el aprendizaje. La evaluación del aprendizaje se centra en investigar, documentar y potenciar el proceso formativo, atendiendo a sus dificultades y oportunidades. El desarrollo de la evaluación en la educación superior se alinea con los enfoques centrados en el aprendizaje. En este marco, Sánchez (2022) distingue entre la evaluación del aprendizaje, la evaluación para el aprendizaje y la evaluación como aprendizaje:

- **Evaluación del aprendizaje**: documenta si el aprendizaje se ha logrado y el nivel alcanzado, poniendo énfasis en los aspectos cuantitativos y numéricos asociados con calificaciones o grados. Se centra en evaluar resultados, desempeños, productos y actividades al término de un periodo formativo. Tal como señala Rincón (2022, p. 43), "busca demostrar el grado en que los estudiantes han alcanzado los objetivos educativos". Entre los instrumentos utilizados destacan los exámenes objetivos y los reportes periódicos.

La evaluación del aprendizaje tiene un carácter sumativo, lo que representa una vulnerabilidad, pues el estudiantado pueden recurrir al uso de la IAGen para resolver pruebas estandarizadas o

tareas que implican la elaboración de productos académicos, como ensayos o análisis en los que habitualmente se evalúa únicamente el resultado final. Esta situación constituye una de las principales preocupaciones del profesorado.

El profesorado debe procurar regular el uso que hace el estudiantado de la IAGen para realizar sus actividades académicas, para que esta se contemple como apoyo y no como sustituto del trabajo propio. Para ello, pueden emplearse estrategias como exámenes orales o actividades en las que sea posible evaluar el proceso —por ejemplo, ensayos o tareas escritas desarrolladas bajo supervisión docente—, de modo que se pueda dar seguimiento a la construcción del aprendizaje.

▶• **Evaluación para el aprendizaje**: identifica las áreas de oportunidad y los logros del estudiantado con el fin de mejorar su desempeño. Este tipo de evaluación "se centra en proporcionar información útil sobre cómo los estudiantes aplican lo que saben; se alinea con la noción de que la evaluación puede ser una herramienta para guiar la enseñanza y mejorar el proceso de aprendizaje" (Rincón, 2022, p. 43). Por tanto, ocurre a lo largo de todo el proceso formativo y se caracteriza por su naturaleza longitudinal y por el flujo constante de información, retroalimentación y diálogo entre docentes y estudiantes. Entre los instrumentos más utilizados se encuentran las retroalimentaciones y las rúbricas.

Diversos informes coinciden en los beneficios que la IAGen puede aportar a la evaluación para el aprendizaje. Sus posibilidades como guía o tutor personalizado, capaz de orientar y ofrecer retroalimentación inmediata en ejercicios y prácticas académicas, brindan al estudiantado oportunidades valiosas de reflexión y autocrítica sobre su proceso de aprendizaje. No obstante, estos informes también destacan la importancia de la figura docente en el acompañamiento y la supervisión, a fin de asegurar que las actividades se alineen con los objetivos de aprendizaje y se desarrollen bajo criterios pedagógicos claros.

▶• **Evaluación como aprendizaje**: se refiere a un proceso "donde el estudiante desempeña un papel activo para relacionar la información con su conocimiento previo y utilizarla para nuevos aprendizajes" (Rincón, 2022, p. 43). Este enfoque desplaza la evaluación hacia la autoevaluación, permitiendo al estudiantado una mayor autorregulación y control intrínseco sobre su propio aprendizaje. La información recopilada orienta al alumnado en el desarrollo del "aprender a aprender", favoreciendo la metacognición y la autonomía. Este tipo de evaluación "refuerza la idea de que la evaluación puede ser una parte integral del proceso de aprendizaje y no simplemente una medición externa" (Rincón, 2022, p. 43) y requiere, por ello, de una práctica reflexiva continua.

Así, mientras que las evaluaciones del aprendizaje proveen evidencia del rendimiento del alumnado y determinan el nivel alcanzado —a través de informes públicos u otros reportes—, las evaluaciones para el aprendizaje ayudan al estudiantado a aprender más y mejor (Moreno, 2016), proporcionando los "cómos", es decir, las estrategias para lograrlo, a través de las evaluaciones como aprendizaje.

La IAGen nos invita a reflexionar y analizar críticamente sus posibilidades; por ello, es indispensable formarnos en su uso ético y desarrollar capacidades metacognitivas y de autorregulación que eviten la dependencia tecnológica y promuevan la autonomía cognitiva. Según Chao Rebolledo (2025, p. 96), el profesorado del futuro "más

que enseñar contenidos, deberá acompañar procesos, facilitar preguntas, habilitar espacios para la metacognición y ayudar a los estudiantes a navegar un mundo de conocimientos en evolución constante". En consecuencia, la incorporación de la IA en la evaluación del, para y como aprendizaje requiere establecer un marco normativo que contemple principios que aseguren una mediación pedagógica adecuada para su implementación.

1.2 Principios Pedagógicos para la Evaluación con IAGen

Para emprender cualquier proceso de evaluación educativa es necesario partir de los principios generales que orientan su diseño e implementación. Sánchez (2022, p. 18), retomando a Miller (2012), identifica los siguientes:

1. Especificar con claridad el objeto de evaluación.
2. Concebir la evaluación no como un fin en sí mismo, sino como un medio para alcanzar una finalidad.
3. Elegir métodos pertinentes de evaluación según su relevancia y las características del objeto evaluado.
4. Diversificar procedimientos e instrumentos para lograr una evaluación más objetiva, útil y efectiva.
5. Conocer las bases técnicas de la evaluación y sus instrumentos, así como las fortalezas y limitaciones de cada método.

Dado que la evaluación educativa es un proceso complejo que requiere diversas estrategias e instrumentos para cumplir con sus principios, la IAGen constituye un conjunto de herramientas con gran potencial para fortalecer la evaluación del, para y como aprendizaje. No obstante, es necesario contar con pautas pedagógicas que clarifiquen los alcances y las limitaciones de su uso y que, como señala Chan (2023), contribuyan a mejorar su impacto en los procesos de enseñanza, aprendizaje y evaluación. A continuación, se presentan los principios que orientan el uso pedagógico de la IA en la evaluación:

▶• *Centralidad del aprendizaje*

El aprendizaje debe ser el eje rector de toda innovación educativa, así se señala en Universidad Internacional de La Rioja (s.f.). En este sentido, la IAGen debe utilizarse como un recurso de apoyo al desarrollo integral del estudiantado, ampliando sus oportunidades de reflexión, autoevaluación y construcción de conocimiento.

Un ejemplo de ello es el uso de *chatbots* especializados que, en lugar de proporcionar respuestas definitivas, formulen preguntas que orienten al alumnado hacia el descubrimiento autónomo. En áreas como medicina, derecho o ingeniería, esto puede traducirse en simulaciones de casos complejos que requieran la aplicación de conocimientos en contextos auténticos. (Guzmán Martínez, M. I., & Marín Uribe, R. E. , 2020).

▶• *Personalización del aprendizaje*

La IAGen posibilita la personalización del aprendizaje mediante la creación de entornos evaluativos interactivos que atienden las particularidades del estudiantado en sistemas de tutoría virtual. Esto permite aprovechar su diversidad, identificar necesidades específicas y anticipar su rendimiento académico. La personalización puede sustentarse en un conocimiento detallado de las fortalezas y áreas de oportunidad de cada estudiante a partir de diagnósticos, preguntas adaptadas, intereses, estilos de estudio y evaluaciones continuas. Estos sistemas pueden ser adaptativos, ajustándose al ritmo y estilo de aprendizaje del alumnado. La información generada es valiosa tanto para docentes como para estudiantes, pues orienta el acompañamiento educativo. Algunos ejemplos de esta personalización

en la evaluación incluyen:

- evaluar conocimientos y habilidades y ajustar ejercicios o actividades a su nivel;
- recomendar lecturas o recursos alineados con sus intereses;
- generar informes individuales de desempeño;
- identificar patrones y tendencias para apoyar decisiones pedagógicas;
- ofrecer retroalimentación inmediata mediante tutorías virtuales.

▶• *Realimentación formativa y personalizada*

Una de las principales ventajas de la IAG es su capacidad para ofrecer retroalimentación inmediata y personalizada. Así, un ensayo académico puede recibir comentarios automáticos sobre su estructura, claridad y coherencia. (Bolaño-García & Duarte-Acosta, 2024)

No obstante, esta retroalimentación debe ser supervisada y complementada por el docente, a fin de evitar que el estudiantado reciba únicamente correcciones técnicas y asegurar que también cuente con orientación crítica y formativa. Tal como señala Jara (2020, citado en Montoya et al., 2024, p. 12): "...es importante mantener un equilibrio entre la personalización y la interacción humana, ya que la educación implica el desarrollo de habilidades sociales y emocionales que requieren el contacto directo con profesores y compañeros."

▶• *Objetividad*

Es uno de los atributos fundamentales de la evaluación educativa, pues implica minimizar la influencia del evaluador y centrarse en criterios claros y transparentes que garanticen una valoración consistente e imparcial.

Como señala Llanos (2025, p. 80): "En el contexto de las evaluaciones tradicionales, los sesgos inconscientes de los evaluadores, las variaciones en la severidad de las calificaciones y otros factores humanos pueden influir significativamente en la consistencia de los resultados, lo que a menudo conduce a injusticias en la evaluación del desempeño estudiantil." Aunque algunos autores denominan esta objetividad como un *desiderátum* López, A.T., & Pedraza, N. Y. (2017) o "evaluación reproductiva" (Calatayud, 1999), su propósito no es estandarizar el conocimiento, sino eliminar sesgos y evitar prejuicios o interpretaciones subjetivas, al basarse en criterios medibles y establecidos. Este atributo resulta especialmente valioso en la evaluación apoyada en IA. Mora Naranjo et al. (2023) sostienen que la IA, al aplicar criterios consistentes y basados en algoritmos, asegura evaluaciones más justas y equitativas, especialmente en contextos de gran escala.

De este modo, el enfoque evaluativo basado en datos y algoritmos reduce la variabilidad propia de la subjetividad humana y promueve mayores niveles de confiabilidad, lo que contribuye a la equidad y fortalece la percepción de legitimidad en los procesos evaluativos.

Como afirma Llanos (2025, p. 80): "Cuando los estudiantes y profesores saben que están utilizando un sistema basado en algoritmos consistentes y transparentes, es más probable que confíen en los resultados y en el valor de la evaluación. Esta confianza es crucial para fomentar un entorno de aprendizaje en el cual los estudiantes se sientan motivados a participar y esforzarse, sabiendo que su rendimiento será evaluado de manera justa."

▶• *Validez y confiabilidad*

Toda evaluación, ya sea tradicional o apoyada en tecnología, debe cumplir con los criterios de validez —es decir, medir aquello que pretende medir— y confiabilidad —ofrecer resultados consistentes. La IAG puede

generar reactivos con gran rapidez, pero estos deben ser revisados y validados por especialistas en contenido para garantizar su calidad. Vivas Urías, M. D., & Ruiz Rosillo, M. A. (2025)

De lo contrario, se corre el riesgo de evaluar competencias superficiales o de introducir sesgos en la calificación. En la UNAM, los cuerpos colegiados y las academias de profesorado podrían fungir como instancias de validación de los materiales generados mediante IA.

▶• *Diversificación e innovación*

La IAG tiene el potencial de ampliar el abanico de instrumentos de evaluación, tales como:

- rúbricas dinámicas que ajustan la retroalimentación según los niveles de desempeño del estudiantado;
- simulaciones interactivas, especialmente útiles en áreas como medicina, odontología o arquitectura;
- generación de escenarios problemáticos que fomentan el pensamiento crítico y la resolución de problemas.

No obstante, la innovación no debe confundirse con uniformidad. La evaluación debe preservar su carácter multimodal, integrando dimensiones cuantitativas y cualitativas, así como modalidades digitales y presenciales.

▶• *Metacognición y pensamiento crítico*

El uso de la IAG no debe fomentar una actitud pasiva en el estudiantado; por el contrario, debe constituir una oportunidad para fortalecer la metacognición, entendida como la capacidad de reflexionar sobre el propio aprendizaje. Esto implica enseñar a los estudiantes a evaluar críticamente las respuestas generadas por la IA, distinguiendo entre sugerencias útiles y posibles errores. De este modo, la tecnología deja de ser un oráculo y se convierte en un detonador del pensamiento crítico. (Digital Education Council, 2025, p. 29) Como señala Rincón (2022, p. 41): “Aunque la IA y la tecnología ofrecen numerosas ventajas, como la personalización del aprendizaje, la accesibilidad y la mejora de la eficiencia al adaptar el aprendizaje a necesidades personales, es fundamental recordar que estos avances son herramientas que deben estar al servicio de los objetivos pedagógicos. Por tanto, la enseñanza y la evaluación deben permanecer en el centro de atención. En este punto, el papel del profesorado cobra vital importancia para encontrar el equilibrio entre el uso de la tecnología y la intervención humana. Como se ha señalado antes, en la evaluación la IA no debe ser un fin en sí mismo, sino una aliada que promueva la mejora constante de la educación y el logro de aprendizajes significativos.”

1.3 Consideraciones éticas del uso de IAGen en la evaluación

Como hemos visto en la sección anterior, la IAGen comienza a posicionarse como un agente transformador en la evaluación educativa, su integración sistemática permite elevar la validez, mejorar la eficiencia, reducir sesgos, fortalecer la retroalimentación y ampliar las posibilidades analíticas. Sin embargo, al mismo tiempo, introduce desafíos complejos en torno a transparencia, auditabilidad, equidad algorítmica y responsabilidad institucional. Es de suma importancia reconocer que el uso y la naturaleza de las IAGen conllevan dilemas éticos, ambientales, laborales y pedagógicos (Ferrarelli, 2024). Tal como lo señalan Shute & Ventura (2022), la IA implica la integración de procesos evaluativos invisibles (*stealth assessment),* lo que exige políticas sólidas de gobernanza, vigilancia ética y participación docente. Esto significa que la IA debe ser adoptada desde un enfoque crítico, informado y éticamente responsable, preservando la centralidad del juicio pedagógico y la rendición de cuentas.

Por ello, a continuación, se presentan algunas consideraciones éticas que contribuyen a orientar su adecuada regulación.

La UNESCO (2025) propone que el enfoque de la IA debe contemplar diez principios fundamentales:

1. proporcionalidad e inocuidad;
2. seguridad y protección;
3. derecho a la intimidad y protección de datos;
4. gobernanza y colaboración adaptativas y de múltiples partes interesadas;
5. responsabilidad y rendición de cuentas;
6. transparencia y explicabilidad;
7. supervisión y decisión humanas;
8. sostenibilidad;
9. sensibilización y educación;
10. equidad y no discriminación.

Además del establecimiento de un marco normativo para la utilización de la IAGen en la evaluación del, para y como aprendizaje, es fundamental integrar un marco ético que coloque en el centro el aprendizaje de los estudiantes, y no únicamente el uso de la tecnología para obtener una calificación aprobatoria.

Este marco ético debe partir del principio de que el uso de la IAGen debe ser situado, colectivo y colaborativo, con la participación de toda la comunidad educativa —instituciones, profesorado y estudiantado—, de manera que se salvaguarden los derechos y la seguridad de todas las partes, incluyendo la protección de datos y la privacidad. El manejo masivo de información sensible exige implementar medidas de seguridad, soberanía digital y consentimiento informado (UNESCO, 2025).

En todo momento, las instituciones y el profesorado deben garantizar que el uso de la IAGen en la evaluación educativa funcione como una herramienta de apoyo para la reflexión, el pensamiento crítico y la construcción de aprendizajes significativos en los estudiantes.

▶• *Equidad y justicia*

En el ámbito ético, uno de los principios centrales es la equidad en el acceso a estas tecnologías. La educación pública en México enfrenta desigualdades en el acceso a dispositivos y conectividad, por ello, la UNAM ha señalado la importancia de reconocer las brechas digitales que persisten en la comunidad académica, por lo cual, el uso de IAGen debe acompañarse de estrategias que reduzcan desigualdades y eviten exclusiones (UNAM, 2025), su implementación en la evaluación requiere políticas institucionales que garanticen condiciones de equidad. Algunas de estas implicaciones incluyen:

- asegurar que las plataformas utilizadas sean accesibles desde diversos dispositivos;
- crear materiales de orientación y apoyo para el uso adecuado de las IAGen por parte de todos los miembros de la comunidad educativa —profesorado, estudiantado y autoridades— (Wargo y Anderson, 2024);
- evitar que la dependencia de servicios comerciales internacionales excluya a estudiantes con menos recursos;
- promover una evaluación crítica de las herramientas de IAGen antes de su utilización, considerando los posibles sesgos y estereotipos sociales y culturales que pueden reproducir (Wargo y Anderson, 2024);
- diseñar programas de apoyo tecnológico para los sectores más vulnerables.

▶• *Transparencia y explicabilidad*

Regular la aplicación de la IAGen en la evaluación educativa permitirá que profesorado, instituciones y estudiantado puedan auditar y comprobar la validez de los instrumentos de evaluación, así como de las respuestas generadas, y definir con claridad su uso en los procesos de enseñanza y aprendizaje.

El estudiantado debe comprender cómo operan los algoritmos que intervienen en su evaluación, pues la opacidad tecnológica puede minar la confianza en el proceso académico. Por ello, es necesario que la UNAM establezca criterios claros de transparencia que expliquen qué parámetros se consideran en las evaluaciones generadas por IA, cómo se ponderan los resultados y qué función conserva el docente en la decisión final. (Porter, 2025)

Asimismo, es fundamental desarrollar materiales de difusión que informen a la comunidad educativa cómo el uso de IAG puede afectar sus experiencias de aprendizaje, promoviendo una postura crítica (Wargo y Anderson, 2025).

▶• *Integridad académica*

La rápida expansión de la inteligencia artificial generativa ha generado una serie de cuestionamientos respecto a su uso. Aunque estudios como el de Sung et al. (2023) señalan que herramientas como ChatGPT tienen un gran potencial para transformar la práctica docente y el aprendizaje, una de las principales inquietudes en el ámbito de la evaluación es su impacto en la integridad académica.

Se ha identificado que un porcentaje significativo de estudiantes ha utilizado ChatGPT en tareas de evaluación, lo que ha llevado a diversas universidades a restringir o prohibir su uso. Como afirman Gallent-Torres et al. (2023, p. 6): "...esto plantea interrogantes sobre cómo garantizar la equidad y la autenticidad en las evaluaciones, y cómo evitar que el fraude académico comprometa el sistema educativo." En este punto resulta pertinente retomar el modelo propuesto por Chan (2023).

▶• *Autenticidad académica*

El plagio asistido por IA constituye una de las preocupaciones centrales en el ámbito educativo. La respuesta institucional no debe ser únicamente punitiva, sino también pedagógica. Algunas acciones recomendadas son:

- rediseñar tareas que fomenten la originalidad y la reflexión crítica;
- valorar los procesos —borradores, esquemas, auto explicaciones— además de los productos finales;
- fomentar la cultura de honestidad académica desde los primeros semestres;
- enfatizar que las IAGen pueden apoyar, pero no sustituir las intervenciones humanas ni las interacciones entre docentes y estudiantes, fundamentales para el aprendizaje (Wargo y Anderson, 2024);
- promover la declaración explícita y justificada de los propósitos con los que se emplean las IAGen en las actividades educativas, impulsando la participación activa del estudiantado (Wargo y Anderson, 2024);
- incentivar procesos de revisión entre pares en actividades asistidas con IA, para preservar la autenticidad de la intervención humana y compartir estrategias de uso responsable de las IAGen (Wargo y Anderson, 2024).
- Implementar planes de evaluación alternativos que recurran a otras formas de evaluar los resultados de aprendizaje, "...tales como discusiones, debates, proyectos, portafolios o prácticas que,

además de permitir la verificación de la autenticidad de la producción intelectual del estudiante frente a la posibilidad de utilizar chatbots, permiten al docente identificar directamente su desempeño". (Chiappe et.al., 2025: p.128).

▶• *Privacidad y protección de datos*

El uso de plataformas de IAGen implica el manejo de datos personales y académicos. La UNAM debe garantizar que la información de estudiantes y docentes no sea utilizada con fines comerciales ni expuesta a riesgos de seguridad digital. Para ello se requieren, entre otras acciones:

- políticas claras de uso de datos;
- protocolos de anonimización;
- preferencia por herramientas desarrolladas o controladas institucionalmente;
- desarrollo de guías que expliquen cómo la propiedad intelectual y la información personal pueden utilizarse para entrenar herramientas de IAGen, así como el impacto que ello puede tener en los resultados futuros (Wargo y Anderson, 2024).

▶• *Sustentabilidad*

El entrenamiento e implementación de sistemas basados en IA e IAGen requieren un consumo energético elevado, lo que contribuye al incremento de la huella de carbono. Asimismo, su uso implica una base material asociada a la producción y desecho del hardware necesario para soportar estas tecnologías, el cual, al finalizar su vida útil, suele ser difícil de reciclar o reutilizar y puede resultar ambientalmente tóxico (Ferrarelli, 2024). El uso de IAGen obliga, por tanto, a considerar los riesgos ambientales que pueden aumentar la vulnerabilidad de ciertas poblaciones. Para atender estos dilemas, se proponen las siguientes acciones:

- promover el uso de IAGen con menor impacto ambiental o con infraestructuras que prioricen la sustentabilidad, reduciendo el consumo de agua o las emisiones de carbono (Wargo y Anderson, 2024);
- favorecer debates dentro de la comunidad educativa para comprender las consecuencias ambientales del uso indiscriminado de las IAGen, e integrar estas reflexiones en planes y programas de estudio pertinentes (por ejemplo, en áreas dedicadas al diseño tecnológico) (Wargo y Anderson, 2024);
- ofrecer talleres que incorporen el eje de la sustentabilidad en el diseño y uso de las IAGen, promoviendo alternativas para reducir sus impactos ambientales (Wargo y Anderson, 2024).

2. Usos Potenciales de la IAGen para la evaluación educativa

La integración de la Inteligencia Artificial Generativa (IAGen) en los procesos de evaluación educativa constituye un avance tecnológico que permite optimizar, escalar y sofisticar la obtención, el análisis y la interpretación de evidencias de aprendizaje. Su aporte se manifiesta en diversas fases del ciclo evaluativo, desde la formulación de resultados de aprendizaje y la planeación didáctica hasta la generación de instrumentos, la administración de pruebas y la retroalimentación automatizada. Asimismo, la IAGen fortalece distintos tipos de evaluación —del, para y como aprendizaje— mediante procedimientos de análisis lingüístico, clasificación, detección de patrones, generación de reactivos y simulación de escenarios complejos.

Este apartado examina los usos potenciales de la IAGen con un enfoque pedagógico y funcional, incluyendo aplicaciones emergentes como la analítica del aprendizaje, la evaluación automatizada de respuestas abiertas y los sistemas de simulación con retroalimentación inmediata. En conjunto, estas capacidades tecnológicas muestran que la IAGen no solo agiliza procesos, sino que constituye un recurso valioso para diseñar procesos de evaluación más pertinentes, equitativos, formativos y alineados con las demandas actuales del contexto educativo, asimismo introduce nuevas posibilidades metodológicas para el diseño de evaluaciones más precisas, adaptativas y basadas en datos.

2.1 Usos potenciales en el ciclo de la evaluación

La incorporación de la IAGen en los procesos de evaluación del aprendizaje ofrece un amplio conjunto de posibilidades que transforman desde la definición de los aprendizajes esperados hasta las estrategias de retroalimentación empleadas para valorar el progreso estudiantil. Para poder visualizar mejor dichas posibilidades, su impacto y poder realizar aplicaciones prácticas de la IAGen en el entorno evaluativo, las enfocaremos dentro del Ciclo de Evaluación Educativa.

Este ciclo es un proceso continuo, sistemático y retroalimentado que permite valorar el aprendizaje, orientar la enseñanza y mejorar los resultados académicos. No es un acto aislado, sino un circuito permanente de decisiones pedagógicas. Lo componen las siguientes etapas o fases:

Planificación de la evaluación. En esta fase se definen:

- Competencias, resultados de aprendizaje o aprendizajes clave que se evaluarán.
- Criterios e indicadores que permiten observar el nivel de desempeño.
- Instrumentos y métodos de evaluación (rúbricas, listas de cotejo, ECOE, exámenes escritos, portafolios, OSCE, observación directa).
- Momento en que se aplicará cada tipo de evaluación (diagnóstica, formativa, sumativa)

Recolección de información. En esta fase se aplica la evaluación mediante los instrumentos que se hayan elegido y definido:

- Pruebas escritas, exámenes objetivos, *quizzes*
- Actividades prácticas
- Observación estructurada
- Ensayos, proyectos o portafolios
- Exámenes clínicos objetivos (ECOE/OSCE)
- Autoevaluación y coevaluación

Valoración e interpretación. En esta fase se:

- Analizan los datos recolectados.
- Comparan con los criterios e indicadores previamente definidos.
- Establece el nivel de desempeño del estudiante.
- Detectan fortalezas, áreas de mejora y brechas formativas.

Retroalimentación. Fase central del ciclo y la de mayor impacto en el aprendizaje, debe ser:

- Clara, específica y oportuna
- Orientada a la mejora
- Basada en evidencias
- Enfocada en el desempeño, no en la persona
- **Idealmente dialogada** (*feedback + feedforward)*

Toma de decisiones. Con la información obtenida se pueden determinar:

- Ajustes en la enseñanza
- Nuevas estrategias didácticas
- Diferenciación del aprendizaje
- Necesidad de tutoría o apoyo
- Revisión de instrumentos de evaluación
- Planes de mejora para la asignatura o curso

Mejora del proceso (metaevaluación). Implica evaluar la calidad del propio proceso de evaluación:

- ¿Son válidos y confiables los instrumentos?
- ¿Los criterios fueron claros?
- ¿La evaluación promovió aprendizaje profundo?
- ¿Hubo coherencia entre enseñanza, evaluación y competencias?
- ¿Se redujeron sesgos?
- ¿Se favoreció la justicia evaluativa?

El ciclo es continuo ya que una vez cerrada una ronda de evaluación, el proceso vuelve a iniciar. Por eso la evaluación no es un evento, sino un sistema dinámico de mejora, y en cada una de las fases o etapas que conforman su ciclo, las IAGen pueden aprovecharse para potenciar la calidad, pertinencia y equidad del proceso evaluativo. A continuación revisaremos sus usos potenciales de acuerdo con los propósitos de cada fase o etapa del ciclo, de acuerdo con la propuesta del ciclo en *The Next Era of Assessment* (DEC, 2025):

Tabla 1. Usos Potenciales de la IA en diferentes fases del ciclo de Evaluación Educativa (DEC, 2025)

Propuesta del Ciclo en The Next Era of Assessment (Digital Education Council)	Usos Potenciales de la IAGen
Etapa 1. **Establecimiento de resultados de aprendizaje.** En esta primera etapa se definen los resultados de aprendizaje, es decir, los objetivos concretos que el alumnado deberá alcanzar al finalizar un curso, unidad o actividad formativa.	La IAGen puede apoyar esta etapa mediante: ▪ el análisis de marcos de competencias nacionales e internacionales, ▪ la identificación de tendencias educativas emergentes, ▪ la revisión de expectativas institucionales y requerimientos del sector profesional, ▪ la extracción de habilidades y actitudes más demandadas en el mundo laboral. Con base en ello, es posible actualizar y formular resultados de aprendizaje más pertinentes, actuales y alineados con las necesidades formativas del estudiantado y de su futura inserción profesional.
Etapa 2. **Planeación didáctica.** En esta etapa se organiza el contenido y se diseña la estructura del curso, así como los métodos para evaluar los aprendizajes.	La IAGen puede contribuir mediante: ▪ la generación de mapas curriculares, ▪ la dosificación y secuenciación de contenidos, ▪ el codiseño de actividades alineadas con los resultados esperados, ▪ la creación de itinerarios personalizados según el perfil del estudiantado, ▪ la propuesta de estrategias e instrumentos de evaluación coherentes con los principios pedagógicos. También es posible utilizar la IAGen para definir criterios éticos de uso de estas herramientas por parte del alumnado. Por ejemplo: ▪ planear actividades iniciales donde se restrinja el uso de generadores de texto para fortalecer habilidades de escritura independiente; ▪ permitir posteriormente su uso solo para revisión estilística; ▪ evaluar la capacidad del estudiantado para distinguir entre aportaciones propias y sugerencias de la IA.
Etapa 3. **Desarrollo de estrategias de evaluación.** En esta fase se establecen los criterios y los instrumentos que permitirán evidenciar el nivel de avance del alumnado.	La IAGen puede apoyar en: ▪ la creación de tareas, exámenes, rúbricas y casos prácticos alineados con los objetivos del curso, ▪ la generación de simulaciones auténticas y guías de reflexión, ▪ la elaboración de reactivos o variaciones de una misma prueba para aumentar la equidad y reducir la copia, ▪ la definición de criterios de evaluación claros y consistentes. Por ejemplo, la IAGen puede generar múltiples versiones de un examen con variaciones numéricas, manteniendo el nivel de dificultad y los objetivos del instrumento.

Etapa 4. **Ejecución de la evaluación.** En esta etapa se aplican los instrumentos previamente diseñados.	La IAGen puede apoyar mediante: ▪ el registro y análisis inmediato de respuestas, ▪ la detección de patrones inusuales o conductas sospechosas en exámenes en línea para fortalecer la integridad académica, ▪ la automatización de procesos administrativos y rutinarios, ▪ la facilitación de tareas de monitoreo en ambientes virtuales. Por ejemplo: ▪ un sistema de supervisión con IAGen puede identificar cambios constantes de pantalla o respuestas idénticas en lapsos atípicos en un examen remoto; ▪ en evaluaciones de proyectos, la IAGen puede revisar código básico y liberar tiempo del profesorado para valorar la creatividad y la calidad conceptual del trabajo.
Etapa 5. **Retroalimentación y revisión.** La etapa final del ciclo incluye la revisión de las evidencias de aprendizaje y la retroalimentación.	La IAGen ofrece ventajas como: ▪ la retroalimentación inmediata, detallada y adaptada a cada estudiante, ▪ la revisión sistemática y masiva para garantizar la claridad de instrucciones e instrumentos, ▪ la generación de reportes de desempeño individuales y grupales, ▪ las sugerencias para mejorar actividades e instrumentos evaluativos. Por ejemplo, una herramienta de IAGen puede revisar un ensayo identificando cohesión, gramática y vocabulario; el profesorado puede complementar esa retroalimentación con un análisis profundo del argumento y la originalidad.

En conjunto, la IAGen puede emplearse a lo largo de todo el ciclo de evaluación, **no** solo para optimizar aspectos logísticos, sino también para impulsar transformaciones pedagógicas significativas. Gracias a estas herramientas, el profesorado puede diseñar experiencias más equitativas, contextualizadas y orientadas al desarrollo de competencias necesarias en un entorno donde la colaboración con sistemas inteligentes será cada vez más habitual.

De esta manera, la IAGen, en su dimensión evaluativa, puede entenderse **como una extensión de la medición educativa**, como un:

- Sistema analítico (capaz de inferir patrones y tendencias).
- Agente pedagógico (que retroalimenta y orienta).
- Mecanismo de garantía de calidad (que regula, detecta anomalías y corrige).
- Entorno adaptativo (que personaliza tanto la presentación como la dificultad del examen).

Estas funciones se alinean con los modelos contemporáneos de evaluación, por ejemplo, la evaluación formativa integrada, la medición adaptativa y la psicometría computacional, lo que además da lugar a una nueva ***ecología de la evaluación*** (Gierl & Lai, 2022).

2.2 Usos de IAGen por tipo de evaluación

Son múltiples las herramientas de inteligencia artificial que pueden apoyar y fortalecer los distintos tipos de evaluación —del, para y como aprendizaje— en diversos niveles educativos y áreas disciplinarias.

Como se señaló en el apartado previo, una misma herramienta de IAGen puede integrarse en distintas fases del ciclo evaluativo para diseñar planeaciones, construir rúbricas, proponer actividades o crear proyectos, entre otras tareas docentes.

La IAGen transforma la evaluación cada vez que se utiliza para asignar calificaciones automáticas con retroalimentación inmediata o para detectar plagio en trabajos académicos (ANUIES, 2025). De acuerdo con la UNESCO (2025), la IAGen tiene el potencial de reconfigurar la evaluación al favorecer un desplazamiento de las pruebas sumativas hacia la valoración de procesos. También promueve la retroalimentación constante y la moderación humana mediante el uso de rúbricas para el seguimiento y evaluación formativa.

▶• *Evaluación del aprendizaje*

En este tipo de evaluación, las IAGen facilitan la evaluación automatizada del trabajo del estudiantado, proporcionando retroalimentación instantánea y precisa a través de instrumentos clásicos de evaluación sumativa, como los exámenes y los ensayos.

Los exámenes, como señala Moreno (2016), tienen una larga tradición en el ámbito escolar como estrategia para demostrar el conocimiento y la comprensión logrados. Estos pueden consistir en preguntas abiertas formuladas por el profesorado y respondidas por el alumnado de manera oral o escrita; en la resolución de problemas; en la defensa de una postura o en la discusión de un tema controvertido. También pueden adoptar la forma de pruebas objetivas con múltiples preguntas o reactivos (ítems), que se responden con una frase breve o incluso con una palabra, letra o número.

Para que tengan valor educativo, es fundamental que los exámenes o pruebas no se limiten a medir la capacidad de reproducción literal de la información, sino que también permitan al profesorado mejorar su enseñanza, orientando o reorientando al alumnado en su proceso de aprendizaje (Moreno, 2016).

Entre las herramientas de evaluación automática de exámenes y trabajos se encuentran las siguientes (Mujica-Sequera, 2024):

GRADESCOPE https://www.gradescope.com/

Permite al profesorado cargar exámenes, analizar respuestas y calificarlas automáticamente, brindando retroalimentación inmediata y correcciones automatizadas.

CROWDMARK https://crowdmark.com/

Facilita la calificación de grandes volúmenes de respuestas mediante el análisis del contenido de exámenes y trabajos impresos o en línea, lo que agiliza el proceso evaluativo.

▶• *Evaluación para el aprendizaje*

Algunos de los instrumentos habituales de evaluación para el aprendizaje pueden clasificarse de la siguiente manera (Castejón et al., 2009, citado en Hamodi, López-Pastor y López-Pastor, 2015):

- Con **predominio de la expresión escrita**: examen, trabajo escrito, ensayo, póster, informe, recensión, proyecto, cuaderno de campo, fichas prácticas, diario, carpeta o dossier, portafolio y memoria.
- Con **predominio de la expresión oral**: exposición, debate, entrevista, entrevista en grupo, grupo

de discusión, mesa redonda, panel de expertos, ponencia y comunicación.

- Con **predominio de la expresión práctica**: representación, demostración o actuación, simulación, desarrollo de proyectos con componente práctico, búsquedas y prácticas supervisadas.

En cualquiera de estos instrumentos es fundamental que el profesorado registre información que le permita valorar el logro alcanzado y las áreas de oportunidad para optimizar el proceso de enseñanza-aprendizaje. Esto puede realizarse mediante listas de cotejo, escalas de comprobación o rúbricas, las cuales pueden co-diseñarse con apoyo de las IAGen.

De acuerdo con Guerra (2025), entre las aplicaciones de inteligencia artificial que pueden resultar útiles para este tipo de evaluación se encuentran las siguientes:

Tabla 2. Aplicaciones de la IA útiles para la Evaluación para el aprendizaje

Eduaide.ai	Permite trabajar en cinco dominios: *Content Generator*, *Assistant*, *Feedback Bot*, *Eduaide Chat* y *Assessment Builder*. Este último resulta especialmente útil para apoyar los procesos de evaluación. Su función principal consiste en crear preguntas de evaluación sobre cualquier materia o tema, alineadas con objetivos o estándares educativos. Asimismo, permite construir cuestionarios con distintos niveles de dificultad, tales como opción múltiple, verdadero o falso, ensayo, preguntas abiertas o ejercicios de completar espacios en blanco.
Learnt.ai	Es una herramienta multipropósito con funciones diversas, entre las que se encuentran *Design*, *Teaching*, *Assessment*, *Evaluation*, *Support* y *Promotion* (Guerra, 2025). Para las tareas de evaluación, permite crear criterios de evaluación para un curso, construir rúbricas, elaborar preguntas sobre un tópico, proporcionar retroalimentación al estudiantado destacando fortalezas y áreas de mejora, diseñar marcos de evaluación y retroalimentación entre pares, así como generar ideas para medir el desempeño académico en diferentes procesos.
To-Teach.ai	Se compone de los módulos *Exercises*, *Worksheets*, *Lesson Plan* y *Tools*. Permite elaborar crucigramas, mapas mentales, exámenes de opción múltiple basados en el contenido de un texto, así como generar preguntas con soluciones o ejemplos sobre un tema específico.
Erubrica	Permite crear, aplicar, editar y descargar rúbricas en línea con apoyo de IA, así como almacenar y registrar los resultados. Además, cuenta con un banco de rúbricas, juegos educativos, listas de cotejo y otros recursos para el profesorado.

▶• *Evaluación como aprendizaje*

La IAGen también puede integrarse en la evaluación como aprendizaje. Además de que las herramientas mencionadas para los otros tipos de evaluación pueden aprovecharse en este enfoque, a continuación, se presentan seis estrategias en que esta integración puede llevarse a cabo, de acuerdo con el *Digital Education Council (2025):*

• Autoevaluación y reflexión guiada por inteligencia artificial.

En este caso, el estudiantado interactúa con una herramienta que plantea preguntas, ofrece retroalimentación y reta a cada persona a explicar lo aprendido. Esta dinámica fomenta la metacognición, entendida como la capacidad de reflexionar sobre la propia forma de aprender. Por ejemplo,

una alumna o un alumno puede utilizar una herramienta de IAGen para poner a prueba un argumento construido en clase o para explicar un concepto complejo como si lo enseñara a un principiante.

• Primero la inteligencia artificial, después la revisión humana.

En este enfoque, la inteligencia artificial genera un primer borrador —ya sea un texto, un ensayo o un fragmento de código— y posteriormente se solicita al estudiantado revisarlo, corregirlo y mejorarlo. Esta metodología resulta útil porque obliga a identificar errores, incoherencias o sesgos y, a partir de ello, construir un producto de mayor calidad. Así, la tarea deja de ser un ejercicio pasivo y se convierte en una oportunidad de análisis crítico.

• Primero el trabajo humano, después la revisión con inteligencia artificial.

A diferencia del caso anterior, en este enfoque el estudiantado inicia la tarea de manera independiente, lo que asegura que pueda demostrar su conocimiento inicial. Una vez finalizado el trabajo, se recurre a la IAGen como revisor que ofrece retroalimentación y señala posibles mejoras. Esto permite comparar el producto original con las sugerencias de la herramienta, evaluando cuáles adoptar y cuáles descartar. Se trata, en esencia, de utilizar la tecnología como un segundo lector o evaluador crítico.

• Materiales generados por inteligencia artificial para su análisis.

La IAGen también puede utilizarse para crear ejemplos, casos o escenarios que el estudiantado debe analizar con las herramientas propias de su disciplina. De este modo, la tarea se centra en aplicar criterios académicos para evaluar la calidad de lo producido por la herramienta. Este método permite, por ejemplo, examinar un texto que imite el estilo de un autor reconocido o analizar razonamientos generados por la inteligencia artificial en torno a un dilema ético o científico.

• Inteligencia artificial como colaborador simulado.

Otra posibilidad es utilizar la inteligencia artificial como un interlocutor que adopta un papel específico, como un paciente en una práctica clínica, un personaje histórico o una persona experta en un área determinada. El estudiantado dialoga con esta simulación para ensayar habilidades de comunicación, toma de decisiones o análisis crítico. De este modo, se crean espacios seguros de diálogo y práctica que resultan especialmente valiosos antes de enfrentar situaciones reales.

• Aprendizaje inmersivo con apoyo de inteligencia artificial.

Finalmente, se plantea el uso de IAGen para crear experiencias inmersivas que permitan explorar un tema con mayor profundidad. El estudiantado puede, por ejemplo, generar un relato en primera persona ambientado en un hecho histórico o diseñar imágenes que representen conceptos abstractos como la resiliencia o la entropía. Al reflexionar sobre estas producciones, se favorece una comprensión más creativa y significativa de los contenidos.

En conjunto, estas seis estrategias de integración del uso de IAGen muestran que esta tecnología puede ser una aliada pedagógica para potenciar la evaluación como aprendizaje, siempre que se utilice con un propósito educativo claro y en coherencia con los principios éticos y pedagógicos establecidos en el primer apartado de esta guía. Más que ofrecer respuestas inmediatas, se busca que profesorado y estudiantado desarrollen

pensamiento crítico, creatividad y capacidad de análisis frente al uso de estas tecnologías, pues tales habilidades son indispensables en un contexto donde convivirán permanentemente con ellas y con nuevos avances tecnológicos.

2.3 Otras aplicaciones útiles de IAGen para la evaluación

En este apartado se presentan diversas aplicaciones de IAGen que pueden resultar útiles para apoyar distintos momentos del proceso evaluativo. Aunque no todas están diseñadas específicamente para fines de evaluación, sus funcionalidades pueden aprovecharse para enriquecer el diseño de instrumentos, fortalecer la retroalimentación y facilitar el análisis del aprendizaje

▶• Analítica del aprendizaje (*Learning Analytics*) y vigilancia del desempeño

La analítica del aprendizaje se refiere a la medición, recopilación, análisis y comunicación de datos sobre el alumnado, con el propósito de mejorar la práctica educativa, optimizar el rendimiento académico y fortalecer la gestión y calidad del modelo educativo (Pedraza, 2022).

Las IAGen se han mostrado como herramientas potencialmente útiles para estos procesos. Por ejemplo, en el Royal College de Canadá se analiza su uso para integrar múltiples fuentes de datos —exámenes escritos, simulaciones y desempeño clínico— y trazar trayectorias individuales de competencia para compararlas con curvas esperadas (Royal College of Physicians and Surgeons of Canada, 2024).

En universidades mexicanas se han publicado estudios que describen innovaciones evaluativas basadas en herramientas digitales capaces de recopilar datos, analizar patrones de respuesta y apoyar la toma de decisiones curriculares (Luna de la Luz & González-Flores, 2020).

El uso de IAGen en labores de analítica del aprendizaje posibilita el diagnóstico temprano de estudiantes con bajo rendimiento, facilita la realización de ajustes curriculares basados en evidencia real y permite generar informes para el profesorado sobre los temas con mayor índice de error.

Sin embargo, deben considerarse diversos desafíos, como la protección de datos personales y la privacidad del estudiantado, la interpretación ética de los resultados —incluida la necesidad de evitar etiquetas o clasificaciones estigmatizantes—, así como los requerimientos de infraestructura tecnológica y formación en ciencia de datos.

▶• Generación de reactivos de opción múltiple (ROM) con IAGen

Las IAGen pueden utilizarse para elaborar cuestionarios de opción múltiple. En algunas escuelas de medicina en Estados Unidos ya se emplea IA para generar cuestionarios de práctica, exámenes intermedios y bancos adicionales de preguntas para el estudiantado (AAMC3).

Un estudio reciente comparó preguntas generadas por ChatGPT-4 con preguntas elaboradas por expertos clínicos para exámenes de emergencias médicas. Aunque las preguntas generadas por IA tendieron a ser más sencillas, mostraron índices de discriminación similares. No obstante, también se identificaron más errores fácticos, elementos irrelevantes y fallas de adecuación en los reactivos producidos por IA (BioMed Central 1).

Entre las ventajas observadas de esta aplicación destacan la posibilidad de generar un gran volumen de reactivos en poco tiempo, la creación automática de distractores (respuestas incorrectas) y el ahorro de tiempo para el profesorado.

Entre sus limitaciones se encuentran la necesidad de una revisión humana rigurosa para evitar errores conceptuales, el ajuste adecuado del nivel cognitivo —ya que la IA tiende a generar preguntas superficiales— y el riesgo de sesgos o de falta de alineación con los objetivos de aprendizaje.

▶• Simulaciones clínicas con retroalimentación automática

Las IAGen aplicadas a los campos de ciencias de la salud pueden apoyar la realización de simulaciones con retroalimentación automatizada. Por ejemplo, el proyecto MedSimAI desarrolla encuentros con pacientes virtuales basados en modelos de lenguaje, capaces de generar retroalimentación automática utilizando escalas clínicas establecidas, como la *Master Interview Rating Scale.* En un estudio piloto con estudiantes de primer año, se observó que valoraron la posibilidad de practicar en múltiples ocasiones y recibir retroalimentación inmediata (arXiv 4).

En el artículo *Authentic assessment in medical education,* los autores exploran cómo combinar la IA con metodologías de evaluación auténtica y colaborativa (*student-as-partners*) para crear escenarios clínicos más realistas, personalizados y flexibles (PubMed 5).

Diversas instituciones médicas están experimentando con "pacientes virtuales" inteligentes que responden dinámicamente a las decisiones del estudiantado y permiten evaluar el razonamiento clínico, la comunicación y la toma de decisiones en escenarios simulados (AAMC 3).

La aplicación de estas herramientas en las áreas de la salud permite practicar habilidades clínicas sin riesgo para el paciente real, evaluar competencias "blandas" como la comunicación, la empatía y el razonamiento clínico, y ofrecer retroalimentación inmediata al estudiantado para favorecer la mejora continua.

Entre las principales retos se encuentran lograr que la simulación sea suficientemente realista y no predecible, detectar decisiones atípicas o errores poco frecuentes que el modelo no anticipa, así como afrontar los costos computacionales, los requerimientos de infraestructura y la necesidad de capacitación del profesorado.

▶• Corrección automática de respuestas abiertas, notas clínicas o "*short answers*"

Las IAGen también se han utilizado para calificar notas médicas, respuestas cortas y producciones narrativas mediante técnicas de procesamiento de lenguaje natural (NLP). Por ejemplo, en evaluaciones *on demand* del NBME (National Board of Medical Examiners), se exploran aplicaciones de IA para la puntuación automática de notas médicas y la generación de reactivos o distractores (nbme.org 6).

En la guía de la AMEE sobre *Health Professions Education* se dedica un capítulo a describir cómo utilizar la IA para apoyar la corrección automática en evaluaciones formativas y sumativas. Asimismo, la guía aborda los riesgos éticos, la importancia de la transparencia y la necesidad de realizar ajustes pedagógicos (geiselmed.dartmouth.edu 7).

Entre los beneficios de esta aplicación de las IAGen se encuentran la retroalimentación inmediata en exámenes escritos con respuesta libre, la reducción de la carga del profesorado en la corrección manual y la posibilidad de escalar el proceso para cohortes numerosas.

No obstante, entre las limitaciones a resolver se encuentran la necesidad de que la IA interprete correctamente respuestas complejas o con lenguaje impreciso, evitar que el estudiantado "entrene al modelo" con errores recurrentes y garantizar una evaluación justa y libre de sesgos en la calificación.

3. Recomendaciones para el profesorado

La incorporación de la Inteligencia Artificial Generativa (IAGen) en los procesos de evaluación representa una oportunidad para fortalecer la práctica docente, siempre que su uso se realice de manera informada y responsable. Aunque estas herramientas pueden apoyar la elaboración de materiales, la retroalimentación y el análisis del desempeño estudiantil, su efectividad depende de la supervisión y el criterio profesional del docente. Por ello, es fundamental conocer sus posibilidades, límites y riesgos, así como las acciones necesarias para garantizar una implementación ética y pedagógicamente pertinente.

Este apartado aborda recomendaciones prácticas para orientar al profesorado en la integración de la IAGen dentro de sus evaluaciones, ofrece lineamientos para diseñar evaluaciones asistidas por IAGen, validar los productos generados, prevenir el uso acrítico por parte del estudiantado y orientar la implementación de casos prácticos dentro de marcos institucionales sólidos y pedagógicamente coherentes, proporcionando pautas claras para aprovechar sus beneficios sin comprometer la validez, la equidad ni la formación integral del estudiantado.

3.1 Cómo diseñar evaluaciones aprovechando la IAGen

En su rol como mediadores del aprendizaje, los docentes deben promover un uso ético, crítico y pedagógico de la IAGen. Esto implica integrar estas herramientas de manera reflexiva, favoreciendo no solo la simplificación de tareas, sino también la reflexión, el análisis y la mejora continua del proceso evaluativo.

La incorporación de la IAGen requiere valorar sus beneficios y limitaciones, considerando siempre el desarrollo de competencias y la autonomía del estudiantado (UNESCO, 2023a).

A continuación, se proponen algunas pautas para iniciar un proceso de evaluación apoyado en IAGen:

▶• Preguntas sobre la integración de la IAGen en la evaluación.

- ¿Cuál es la intencionalidad pedagógica de incorporar herramientas de IAGen en este proceso de evaluación?
- ¿Qué expectativas tiene respecto a su uso?
- ¿Qué decisiones se delegarán a la herramienta y cuáles permanecerán bajo la responsabilidad del docente?
- ¿Qué errores o sesgos podrían derivarse del uso de IAGen en este proceso y cómo podrían abordarse?
- ¿De qué manera se documentará el proceso de incorporación de estas herramientas?

▶• Preguntas sobre el propio proceso de evaluación.

- ¿Cuál es el objeto de la evaluación?

- ¿Qué tipo de información requiere recopilar?
- ¿Qué instrumentos de evaluación resultan más pertinentes?
- ¿De qué manera puede la IAGen apoyar en la recolección, el análisis o retroalimentación?

▶• Antes de seleccionar herramientas de IAGen, el docente debe identificar:

- Nivel educativo, asignatura, subdisciplina y semestre.
- Perfil del estudiantado.
- Estándares u objetivos de aprendizaje.
- Tipo de evaluación (diagnóstica, formativa o sumativa).
- Uso y comunicación de los resultados.

▶• Aspectos clave para la selección:

- Utilidad para el tipo de evaluación pretendida.
- Tipos de instrumentos que puede generar (rúbricas, reactivos, listas de cotejo, entre otros).
- Ventajas y limitaciones de cada herramienta.
- Privacidad, precisión, accesibilidad y costo.
- Posibles combinaciones con otras herramientas.

▶• Co-creación con IAGen: diseño de criterios e instrumentos.

Se puede solicitar a la IAGen sugerencias de criterios para evaluar los procesos de aprendizaje. A partir de los resultados generados, el docente puede añadir o eliminar criterios según los objetivos de aprendizaje establecidos.

Ejemplos de *prompts* para la co-creación de criterios de evaluación:

"Propón criterios de evaluación para esta planeación."

"Describe los criterios de evaluación para esta actividad."

"Agrega un criterio para evaluar el pensamiento crítico."

▶• Co-creación de instrumentos de evaluación.

Para la co-creación de instrumentos se puede solicitar a la IAGen que genere propuestas iniciales que posteriormente deberán ser revisadas y ajustadas por el docente.

Ejemplos de *prompts* para la co-creación de instrumentos de evaluación:

"Eres un profesor que imparte la materia de mercadotecnia a alumnos de 5° semestre de licenciatura, eres especialista en el tema y en evaluación educativa. Genera una rúbrica con 3 niveles de desempeño para el proyecto final que consiste en..."

"Diseña una lista de cotejo que sirva de guía a los estudiantes para el desarrollo de su proyecto final que consiste en... y que cubra los criterios necesarios para verificar el logro de los aprendizajes esperados que son..."

Los instrumentos generados deben ajustarse manualmente para adecuarse al contexto, nivel y necesidades del estudiantado, cuidando y corrigiendo siempre los sesgos que reproducen estas tecnologías.

▶• Consideraciones importantes

- Las versiones gratuitas de herramientas tienen funciones limitadas.
- Es necesario contar con consentimiento informado si se cargan datos o productos personales.
- Reconocer y comunicar el uso de la IAGen en la evaluación.
- Identificar los posibles sesgos y limitaciones disciplinarias.

3.2 Cómo revisar y validar las respuestas generadas por IAGen

La revisión crítica y la interpretación de los resultados generados por herramientas de IAGen son procesos indispensables dentro de la formación académica. En este sentido, la supervisión docente no solo garantiza la validez y el uso ético de dichos resultados, sino que también permite identificar limitaciones presentes en el proceso, ya sean técnicas, metodológicas o derivadas del propio aprendizaje del estudiante. A partir de este análisis, los resultados se convierten en un recurso formativo que favorece la metacognición y, al mismo tiempo, brindan al docente insumos para mejorar las estrategias de enseñanza.

Según Quinde-Moncerrate, Quinde-Zambrano y Franco-Arroyo (2025), la IAGen debe utilizarse como un recurso complementario orientado a prácticas éticas, inclusivas y pedagógicamente significativas. El uso didáctico de la IAGen debe promover:

- La reflexión sobre la elección de herramientas y sus limitaciones.
- La valoración de las necesidades del estudiantado, permitiendo ajustar contenidos, estrategias y ritmos de trabajo.
- El análisis ético, social y cultural de la tecnología.
- La importancia de evaluar el proceso, no solo el producto generado con la IAGen.
- La comprensión del funcionamiento y de los límites de la IAGen.

3.3 Prevención del uso indebido o acrítico por parte del estudiantado

Cuando se fomenta el uso de IAGen entre el estudiantado, es importante también propiciar el cuestionamiento sobre:

- Las limitaciones e implicaciones éticas, sociales, culturales y ambientales.
- El procedimiento de interacción entre el estudiantado y la IAGen.
- La comprensión del funcionamiento de la IAGen.
- La integración de los conocimientos previos con los resultados generados por la IAGen.

Para evitar el uso acrítico por parte del estudiantado y para no limitar el proceso de evaluación del aprendizaje al resultado o producto final generado con la IAGen, en la **Tabla 3** se presentan cinco dimensiones evaluables que pueden servir como indicadores del grado de incorporación crítica y creativa que cada estudiante realiza en una actividad específica dentro de su proceso educativo.

Tabla 3. Ejemplo de rúbrica para evaluar la incorporación crítica y creativa de la IAGen en el proceso educativo. (Videla, 2025, p. 17)

Dimensión	Nivel alto	Nivel medio	Nivel inicial
Pensamiento crítico	Detecta y analiza sesgos, propone mejoras y contrasta fuentes.	Identifica errores o limitaciones, pero sin mucha profundidad.	Acepta las respuestas de la IA sin cuestionamientos.
Creatividad	Propone un producto innovador, que articula varios lenguajes y herramientas.	El producto es funcional y adecuado, pero predecible.	Se limita a replicar un modelo sin aportar valor agregado.
Comprensión de la IA	Explica con claridad que hizo la IA y qué hizo él/ella, con fundamentos técnicos.	Tiene nociones generales pero confunde roles o procesos.	No diferencia el rol de la IA del humano en la producción.
Ética y ciudadanía	Incluye una reflexión profunda sobre los impactos de la IA.	Muestra preocupación ética pero sin fundamentación sólida.	No incluye perspectiva ética ni social del uso de la IA.

3.4 Casos prácticos y sugerencias de implementación

La adopción de nuevas tecnologías en educación depende directamente de la alfabetización digital y la capacitación docente. Un estudio de Báez-Vizcaino (2025) muestra que herramientas interactivas como Quizizz y Kahoot presentan una alta aceptación debido a su facilidad de uso y a la retroalimentación inmediata que brindan al estudiantado, lo que contribuye a una mayor motivación académica. En contraste, plataformas más complejas como JigSpace y Chromeville Science enfrentan barreras de uso, especialmente por la falta de dispositivos adecuados y de habilidades tecnológicas avanzadas.

En el caso de herramientas de IAGen, como ChatGPT, se observa una rápida adopción por parte del estudiantado, lo que evidencia la urgencia de fortalecer las competencias digitales docentes, no solo para su dominio operativo, sino también para orientar su uso pedagógico, ético y crítico. Esto coincide con estudios como el de Tondeur et al. (2017), que señalan que el uso de la tecnología depende del propósito formativo, invitando a extender su aplicación más allá del aula.

Diversas instituciones educativas de nivel superior han implementado programas, talleres y seminarios para capacitar a profesores y estudiantes en el uso formativo de la IAGen. Entre ellas destacan la Universidad Central de Ecuador, la Universidad Laica Vicente Rocafuerte (ULVR), la Universidad Andina Simón Bolívar (UASB) y la UNAM, donde se han desarrollado iniciativas orientadas a la integración de la IAGen para la creación de contenidos académicos, el diseño instruccional y la generación de instrumentos de evaluación basados en rúbricas con supervisión docente.

Estas experiencias muestran que la IAGen puede potenciar significativamente la educación superior,

especialmente en áreas como el aprendizaje personalizado, el acompañamiento tutorial automatizado, la generación de contenido académico y la optimización de procesos administrativos y evaluativos. No obstante, dicho potencial solo puede aprovecharse plenamente cuando se vincula a procesos formativos institucionales, políticas de uso y programas de actualización docente continua.

Es necesario concebir a la IAGen como un recurso complementario que fortalezca la toma de decisiones pedagógicas, sin reemplazar el criterio profesional ni la interacción humana. Para su implementación en licenciatura, se recomienda desarrollar casos prácticos en los que los estudiantes utilicen IAGen para análisis, diseño, resolución de problemas y generación de propuestas, acompañando estas actividades con criterios éticos, verificación de fuentes, reflexión crítica y estrategias de evaluación que consideren la autenticidad del aprendizaje (Campuzano-Vázquez, 2025).

La integración de la IAGen en los procesos de evaluación requiere planificación, criterio pedagógico y una implementación gradual. Las siguientes sugerencias ofrecen orientaciones prácticas para que el profesorado incorpore la IAGen de manera ética, segura y alineada con los objetivos formativos. Más que centrarse en la herramienta, es importante tomar decisiones informadas, supervisar los resultados y mantener el papel central del juicio docente en cada etapa del proceso evaluativo.

▶• *Comenzar con claridad pedagógica, no con la tecnología*

La IAGen debe integrarse no como un añadido, sino como parte del diseño curricular y evaluativo. Los ítems, simulaciones o actividades generadas deben alinearse con los objetivos formativos y contribuir al mapeo de competencias definidas (diagnóstico, comunicación, ética, entre otros).

Antes de seleccionar herramientas de IAGen, es necesario definir con precisión:

- La competencia o el resultado de aprendizaje que se desea evaluar (conocimiento, razonamiento, habilidades comunicativas o actitudes).
- La etapa de evaluación correspondiente (diagnóstica, formativa o sumativa).
- El tipo de evidencia más adecuada para valorar dicha competencia (preguntas de opción múltiple —MCQ—, estudio de casos, simulación, ensayo reflexivo o desempeño práctico).

La IAGen debe servir al objetivo pedagógico, no sustituirlo ni distorsionarlo.

▶• *Utilizar IA como apoyo, no como sustituto de la evaluación docente*

Ningún sistema de IAGen debe operar sin validación humana y docente; los errores fácticos o conceptuales siguen siendo una preocupación real, especialmente en evaluaciones de alto riesgo.

La IAGen es más efectiva cuando complementa la labor del profesorado. Se recomienda:

- Generar borradores de preguntas que el docente revise y ajuste.
- Proponer escenarios que puedan adaptarse al contexto local.
- Calificar respuestas abiertas mediante rúbricas predefinidas, con revisión final humana en los casos pertinentes.
- Ofrecer retroalimentación inmediata que posteriormente se profundice en clase.

El juicio experto y pedagógico del docente sigue siendo insustituible.

▶• *Empezar con usos acotados y de bajo riesgo*

En lugar de aplicar la IAGen directamente en evaluaciones de alta relevancia —como los ECOEs finales—, es recomendable:

- Utilizarla en evaluaciones formativas (cuestionarios diagnósticos, prácticas, autoevaluaciones).
- Generar bancos de preguntas de práctica para el estudiantado.
- Implementar simulaciones clínicas virtuales como complemento de las rotaciones.

Estas acciones permiten que profesorado y estudiantado se familiaricen gradualmente con la herramienta, identifiquen posibles problemas y realicen mejoras antes de escalar su uso.

▶• *Revisar críticamente la producción de IAGen*

Toda pregunta, caso clínico, rúbrica o retroalimentación generada mediante IAGen debe pasar por una revisión docente. Se sugiere, por tanto:

- Verificar la exactitud científica, ya que los errores fácticos siguen siendo frecuentes.
- Ajustar el nivel cognitivo de las preguntas o actividades (evitar formulaciones superficiales cuando se busca evaluar razonamiento clínico).
- Asegurar la pertinencia cultural y lingüística de los materiales generados.
- Corroborar que los productos estén alineados con los objetivos de la asignatura y del currículo.

Es útil considerar a la IA como un "asistente junior" que requiere supervisión experta y validación permanente.

▶• *Integrar la IA con criterios de evaluación claros y transparentes*

La transparencia en el uso de la IAGen favorece la confianza, la equidad y la comprensión del proceso evaluativo. Por ello, se recomienda:

- Explicar al estudiantado cómo, por qué y con qué fines se utiliza la IAGen en el curso o actividad.
- Definir y comunicar con antelación las rúbricas o criterios utilizados en cualquier calificación automática o asistida por IA.
- Establecer mecanismos de revisión o apelación cuando exista desacuerdo con la evaluación generada o apoyada por la IA.
- Asegurar que la herramienta utilizada no opere como una "caja negra" y que sus criterios sean comprensibles para el profesorado y el estudiantado.

▶• *Formarse y mantenerse actualizado*

Muchos docentes requieren formación en el uso de la IAGen, en la interpretación de sus resultados y en la supervisión de los modelos. Por ello, su integración al proceso educativo demanda un respaldo institucional sólido, tanto en infraestructura (servidores, licencias, conectividad) como en políticas que regulen y orienten su uso pedagógico y ético. Dado que la IAGen educativa evoluciona rápidamente, se recomienda:

- Participar en talleres institucionales sobre IA y evaluación.
- Compartir experiencias con colegas mediante comunidades docentes o grupos de innovación educativa.
- Consultar guías reconocidas —por ejemplo, AMEE Guide No. 178 (2025), sobre IA en evaluación en educación en salud—.

- Practicar con herramientas en entornos seguros, como *sandboxes* o espacios de prueba. La alfabetización digital docente forma ahora parte esencial de las competencias profesionales en educación.

▶• *Cuidar la dimensión ética y la privacidad estudiantil*

Como se señala en los criterios éticos de la primera sección de esta guía, es fundamental diseñar salvaguardas —monitoreo, límites y supervisión— que protejan la dimensión ética del proceso evaluativo y eviten el uso indebido de las herramientas de IAGen. Por ello, se recomienda:

- Verificar que la herramienta utilizada cumpla con la legislación local sobre protección de datos personales (por ejemplo, en México: LFPDPPP; en la Unión Europea: GDPR).
- Evitar almacenar información sensible del estudiantado en plataformas no seguras o sin respaldo institucional.
- Revisar de manera sistemática los posibles sesgos algorítmicos y auditar periódicamente las herramientas empleadas.
- Comunicar con claridad al estudiantado qué datos se recaban, con qué fines y cómo serán protegidos.
- La ética debe formar parte de cada etapa de la innovación educativa.

▶• *Promover un uso reflexivo de IA también en los estudiantes*

Las herramientas de IAGen se entrenan con grandes volúmenes de datos que pueden contener sesgos geográficos, demográficos o culturales. Por ello, es esencial auditar los modelos y garantizar que no favorezcan ni penalicen injustamente a determinados grupos. El uso crítico, la transparencia en el empleo de la IAGen y la existencia de mecanismos de apelación para profesorado y estudiantado son elementos clave.

- Enseñar al estudiantado a utilizar la IAGen como herramienta de autoevaluación y práctica, no como un atajo.
- Diseñar actividades donde el estudiantado compare su razonamiento con el de un modelo de IA, identifique errores y fortalezca su pensamiento crítico.
- Fomentar el uso ético: distinguir entre asistencia legítima y fraude académico.

Formar un estudiantado crítico implica enseñarle a interactuar de manera reflexiva e inteligente con la IAGen.

▶• *Evaluar y ajustar continuamente*

Después de implementar la IAGen en alguna fase del proceso evaluativo, se recomienda:

- Recopilar datos de desempeño y retroalimentación tanto del estudiantado como del profesorado.
- Analizar los errores recurrentes de la IAGen y del propio estudiantado.
- Ajustar las rúbricas, los prompts o la forma de integración de la herramienta.
- Documentar la experiencia para compartirla y mejorarla institucionalmente.

La integración de la IAGen no es un evento puntual, sino un proceso permanente de revisión y mejora.

▶• *Articular la innovación con la política institucional*

Finalmente, es recomendable:

- Coordinar con las autoridades académicas para asegurar que el uso de la IAGen esté alineado con las normas y lineamientos institucionales.
- Participar en la formulación y actualización de políticas internas sobre el uso de IA en la enseñanza y la evaluación.

- Garantizar que la evaluación apoyada por IA no reproduzca inequidades ni sustituya el acompañamiento pedagógico del profesorado.

La IAGen no reemplaza al buen docente: potencia su capacidad para evaluar con mayor evidencia, reducir la carga repetitiva y dedicar más tiempo al acompañamiento del estudiantado. Como señala la literatura especializada, “es esencial que los profesores desarrollen un pensamiento crítico que les permita discernir la calidad y pertinencia de las herramientas basadas en IA, así como su impacto en el aprendizaje y el desarrollo integral de sus estudiantes” Sereño, F. (2024, p.8). Por ello, es fundamental cultivar una actitud de aprendizaje continuo frente a la IA. El uso educativo de estas herramientas requiere no solo habilidades técnicas, sino también reflexión, criterio profesional y responsabilidad ética para adaptarse a los cambios tecnológicos de manera crítica y fundamentada.

4. Recomendaciones para el estudiantado

La integración de la IAGen en el aprendizaje estudiantil abre un panorama de posibilidades para fortalecer el estudio autónomo, la comprensión profunda y el desarrollo de habilidades académicas relevantes para el entorno digital contemporáneo. Sin embargo, para aprovechar plenamente estos beneficios, es indispensable que el estudiantado adopte un uso crítico, ético y responsable de estas herramientas.

Este apartado proporciona recomendaciones prácticas que orientan a los estudiantes en el uso adecuado de la IAGen para estudiar, practicar, reflexionar y citar correctamente los contenidos generados. El objetivo es promover un aprendizaje informado que evite la dependencia tecnológica y fomente la autorregulación, el pensamiento crítico y la integridad académica, asegurando que la IA funcione como un apoyo complementario y no como un sustituto del esfuerzo intelectual.

4.1 Cómo usar IAGen para estudiar, practicar y reflexionar

En un estudio que se cita previamente, realizado en la Universidad Técnica de Machala, León & Apolo, (2025, p.4), subraya la importancia de que los estudiantes adopten un rol activo, crítico y ético frente al uso de la inteligencia artificial en su propio proceso de aprendizaje.

Se recomienda utilizar estas herramientas como un apoyo complementario para reforzar conocimientos, practicar habilidades o mejorar la organización del estudio. Sin embargo, su uso debe basarse en una participación consciente y crítica, evitando depender de ellas como sustituto del razonamiento o del esfuerzo personal.

Los estudiantes deben desarrollar pensamiento crítico, contrastar la información generada por los sistemas de IA y verificar sus fuentes, recordando que estos modelos pueden ofrecer datos erróneos o desactualizados.

También es importante mantener una conducta ética en la elaboración de trabajos académicos, citando adecuadamente los aportes obtenidos mediante la IA y respetando la propiedad intelectual.

El estudio propone que los estudiantes participen en programas de capacitación institucional que les permitan comprender el funcionamiento, los beneficios y las limitaciones de la IA, fortaleciendo así su preparación para los retos académicos y profesionales del entorno digital contemporáneo.

En síntesis, el estudiantado debe realizar un uso crítico de la IAGen.

Para ello, es fundamental comprender qué es la IAGen y cómo funciona, ya que esto permitirá identificar el tipo de respuestas que puede generar y detectar posibles sesgos. Asimismo, es importante recordar que

estas herramientas tienen limitaciones y pueden producir respuestas imprecisas.

"Los modelos de lenguaje tienden a generar hechos incorrectos pero verosímiles [...] pueden inventar hechos que suenan completamente plausibles e incorporarlos sin ninguna advertencia" Caldeiro, (2024, p. 28).

Por ello, se recomienda validar y contrastar la información generada con la IAGen, ya que la utilidad y eficiencia de esta herramienta dependerán de la capacidad del estudiantado para analizar de forma crítica la información obtenida.

4.2 Cómo citar, verificar y no depender exclusivamente de la IA

En el artículo *"La inteligencia artificial en la educación: ¿transformación o infoxicación?* Un análisis crítico de la nueva frontera educativa", Suárez Estavillo (2025) analiza diversos desafíos éticos, como la privacidad de los datos y los sesgos algorítmicos presentes en los resultados generados por la IA. A continuación, se mencionan algunos de ellos:

- Verificar siempre las fuentes originales. Si la IA cita un artículo o documento, verifica siempre la existencia y autenticidad de la fuente original. No confíes únicamente en el resultado generado por la IA.

- Citar contenido generado por IA de manera transparente. En diversos estilos de citación (APA o Chicago, por ejemplo) se recomienda reconocer de forma explícita que una parte del texto fue generada por un modelo de IA, indicando el nombre del modelo, la fecha de generación y otros elementos pertinentes.

- No otorgar "autoría" completa a la IA. En general, los trabajos académicos deben tener un autor humano. La IA puede emplearse como herramienta o fuente auxiliar, pero no debe considerarse como "autor" en la producción académica.

- Reconocer las limitaciones de la IA. Los modelos pueden generar contenido convincente que no es verídico o que carece de respaldo. Por ejemplo, un estudio sobre motores generativos encontró que solo alrededor del 51.5 % de las oraciones estaban completamente respaldadas por las citas proporcionadas.

- Usar criterios de confianza para evaluar fuentes. Herramientas como la prueba CRAAP (actualidad, relevancia, autoridad, precisión y propósito) ayudan a determinar si una fuente es fiable.

Asimismo, enfoques como el método SIFT ("Stop; Investigate; Find better coverage; Trace claims") promueven contrastar la información con múltiples fuentes en lugar de quedarse únicamente con una.

- Aplicar control humano y juicio crítico. Aunque la IA puede sugerir ideas, ampliar borradores o apoyar en la organización del texto, siempre se debe revisar y modificar el contenido final, asegurando su veracidad, coherencia y originalidad.

4.3 Recomendaciones para citar el uso de IAGen

La citación de contenido generado por IAGen todavía se encuentra en discusión académica. Sin embargo, en términos generales, suele citarse como una comunicación personal, dado que implica una interacción directa con un sistema algorítmico considerado como una herramienta intelectual y no como un autor. Por ello, se recomienda mencionar al desarrollador del modelo, seguido del año de consulta entre paréntesis.

Por ejemplo, en estilo APA se sugiere utilizar el nombre del desarrollador como autor, seguido de la fecha de generación y del nombre de la herramienta de IA empleado como título (Google AI, 2025).

Otra opción es integrar la referencia en el cuerpo del

texto. Por ejemplo: "De acuerdo con Google AI (2025), para citar contenido generado por IA en estilo APA se debe utilizar el nombre del desarrollador como autor, la fecha de generación y el nombre de la herramienta como título."

La Universidad de Guanajuato (UG, 2024) sintetiza varias estructuras y ejemplos para citar contenido generado por IAGen conforme a la versión más reciente del estilo APA (2025).

Para citar textos generados mediante IAGen en estilo APA (2025), la estructura general de referencia es la siguiente:

- **Desarrollador**: entidad responsable del modelo (por ejemplo, OpenAI para ChatGPT; Google AI para Gemini).
- **Fecha**: año de consulta.
- **Nombre del modelo**: nombre de la herramienta en cursivas, seguido de la versión si está disponible (por ejemplo, *Claude 2*).
- **Descripción entre corchetes**: tipo de modelo empleado, como [Modelo de lenguaje de gran tamaño] o [Large language model].
- **Fuente**: URL de acceso a la herramienta.

Ejemplos de referencias en estilo APA (2025):

- **OpenAI**. (2025). *GPT-4o* [Modelo de lenguaje de gran tamaño]. https://chat.openai.com
- **Google AI**. (2025). *Gemini 1.5 Pro* [Modelo de lenguaje de gran tamaño]. https://gemini.google.com/
- **Anthropic**. (2025). *Claude 4 Sonnet* [Large language model]. https://claude.ai/new

También es posible citar una conversación completa con la herramienta, ya sea como apéndice o como material suplementario. Para ello, se recomienda incluir el texto íntegro generado por la IAGen junto con el *prompt* utilizado, siguiendo la estructura sugerida por APA para conversaciones con IA:

Desarrollador. (año, mes día). **Título del chat** [Descripción, Chat con IA generativa]. **Nombre del modelo**. **URL**.

Ejemplo:
Perplexity AI. (2025, May 20). *High school grammar topics [Generative AI chat]. Perplexity.* https://www.perplexity.ai/search/a457cb8c-c663-4c9b-b34e-cb03d8108b35

Para citar imágenes o figuras generadas por IAGen, se debe seguir el estilo APA correspondiente. Esto incluye incorporar el título y el pie de figura, colocados según la indicación del formato. El pie debe incluir la fuente de la imagen, el año de generación y, cuando sea necesario, notas explicativas adicionales (UG, 2024).

Los elementos que deben aparecer en el pie de figura son los siguientes (UG, 2024):

- ***Prompt* sintético**: si el *prompt* es corto (menos de 80 caracteres), puede incluirse en cursivas precedido por la palabra *"Prompt"*.
- **Indicador de uso de IA**: referencia explícita como [Imagen generada con IA] o [Ilustración generada por IA].
- **Desarrollador**: institución o empresa responsable del modelo (por ejemplo, *OpenAI* para DALL·E).
- **Fecha**: año de generación de la figura.
- **Modelo utilizado**: nombre y, de ser posible, la versión del modelo.

- **Descripción del modelo entre corchetes**: por ejemplo, [Modelo de lenguaje multimodal] o [Modelo de difusión].
- **Fuente**: URL de acceso a la herramienta utilizada.

Ejemplo:

[Incluir Título de la figura] *Docente universitario*

[Incluir Imagen o figura]

[Pie de figura] Nota. *Prompt:* imagen de *"docente universitario impartiendo clase frente a un grupo de estudiantes"* [Imagen generada con IA]. Microsoft AI, 2024, Kosmos-1 [Modelo de lenguaje multimodal]. https://designer.microsoft.com/

Para referenciar una figura generada mediante IAGen en la lista de referencias, se recomienda emplear el siguiente formato (UG, 2024):

- **Desarrollador**: nombre de la empresa o institución responsable del modelo de IA (por ejemplo, *OpenAI* para DALL·E o *Microsoft AI* para *Microsoft Designer).*
- **Fecha**: año en que se generó la figura.
- **Título o *prompt* sintético**: si la figura cuenta con título, este debe colocarse en cursivas. En caso de no existir, puede incluirse un *prompt* sintético (menor a 80 caracteres), antecedido por la palabra *"Prompt"* y escrito también en cursivas. Si el prompt es extenso, puede omitirse.
- **Indicativo de uso de IA**: se debe señalar explícitamente que la imagen fue generada mediante IA, por ejemplo: *[Imagen generada por IA]* o *[Ilustración generada por IA].*
- **Nombre del modelo multimodal**: incluir el nombre y, cuando corresponda, la versión del modelo empleado para generar la imagen (por ejemplo: *Kosmos-1,* utilizado en Microsoft Designer).
- **Descripción del modelo**: incluir entre corchetes una breve descripción, como *[Modelo de lenguaje multimodal].*
- **Fuente**: proporcionar la URL de acceso a la herramienta o plataforma de IA.

Ejemplo de referencia en lista de referencias:

Microsoft AI. (2024). *Estudiante en biblioteca. Prompt*: *imagen de estudiante realizando tarea en biblioteca* [Imagen generada con IA]. Kosmos-1 [Modelo de lenguaje multimodal]. https://designer.microsoft.com/

En este caso, al incluir la referencia completa de la figura en la lista de referencias, no es necesario repetir todos los elementos en una nota al pie. Basta con indicar el título de la imagen y, en el pie de figura, el desarrollador y la nota *[imagen generada con IA].*

Ejemplo siguiendo el caso anterior:

[Título de la figura] *Estudiante en biblioteca.*

[Imagen o figura]

[Pie de figura] Microsoft AI, 2024, *[imagen generada con IA].*

Referencias consultadas:

APA. (2025). *Citing generative AI in APA Style: Part 1—Reference formats.* https://apastyle.apa.org/blog/cite-generative-ai-references

UG. (2024). *Recomendaciones para citar contenido generado por Inteligencia Artificial en estilo APA.* Universidad de Guanajuato. https://sedigital.ugto.mx/wp-content/uploads/2025/06/Citar-contenido-generado-por-IA-en-estilo-APA.pdf

4.4 Actividades que promueven la metacognición con ayuda de la IA

El estudio de Quinde-Moncerrate, Quinde-Zambrano y Franco-Arroyo (2025) destaca que la Inteligencia Artificial Generativa (IAGen) ofrece oportunidades relevantes para fortalecer el aprendizaje autónomo, personalizado y significativo. De acuerdo con los autores, su integración en el proceso formativo puede favorecer:

- La investigación independiente, al facilitar la exploración guiada de temas, conceptos y fuentes de información.
- La retroalimentación inmediata, que permite comprobar el nivel de comprensión y corregir errores en tiempo real.
- El acceso a contenido adaptado, ajustado a los ritmos, necesidades y estilos de aprendizaje del estudiantado.
- El desarrollo de habilidades cognitivas y digitales, indispensables en los entornos educativos contemporáneos.

Sin embargo, los autores subrayan que estos beneficios requieren un uso ético, responsable y no dependiente de la tecnología, especialmente en actividades que exigen juicio crítico y autonomía intelectual.

▶• La metacognición y su relación con la IAGen

La metacognición, término introducido por Flavell (1979), se refiere a la capacidad de evaluar críticamente los propios procesos de pensamiento y aprendizaje. Este constructo implica reconocer cómo aprendemos, qué estrategias utilizamos, cuáles son más efectivas y cuáles requieren ajustes para optimizar la comprensión y el desempeño académico. Por ello, constituye un eje fundamental para la educación, la resolución de problemas y la toma de decisiones informadas.

De acuerdo con Walker et al. (2025), la metacognición integra dos componentes centrales:

- **Conocimiento metacognitivo**: corresponde a la conciencia y comprensión que una persona posee sobre sus propios procesos cognitivos. Incluye la identificación de fortalezas y debilidades, de capacidades y limitaciones, así como la valoración de la eficacia de diversas estrategias de aprendizaje o resolución de problemas.
- **Regulación metacognitiva**: implica supervisar y controlar las propias actividades cognitivas mediante la planificación, el monitoreo del progreso y la evaluación de los resultados. Este proceso permite realizar ajustes oportunos para mejorar la calidad del aprendizaje.

Asimismo, Zakrajsek (2025) señala tres áreas metacognitivas que pueden fortalecerse para favorecer la autorregulación del aprendizaje: la planificación, el seguimiento y la reflexión.

- **Planificación**: permite anticipar y organizar el proceso de aprendizaje, establecer objetivos y criterios de éxito, seleccionar estrategias pertinentes e identificar las condiciones y contextos que favorecen un aprendizaje más efectivo.
- **Seguimiento**: consiste en monitorear el progreso para verificar si los objetivos se están cumpliendo, si las estrategias funcionan como se esperaba y si es necesario realizar ajustes durante el proceso.
- **Reflexión**: se lleva a cabo al finalizar una actividad académica o una sesión de estudio. Implica analizar el proceso seguido para determinar si se alcanzaron los objetivos, qué resultados se lograron, qué nuevas necesidades surgieron y qué modificaciones deben considerarse para mejorar el aprendizaje en futuras tareas.

▶• Estrategias prácticas para cada área de la metacognición con IAGen

El uso de herramientas de Inteligencia Artificial Generativa (IAGen) puede apoyar significativamente los procesos metacognitivos de planificación, seguimiento y reflexión del aprendizaje. A continuación, se proponen diversas acciones para potenciar cada una de estas dimensiones.

Para facilitar la planificación metacognitiva

- Solicite a la herramienta de IAGen apoyo para **definir objetivos de aprendizaje** para una sesión de estudio y para identificar **estrategias de aprendizaje pertinentes** según el contenido que debe abordar.
- Pida a la herramienta que le ayude a **elaborar un plan de estudio específico**, que incluya metas de aprendizaje, fechas límite, pausas programadas y estrategias recomendadas. Mientras más precisa sea la solicitud, más útil será el plan generado.
- Solicite a la herramienta que **identifique las principales dificultades cognitivas** asociadas al contenido por estudiar y que sugiera estrategias para superarlas.

Para facilitar el seguimiento del aprendizaje

- Utilice herramientas de evaluación educativa basadas en IA para **generar cuestionarios cerrados** que permitan comprobar su nivel de dominio sobre el contenido.
- Solicite a la herramienta **estrategias alternativas de aprendizaje** cuando experimente dificultades con un contenido curricular e implemente aquellas que considere más adecuadas para avanzar hacia los objetivos establecidos.
- Use una herramienta de **seguimiento de proyectos o tareas**, basada en metas u objetivos, que le permita visualizar su progreso a lo largo del tiempo y verificar el cumplimiento de sus metas de aprendizaje.

Para facilitar la reflexión metacognitiva

- Emplee una herramienta que le ayude a **formular preguntas de reflexión** sobre su proceso de aprendizaje y a autoevaluar su progreso. Pida retroalimentación para identificar fortalezas, áreas de mejora y patrones de desempeño.
- Solicite a la herramienta **nuevas estrategias o soluciones** para atender las dificultades o debilidades detectadas durante el proceso de estudio.
- Utilice una herramienta conversacional que le permita **analizar su respuesta emocional** al finalizar una sesión de estudio y que ofrezca retroalimentación relacionada con el componente afectivo del aprendizaje.

Si bien las IAGen pueden fortalecer la planificación, el monitoreo y la reflexión del aprendizaje, también es importante reconocer sus limitaciones y riesgos (ESPOL, 2023). En ocasiones, estas herramientas pueden sugerir **estrategias inadecuadas**, interpretar de forma incorrecta las solicitudes del usuario, **ofrecer orientación poco pertinente** o generar respuestas repetitivas que deriven en frustración o desmotivación. Por ello, su uso debe complementarse con juicio crítico y supervisión docente cuando corresponda.

5. Ejemplos Prácticos

El avance de la Inteligencia Artificial Generativa (IA-Gen) ha transformado no solo la teoría de la evaluación educativa, sino también su puesta en práctica cotidiana. Por ello, en este apartado se presentan ejemplos concretos y aplicaciones reales que muestran cómo la IA puede integrarse de manera efectiva en distintos componentes del proceso evaluativo. Desde plantillas diseñadas para construir actividades con apoyo de IA —como las tutorías inteligentes y los exámenes escritos enriquecidos con análisis automatizados— hasta rúbricas adaptadas al uso de herramientas generativas y estrategias de *prompting* orientadas a formular instrucciones de alta calidad, este capítulo ofrece una visión operativa y aplicada. Se presentan algunos casos de éxito, que muestran que la IA no solo optimiza tareas puntuales, sino que reconfigura de manera estructural la planeación, construcción, aplicación, análisis y retroalimentación de las evaluaciones, permitiendo prácticas más personalizadas, eficientes y basadas en datos.

5.1 Plantillas de actividades evaluativas con IA

▶• Sistemas de Tutoría Inteligente (ITS)

Los sistemas de tutoría inteligente constituyen una valiosa herramienta de apoyo para el docente, ya que permiten personalizar el aprendizaje de acuerdo con las necesidades, ritmos y progresos de cada estudiante. Estos sistemas analizan de manera detallada el comportamiento del alumno al interactuar con ejercicios o problemas, identificando patrones en sus estrategias de resolución. A partir de este análisis, el sistema estima el nivel de competencia cognitiva del usuario y ofrece retroalimentación oportuna y específica. Como resultado, las tutorías inteligentes facilitan la adaptación dinámica del plan de trabajo individual, fortaleciendo el acompañamiento pedagógico y la mejora continua del aprendizaje.

▶• Exámenes escritos

Al emplear una de las herramientas más conocidas en los procesos de evaluación —los exámenes escritos, particularmente aquellos que utilizan reactivos de opción múltiple (ROM)— es posible identificar diversas aplicaciones prácticas y casos de éxito impulsados por la IA. Con el surgimiento de la Inteligencia Artificial Generativa (IAG) y sus variantes basadas en aprendizaje automático (machine learning), procesamiento de lenguaje natural (NLP) y modelos generativos, la evaluación educativa atraviesa un punto de inflexión histórico.

El impacto de estas tecnologías no es simplemente lineal ni meramente instrumental; es un impacto **estructural**. La IA habilita nuevas dinámicas de retroalimentación continua, análisis avanzado de datos, detección temprana de patrones, calibración psicométrica más precisa, personalización del aprendizaje y una gestión curricular respaldada por evidencia.

Los casos de éxito reportados por instituciones como el *National Board of Medical Examiners* (NBME), el *Educational Testing Service (ETS)* y plataformas internacionales como Duolingo o Coursera muestran que la IA puede mejorar de manera significativa la **validez, confiabilidad, equidad** y **eficiencia** de las prácticas evaluativas (NBME, 2022; ETS, 2018; Duolingo, 2021).

▶• El ciclo de un examen como modelo robusto de análisis

Para comprender la contribución de la IAG y aplicarla cotidianamente, se adopta el ciclo de un examen como referencia analítica:
Planificación → construcción → revisión → aplicación → calificación → análisis → retroalimentación → decisiones → metaevaluación.

Este modelo permite identificar puntos críticos donde la IA incide con mayor impacto y por ello puede aplicarse.

Aplicaciones de la IAG en el ciclo de un examen:

Planificación. La planificación constituye el núcleo epistemológico del examen, pues determina la relación entre el constructo a medir y el instrumento que lo operacionaliza.

- **IA en la construcción de *blueprints***
  El NBME integra análisis automatizados para ajustar la distribución temática del USMLE con base en evidencia longitudinal (NBME, 2022).
  La IA analiza millones de respuestas históricas y su asociación con niveles de competencia, refinando la sensibilidad del *blueprint* a los cambios curriculares.

- **Alineación curricular automatizada**
  ETS y Khan Academy han mostrado que algoritmos supervisados pueden identificar desalineaciones entre competencias, actividades de aprendizaje y evaluación (ETS & Khan Academy, 2022).
  Esto implica una transición desde un enfoque artesanal hacia un modelo **currículo-evaluación basado en datos**.

Construcción del examen. La IA interviene en la generación, transformación y calibración inicial de reactivos.

- **Generación automática de ítems (NLP + IA generativa)**
  El Duolingo English Test es un caso paradigmático: utiliza IA generativa para crear reactivos equivalentes que luego son validados psicométricamente (Duolingo, 2021).
  Esto ha permitido un volumen de producción inédito sin comprometer la calidad.

- **Ajuste del nivel cognitivo**
  Pearson VUE ha demostrado que la IA puede modificar la complejidad cognitiva de un ítem manteniendo su validez de contenido (Pearson VUE, 2021).
  Esto es especialmente relevante en contextos que requieren medir razonamiento clínico o toma de decisiones.

Revisión y validación. La IA permite un control de calidad más riguroso y sistemático.

- **Análisis automatizado de sesgos**
  Modelos de IA identifican patrones lingüísticos asociados a estereotipos o sesgos culturales. NBME reporta mejoras significativas en la detección temprana de ítems discriminatorios (NBME, 2020).

- **Validación técnica preliminar**
  En Code.org se documenta que la IA reduce en un 30% el rechazo de ítems por fallos de plausibilidad o problemas de redacción (Code.org, 2023).
  Esto genera un proceso de validación más eficiente y consistente.

Aplicación del examen. La IA redefine la seguridad, autenticidad y adaptabilidad de la evaluación.

- **Supervisión inteligente (AI proctoring)**
  Coursera utiliza IA para autenticar identidades, monitorear comportamientos y detectar patrones anómalos sin intervención humana directa (Coursera,

2020). Este tipo de supervisión fue esencial durante la expansión global de la educación remota.

- **Generación dinámica de versiones del examen.** El examen de Duolingo utiliza IA adaptativa para producir versiones únicas garantizando equivalencia psicométrica (Duolingo, 2021). Esto reduce la probabilidad de filtración y aumenta la justicia evaluativa.

Calificación. La IA opera como un evaluador automatizado altamente consistente.

- **Evaluación de escritura (e-rater®)**

ETS ha demostrado que el sistema e-rater® presenta altos niveles de concordancia con evaluadores humanos en tareas complejas de escritura (Berman & Estrada, 2021).

- **Evaluación clínica asistida por IA**

La Universidad de Kyoto ha experimentado con IA multimodal (video, audio, postura) en OSCE/ECOE, logrando reducir la variabilidad interevaluador (Kyoto University, 2022). Esto abre la puerta a una **objetivación** ampliada de competencias clínicas.

Análisis de resultados

- **Psicometría computacional**

ACTNext utiliza modelos híbridos que combinan IRT y aprendizaje automático para detectar ítems defectuosos, estimar confiabilidad y analizar patrones de desempeño (ACTNext, 2019).

- **Analítica de patrones de error**

Instituciones nacionales, como el Instituto Nacional de Evaluación Educativa en Chile, han empleado IA para mapear errores comunes y orientar ajustes curriculares (Benítez & García, 2022).

Realimentación

- **Realimentación personalizada**

El Duolingo English Test incorpora retroalimentación inmediata basada en patrones predictivos (Duolingo, 2021).

- **Tutores inteligentes adaptativos**

Khan Academy, con Khanmigo, ofrece retroalimentación explicativa paso a paso, mejorando la profundización conceptual (ETS & Khan Academy, 2022).

Toma de decisiones

- **Analítica institucional**

Finlandia integra IA en plataformas nacionales para orientar decisiones curriculares y políticas educativas (Finland Ministry of Education, 2021).

- **Detección temprana de riesgo académico**

La Universidad de Arizona ha demostrado que los modelos predictivos reducen de manera significativa la deserción (Abdelnour-Nocera & Sharp, 2023).

Metaevaluación

- **Recalibración continua del banco de ítems**

Duolingo utiliza IA para ajustar la dificultad y discriminación de sus reactivos en tiempo real (Duolingo, 2021).

- **Detección de filtraciones o colusiones**

El NBME emplea IA para identificar patrones que sugieren acceso preexistente a reactivos, protegiendo la validez del examen (NBME, 2022).

▶• Resumen de casos de éxito de IA en cada fase del ciclo de un examen

Planificación del examen

▪ **Caso de éxito 1 –** NBME (National Board of Medical Examiners, EE. UU.)
El NBME utiliza IA y análisis automatizado para construir *blueprints* dinamizados para exámenes de alto impacto (USMLE).
Logro: alineación estricta entre competencias, peso temático y niveles cognitivos.

▪ **Caso de éxito 2 –** Khan Academy + ETS
ETS integró IA para analizar grandes bases de datos de desempeño y ajustar proporciones temáticas en exámenes estandarizados.
Logro: blueprints basados en evidencia de aprendizaje real.

Construcción de reactivos

▪ **Caso de éxito 1 –** Duolingo English Test (DET, 2021)
Utiliza IA generativa y de evaluación automatizada para crear reactivos equivalentes, validados estadísticamente.
Logro: producción masiva de ítems con calidad psicométrica.

▪ **Caso de éxito 2 –** Pearson VUE (2021)
Usa IA para generar borradores de preguntas en certificaciones profesionales.
Logro: reducción del tiempo de construcción de ítems en un 40%.

Revisión y validación

▪ **Caso de éxito 1 –** USMLE (2023) / NBME Psychometrics Lab
Implementaron IA para **detectar sesgos** y errores en ítems clínicos.
Logro: reducción significativa de ítems con problemas de discriminación.

▪ **Caso de éxito 2 –** OpenAI + Code.org (educación básica y media)
IA identifica lenguaje complejo, ambigüedades y sesgos en reactivos antes de su uso.
Logro: reducción de 30% de reactivos rechazados por problemas de formulación.

Aplicación del examen

▪ **Caso de éxito 1 –** Coursera Proctored Exams (Proctortrack / ProctorU)
IA se usa para verificación de identidad y comportamientos anómalos durante exámenes en línea.
Logro: exámenes masivos seguros para miles de estudiantes en entornos MOOC.

▪ **Caso de éxito 2 –** Duolingo English Test (2021)
Gestión automatizada de versiones, adaptabilidad y seguridad.
Logro: examen totalmente en línea aceptado por más de 4,000 universidades.

Calificación

▪ **Caso de éxito 1 –** ETS e-rater®
Califica automáticamente ensayos y respuestas abiertas en TOEFL y GRE.
Logro: alta consistencia con evaluadores humanos.

▪ **Caso de éxito 2 –** AI scoring en OSCE (Universidad de Kyoto, Japón, 2023)
IA analiza video y voz en estaciones clínicas.
Logro: evaluación estandarizada con reducción en variabilidad inter evaluador.

Análisis de resultados

▪ **Caso de éxito 1 –** ACTNext (2019) (Iowa, EE. UU.)
Usa modelos de IA (IRT + machine learning) para analizar ítems y predecir rendimiento.
Logro: mejora en la confiabilidad y estabilidad de los puntajes.

▪ **Caso de éxito 2 –** Instituto Nacional de Evaluación Educativa (Chile)
Utiliza IA para análisis psicométrico automatizado de pruebas nacionales.
Logro: optimización del banco de reactivos y detección temprana de preguntas defectuosas.

Realimentación al estudiante

▪ **Caso de éxito 1 –** Duolingo English Test (2021)
Ofrece realimentación inmediata y personalizada basada en IA.
Logro: mejora documentada en la preparación del usuario sin intervención humana.

▪ **Caso de éxito 2 –** Khan Academy + GPT-4 Tutor
Realimentación esperada, paso a paso, basada en errores reales del estudiante.
Logro: incrementos significativos en la tasa de aciertos posteriores.

Toma de decisiones académicas

▪ **Caso de éxito 1 –** Unidades de Evaluación de Finlandia
IA analiza tendencias nacionales para modificar currículos y pruebas.
Logro: toma de decisiones informada a nivel país.

▪ **Caso de éxito 2 –** Universidad de Arizona
IA identifica estudiantes en riesgo a partir de resultados de evaluaciones.
Logro: reducción de deserción y mejoras en rendimiento académico.

Metaevaluación y mejora continua

▪ **Caso de éxito 1 –** Duolingo English Test (sistema adaptativo contínuo)
Analiza el desempeño mundial para optimizar ítems de manera iterativa.
Logro: el examen se recalibra automáticamente cada ciclo.

▪ **Caso de éxito 2 –** NBME (2022) (USMLE test evolution)
Emplea IA para evaluar el desempeño de ítems en exámenes pasados y mejorar la validez futura.
Logro: banco de reactivos depurado y actualizado continuamente.

A nivel de ensayo experimental, en la Facultad de Medicina de la UNAM, se está aplicando la IAG en áreas tales cómo: desarrollo de pacientes virtuales para aplicación de exámenes prácticos, elaboración y corrección de reactivos de opción múltiple, análisis comparativo de calificación de respuestas abiertas y aplicaciones de psicometría.

5.2 Rúbricas que integran el uso de IAGen

La integración de herramientas de IAGen en procesos evaluativos ha impulsado la creación de rúbricas específicas que contemplan el uso responsable, ético y metodológico de estas tecnologías. Diversos autores y organismos educativos (Becker et al., 2023; Holmes et al., 2022; UNESCO, 2023a) coinciden en que, dada la presencia creciente de la IA en la producción académica, la evaluación debe enfocarse no solo en el producto final, sino también en la calidad del proceso, la transparencia del estudiante, el pensamiento crítico, y la habilidad para verificar y complementar lo generado por la IA.

En este sentido, las rúbricas continúan siendo una de las herramientas de evaluación más valiosas para docentes de todos los niveles, ya que permiten comunicar expectativas, evaluar de manera transparente y favorecer la retroalimentación formativa (Brookhart, 2018). La integración de la IAGen no sustituye la función pedagógica de la rúbrica, sino que amplía sus posibilidades: facilita su diseño, automatiza parte del análisis de desempeño, sugiere mejoras en criterios y descriptores, y permite generar retroalimentación personalizada basada en los niveles alcanzados (Holmes

et al., 2022; Zawacki-Richter et al., 2019). De tal forma que, las rúbricas con IA no son "rúbricas hechas por la IA", sino rúbricas diseñadas por el docente que incorporan herramientas de IA para mejorar su precisión, claridad y utilidad didáctica (UNESCO, 2023a).

La IAGen puede actuar como un asistente especializado que ayuda al docente a:

- Generar borradores de criterios de evaluación alineados a competencias, objetivos de aprendizaje o perfiles de egreso (ISTE, 2023).
- Reescribir descriptores para asegurar claridad, lenguaje progresivo y ausencia de sesgos (ACARA, 2023).
- Verificar el alineamiento entre la actividad, la rúbrica y el nivel cognitivo esperado (Holmes et al., 2022).
- Sugerir evidencias de desempeño coherentes con cada nivel (Becker et al., 2023).
- Adaptar la rúbrica a distintos grados de complejidad (básica, intermedia o avanzada) (UNESCO, 2023a).
- Evaluar muestras de trabajos (con supervisión docente) para identificar patrones de error o áreas de oportunidad comunes (Zawacki-Richter et al., 2019).

Ejemplo práctico de *prompt*

> "Actúa como experto en educación superior. Revisa los criterios de esta rúbrica para evaluar un ensayo académico de licenciatura. Sugiere mejoras en claridad, nivel cognitivo y alineación curricular. Mantén el formato de nivel 'básico–competente–destacado'."

Un enfoque efectivo es diseñar rúbricas "combinadas o híbridas", es decir, una combinación de juicio docente más el análisis de IAGen, en donde:

- El docente define los criterios principales (p. ej., argumentación, uso de fuentes, estructura, originalidad), lo cual coincide con las recomendaciones de evaluación auténtica y diseño de rúbricas centradas en el juicio profesional del profesorado (Brookhart, 2018).
- La IA colabora mediante revisiones, ejemplos de buenos y malos desempeños, o explicaciones detalladas para cada nivel, tal como señalan diversos análisis sobre el uso de IA en educación superior y su potencial para apoyar tareas de retroalimentación y análisis de desempeño (Holmes, Bialik & Fadel, 2022; Zawacki-Richter et al., 2019).
- El proceso de calificación continúa siendo responsabilidad del docente, pero la IA ayuda a generar retroalimentación más rica y diferenciada, así la IA funciona como un apoyo pedagógico y no como sustituto del juicio humano (UNESCO, 2023a).

Esta forma de diseñar las rúbricas permite mantener el control pedagógico, a la vez se aprovecha la capacidad de la IAGen para analizar patrones y formular retroalimentación inmediata (Holmes et al., 2022; Zawacki-Richter et al., 2019).

Ejemplo de rúbrica que se puede utilizar para que el estudiantado mejore su actividad y que el docente puede emplear para generar la retroalimentación:

Actividad: Ensayo argumentativo (1 500 palabras)
Herramienta de apoyo: ChatGPT o similar para retroalimentación formativa previa.

Tabla 4. Ejemplo de Rúbrica para ser usado como retroalimentación del profesorado y alumnado

Criterio	Incipiente	Básico	Competente	Destacado	Apoyo de IA recomendado
Claridad y coherencia argumentativa	Presenta ideas confusas y sin relación lógica.	Argumentos sin desarrollar.	Argumentos claros y con evidencia suficiente.	Argumentación sólida y persuasiva.	IA sugiere reorganización y mejora de transiciones.
Uso de fuentes	No utiliza fuentes confiables.	Usa pocas o mal citadas (1-2).	Usa fuentes pertinentes con citas correctas (3-4).	Integra fuentes variadas y especializadas en el tema (5 y más).	IA revisa estilo de citación y detecta inconsistencias.
Profundidad crítica	No hay análisis.	Hay análisis limitado.	Presenta análisis reflexivo.	Ofrece análisis complejo y multidimensional.	IA identifica vacíos argumentativos y recomienda preguntas guía.
Redacción académica	Más de tres errores frecuentes.	Redacción básica con menos de tres errores.	Estilo fluído y correcto.	Estilo académico sólido y preciso.	IA corrige estilo, tono y precisión terminológica.

Prompt para mejorar la actividad (estudiante)

Actuarás como asistente académico que me ayude a mejorar mi ensayo argumentativo, pero NO lo escribes en mi lugar.

Te proporcionaré mi borrador. Tu tarea es:

1. Evaluarlo usando esta rúbrica: Claridad y coherencia argumentativa, uso de fuentes, profundidad crítica, redacción académica.
2. Decirme qué nivel tengo en cada criterio (*incipiente, básico, competente o destacado*).
3. Sugerirme mejoras concretas sin reescribir el texto completo.
4. Darme ejemplos de mejoras, pero siempre parciales (1–2 frases), nunca textos *completos que sustituyan mi trabajo.*
5. Señalar posibles inconsistencias, vacíos argumentativos y oportunidades para fortalecer mis fuentes.

Cuando estés listo, te enviaré mi borrador.

Prompt para retroalimentación (docente)

Actúa como docente de nivel licenciatura. Necesito retroalimentación formativa basada en esta rúbrica (claridad argumentativa, uso de fuentes, profundidad crítica, redacción académica).

Analiza mi texto y produce un informe con la siguiente estructura:

1. Diagnóstico por criterio:
   - Nivel alcanzado (incipiente, básico, competente, destacado).
   - Evidencias textuales que justifican el nivel.
2. Recomendaciones puntuales de mejora:
   - Reescrituras sugeridas *solo de frases específicas*, no del ensayo completo.
   - Señala vacíos argumentativos y cómo podrían desarrollarse.
   - Propón 3–5 preguntas guía para profundizar el análisis crítico.
3. Uso ético y responsable de IA:
   - Indica si el texto conserva voz propia.
   - Sugiere cómo equilibrar apoyo de IA y autoría personal.

A continuación, te proporcionaré mi texto.

La integración de la IAGen en los procesos de evaluación no implica renunciar al juicio profesional del docente, sino fortalecerlo mediante herramientas que amplían la precisión, la transparencia y el carácter formativo de las rúbricas. Al combinar la experiencia pedagógica con las capacidades analíticas de la IAGen, se promueve una evaluación más equitativa, coherente y orientada al aprendizaje, al tiempo que se fomenta en el estudiantado una cultura de uso crítico, ético y reflexivo de estas tecnologías.

5.3 Prompting educativo: cómo formular instrucciones claras y efectivas para la IA

Los *prompts* son indicaciones o instrucciones que se proporcionan a las herramientas de Inteligencia Artificial Generativa (IAGen) para que ejecuten una tarea específica, como la creación de un texto o una imagen, la identificación de información relevante o el procesamiento de datos. La calidad, claridad y precisión con que se redacta cada instrucción influyen directamente en la pertinencia y utilidad de la respuesta generada por el sistema. Por ello, es recomendable incluir el contexto en el que se utilizará la tarea, así como definir el **formato,** la **modalidad,** el **nivel de detalle** y otros elementos deseables que se espera obtener en la salida generada.

Con relación a la elaboración de *prompts* en herramientas como ChatGPT, *La Gran EnciclopedIA: recursos interactivos de IA* (2025) identifica seis esquemas básicos que pueden facilitar la redacción de instrucciones efectivas. La lógica organizativa de estos esquemas también puede aplicarse a otras herramientas *de IAGen con funcionamiento similar.*

P-A-S (Problema - Análisis - Solución)

P→ Describe el problema en detalle.
A→ Analiza las posibles causas y contexto.
S→ Propón una solución óptima basada en el análisis.

2. **D-E-P (Datos - Explicación - Propuesta)**

D→ Proporciona los datos relevantes.
E→ Explica el contexto en el que se aplican.
P→ Ofrece una propuesta o solución.

3. **R-I-A (Rol - Instrucción - Alcance)**

R→ Especifica el rol que debe asumir ChatGPT.
I → Da una instrucción clara y detallada.
A→ Define el alcance o limitaciones de la respuesta.

4. **P-O-S-E (Problema - Objetivo - Solución - Ejemplo)**

P→ Explica el problema a resolver.
O→ Indica el objetivo deseado.
S→ Pide una solución práctica.
E→ Incluye un ejemplo de aplicación.

5. **E-F-R (Ejemplo - Formato - Refinamiento)**

E→ Da un ejemplo del tipo de respuesta esperada.
F→ Especifica el formato en que deseas la respuesta.
R→ Pide una versión mejorada o refinada.

6. **C-O-T (Contexto - Objetivo - Tarea)**

C→ Define el contexto con información relevante.
O→ Establece el objetivo claro de la solicitud.
T→ Especifica la tarea exacta que debe realizar la IA

7. **Otro más es RAFA**

R→ Rol (actúa como...)
A→ Acción (realiza...)
F→ Formato (como lista, como ensayo, cómo artículo, como cuadro...)
A→ Ámbito (desde la perspectiva, en el área de, bajo el ámbito de...)

A continuación, se presentan algunos ejemplos de *prompts* relacionados con diversas tareas de evaluación educativa. Estos han sido tomados y modificados de la *Gran EnciclopedIA: Recursos interactivos de IA* (2025, pp. 180-194).

Preguntas para mejorar las evaluaciones de ensayos y trabajos escritos

***PROMPT* 1**

Como especialista en inteligencia artificial aplicada a la educación, diseña un modelo para evaluar automáticamente ensayos y tareas con precisión y sin sesgo para la materia/tema [nombrar materia/tema]. Los objetivos de la asignatura/tema son [enlistar objetivos].

***PROMPT* 2 (continuación del anterior)**

¿De qué manera este modelo de apoyo para la evaluación de ensayos basado en IA, es transparente y justo para el estudiantado?

***PROMPT* 3 (continuación del anterior)**

¿Cómo se puede validar que los resultados generados por IA sean tan precisos como la evaluación humana?

Preguntas para automatizar un sistema de evaluación

***PROMPT* 4**

Actúa como un especialista en analítica educativa. Diseña un sistema basado en datos para evaluar el rendimiento académico de los estudiantes de la carrera [Nombre de la carrera], en el curso [Título del curso], cuyos objetivos de aprendizaje son [Nombrar objetivos].

***PROMPT* 5 (continuación del anterior)**

¿Qué métricas clave o indicadores deben considerarse para medir el progreso de los estudiantes en el curso [Título del curso] y detectar estudiantes en riesgo?

***PROMPT* 6 (continuación del anterior)**

¿Cómo se puede usar y aprovechar esta información de las métricas para diseñar planes de intervención personalizados y orientados a mejorar el rendimiento de los estudiantes en el [Título del curso]?

Preguntas para identificar y prevenir el riesgo de abandono escolar

***PROMPT* 7**

Como experto en análisis de retención estudiantil, desarrolla un modelo predictivo que identifique a los estudiantes con mayor riesgo de abandonar sus estudios en el curso [Título del curso], cuyo programa de estudio es [adjuntar el programa o la planeación propuesta]. Los perfiles de inicio de los estudiantes son [adjuntar base de datos, omitiendo información sensible y datos personales].

***PROMPT* 8 (continuación del anterior)**

¿Cómo validar la precisión del modelo?

***PROMPT* 9 (continuación del anterior)**

A partir del modelo ¿qué acciones preventivas pueden tomarse para prevenir el riesgo de abandono?

6. Advertencias y limitaciones

Es importante reconocer que la incorporación de la IA en los procesos de evaluación implica retos significativos. Aunque, como se ha mencionado previamente, estas tecnologías pueden contribuir a reducir ciertos sesgos presentes en el juicio humano, también es cierto que los sistemas de IA pueden reproducir o incluso amplificar sesgos si los algoritmos no han sido diseñados, entrenados o supervisados de manera adecuada. En este sentido, Llanos (2025, p. 81) advierte que *"la calidad de los datos utilizados para entrenar estos sistemas es de suma importancia; datos sesgados pueden llevar a resultados igualmente sesgados, lo que podría perpetuar las desigualdades existentes. Por lo tanto, es crucial que las instituciones educativas no solo incorporen tecnologías de IA, sino que, al mismo tiempo, se aseguren de que estas se apliquen de manera ética y responsable".*

De manera complementaria, el principio de objetividad —frecuentemente mencionado como una de las ventajas de la evaluación automatizada— también ha sido cuestionado críticamente. Al respecto, el mismo autor señala que *"el uso y la implementación de la IA en la evaluación educativa sugiere interrogantes éticas y pedagógicas que no pueden ser ignoradas. La objetividad prometida por estos sistemas automatizados debe equilibrarse con la necesidad de mantener la interacción humana en el proceso educativo, así como una consideración crítica sobre los sesgos inherentes que pueden surgir de algoritmos mal diseñados"* (Llanos, 2025, p. 75).

En síntesis, la evaluación apoyada en IA ofrece oportunidades relevantes, pero exige un análisis ético riguroso y una implementación responsable que preserve el sentido pedagógico y humano de los procesos formativos.

6.1 Alucinaciones, sesgos, problemas de privacidad y acceso desigual

El uso de herramientas de Inteligencia Artificial Generativa (IAGen) en entornos educativos ofrece múltiples posibilidades pedagógicas, pero también plantea desafíos que deben ser comprendidos y gestionados por el profesorado. Conocer estos riesgos es esencial para orientar a los estudiantes hacia un uso crítico, ético y metodológicamente adecuado de la IAGen.

En este apartado revisaremos los más relevantes: las alucinaciones, los sesgos algorítmicos, los riesgos de privacidad y las brechas de acceso (UNESCO, 2023; OECD, 2023).

▶• Las *alucinaciones*

Son respuestas incorrectas, inventadas o no verificables que los modelos generan con aparente seguridad. Estas ocurren debido a las limitaciones inherentes a los modelos probabilísticos y a la forma en que predicen texto a partir de patrones lingüísticos, más que de comprensión factual (Ji et al., 2023).

De no desarrollarse habilidades de verificación; las citaciones falsas, los datos inexactos o explicaciones inventadas pueden integrarse en trabajos académicos. Por ello, es indispensable que los profesores

consideren y enseñen estrategias de contraste con fuentes académicas confiables y fomenten una cultura del escepticismo informado (UNESCO, 2023a).

▶• Sesgos algorítmicos: reproducción de desigualdades

Los modelos generativos aprenden de grandes cantidades de datos disponibles en internet, los cuales contienen sesgos culturales, de género, raciales y socioeconómicos. Esto puede derivar en *outputs* que reproducen estereotipos o distorsionan representaciones sociales (Bender et al., 2021; Weidinger et al., 2022). Para el trabajo docente, esto significa que las actividades que involucren generación de textos, análisis de casos o propuestas creativas deben revisarse con mirada crítica, enseñando al estudiantado a reconocer que la IAGen no es neutral y que sus resultados requieren interpretación contextual.

▶• Problemas de privacidad: datos personales y riesgos institucionales

El uso de plataformas de IA en contextos escolares plantea cuestiones importantes sobre la privacidad, especialmente si los estudiantes ingresan datos personales, fragmentos de evaluaciones, o información sensible. Organismos como (UNESCO, 2023a) y la OECD (2023) señalan que, sin prácticas adecuadas de protección de datos, puede comprometerse la seguridad de la información personal y académica.

Por ello, es fundamental orientar al estudiantado a:

- evitar ingresar datos privados, nombres completos o información sensible;
- no cargar trabajos íntegros;
- utilizar versiones educativas o institucionales cuando existan, ya que cuentan con protocolos de protección más estrictos (OECD, 2023).

▶• Acceso desigual: una nueva brecha educativa

El acceso diferenciado a dispositivos, conectividad, versiones *premium* de modelos o habilidades digitales crea nuevas brechas que impactan en la calidad del aprendizaje (UNESCO, 2023a; OECD, 2021). Esto implica que algunos estudiantes pueden beneficiarse ampliamente de la retroalimentación asistida por IA, mientras que otros quedan en desventaja.

El rol docente, por tanto, es mediar estas desigualdades mediante:

- acceso equitativo a herramientas aprobadas por la institución,
- capacitación básica para todos los estudiantes,
- diseño de actividades que no dependan exclusivamente de IAGen para su éxito.

6.2 Necesidad de supervisión y juicio docente

Las herramientas de IAGen ofrecen oportunidades significativas para mejorar la enseñanza y el aprendizaje, no obstante, el juicio profesional del docente continúa siendo insustituible. La supervisión humana es esencial para asegurar que el uso de la IA se mantenga alineado con los principios pedagógicos, éticos y disciplinarios que guían la educación superior (UNESCO, 2023b; Holmes, Bialik & Fadel, 2022). Esto en palabras de Rincón (2022, p. 41) "...es fundamental recordar que estos avances son herramientas que deben estar al servicio de los objetivos pedagógicos, por tanto, la enseñanza y la evaluación deben permanecer en el centro de atención".

Las herramientas de IAGen pueden sugerir ideas, identificar patrones o generar retroalimentación preliminar,

pero no tienen la capacidad de interpretar el contexto educativo, los objetivos del curso ni las particularidades del estudiantado. Como advierten Zawacki-Richter et al. (2019), la IA puede apoyar el análisis del desempeño, pero no reemplazar la comprensión profunda que el profesorado posee sobre los contenidos, las intencionalidades formativas y las necesidades de cada grupo. Las decisiones de evaluación, acreditación de aprendizajes o ajustes pedagógicos deben estar siempre bajo el juicio docente (UNESCO, 2023b).

De tal forma que la supervisión docente no solo implica revisar los productos generados por la IA, sino también guiar a los estudiantes en su uso responsable. La OECD (2023) advierte que, sin acompañamiento, el estudiantado puede depender excesivamente de estas herramientas para generar ideas, estructurar textos o resolver actividades, lo cual puede limitar el desarrollo de habilidades críticas, argumentativas y de autoría académica.

Por otro lado, la evaluación en educación superior no consiste únicamente en verificar si un producto está "bien escrito" o "correctamente estructurado"; implica valorar procesos cognitivos como la argumentación, la interpretación disciplinar y la originalidad. Estas dimensiones requieren una interpretación humana que las herramientas generativas no pueden realizar con validez educativa (Brookhart, 2018). En este sentido, el juicio docente actúa como una salvaguarda para que el uso de IAGen no distorsione la integridad académica, el aprendizaje auténtico y la equidad en la evaluación.

Lo recomendable es diseñar actividades en las que la IA sea un apoyo, pero no un sustituto del proceso intelectual del estudiante (UNESCO, 2023b; ISTE, 2023). Esto incluye solicitar análisis reflexivos, justificar decisiones tomadas con la ayuda de IAGen, o incorporar momentos de trabajo en clase donde la interacción humana prevalezca.

El docente, a través de su juicio profesional, determinará cuándo y cómo la IA agrega valor al proceso formativo, y cuándo puede comprometer la autenticidad del aprendizaje. "La Inteligencia Artificial (IA) en la evaluación no debe ser un fin en sí misma sino un poderoso aliado que impulsa la constante mejora de la educación y la obtención de logros educativos significativos' (Rincón, 2022 p. 45).

7. Referencias y otros recursos

Referencias

1. ACARA. (2023). *AI in Australian education: Literacy framework*. Australian Curriculum, Assessment and Reporting Authority.

2. ACTNext. (2019). *Advancing educational measurement with artificial intelligence* (ACT Research Report Series). ACT, Inc.

3. ANUIES. (2025). *Inteligencia artificial en la educación superior*. Asociación Nacional de Universidades e Instituciones de Educación Superior.

4. Avolio de Cols, S., & Lacolutti, M. (2006). *Enseñar y evaluar en formación por competencias laborales: Orientaciones conceptuales y metodológicas* Banco Interamericano de Desarrollo. https://agmerparana.com.ar/wp-content/uploads/2017/10/AVOLIO-DE-COLS-Susana-Ense%C3%B1ar-y-evaluar-en-formacion-por-c.pdf

5. Becker, B., Holmes, W., & Smith, S. (2023). *AI literacy for educators*. MIT Press.

6. Bender, E. M., Gebru, T., McMillan-Major, A., & Mitchell, M. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. https://doi.org/10.1145/3442188.3445922

7. Berman, R., & Estrada, P. (2021). Automated essay scoring: A review of the e-rater® system. *ETS Research Report Series, 2021*(1), 1–23. https://doi.org/10.1002/ets2.12314

8. Bolaño-García, M., & Duarte-Acosta, N. (2024). Una revisión sistemática del uso de la inteligencia artificial en la educación. *Revista Colombiana de Cirugía, 39*(1), 51–63. https://www.redalyc.org/journal/3555/355577357005/html/

9. Brookhart, S. M. (2018). *How to create and use rubrics for formative assessment and grading*. ASCD.

10. Caldeiro, G. (2024). Inteligencia artificial generativa y educación: Hacia un nuevo paradigma. *El Faro. Revista Digital de Docencia Universitaria, 1*(1), 22–43. https://revistaelfaro.uflo.edu.ar/index.php/elfaro/article/view/15

11. Calatayud S, M. A. (1999). *La creencia en la objetividad de la evaluación: Una ilusión posible*. Aula Abierta, (73), 205–221.

12. Campuzano-Vásquez, J., Murillo-Guevara, N., & Sarango-Pintado, D. (2025). Uso de la inteligencia artificial en la educación superior: Estudio de caso Universidad Técnica de Machala. *INNOVA Research Journal, 10*(2), 24–45.

13. Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education, 20*(1), 1–25. https://doi.org/10.1186/s41239-023-00408-3

14. Chao Rebolledo, C. (2025). La educación que viene: Inteligencia artificial y la reinvención radical de la educación superior. En ANUIES, *Inteligencia artificial en la educación superior* (pp. 194–201). ANUIES.

15. Chiappe, A., Sanmiguel, C., & Sáez Delgado, F. M. (2025). *IA generativa versus profesores: reflexiones desde una revisión de la literatura. Pixel-Bit. Revista de Medios y Educación*, (72), 119–137. https://doi.org/10.12795/pixelbit.107046

16. Code.org. (2023). *Code.org*. https://code.org/es-ES

17. Coursera. (2020). *Ensuring integrity in online assessments: AI-driven proctoring solutions*. Coursera for Campus.

18. Digital Education Council (DEC). (2025). *Global AI Faculty Survey 2025: AI meets academia—What faculty think*. https://www.digitaleducationcouncil.com/post/the-next-era-of-assessment-a-global-review-of-ai-in-assessment-design

19. Duolingo. (2021). *The Duolingo English Test: Technical Manual*. Duolingo Research. https://englishtest.duolingo.com/research

20. ESPOL. (2023). *IA generativa como entrenador metacognitivo*. Centro de Investigaciones y Servicios Educativos. https://www.cise.espol.edu.ec/es/ia-como-entrenador-metacognitivo

21. Ferrarelli, M. (2024). *Inteligencia artificial generativa y educación: dilemas éticos, ambientales, laborales y pedagógicos*. Revista Iberoamericana de Educación, 94(1), 1–15

22. Finland Ministry of Education. (2021). *National analytics platform for curriculum decision-making*.

23. Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

24. Gallent-Torres, C., Zapata-González, A., & Ortego-Hernando, J. L. (2023). El impacto de la inteligencia artificial generativa en educación superior: Una mirada desde la ética y la integridad académica. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa, 29*(2), 1–20. https://doi.org/10.30827/relieve.v29i2.29134

25. Gierl, M. J., & Lai, H. (2022). Using artificial intelligence to support test development: A review and research agenda. *Frontiers in Education, 7*, Artículo 852234. https://doi.org/10.3389/feduc.2022.852234

26. Gran EnciclopedIA. (2025). *La Gran EnciclopedIA: Recursos interactivos de IA*. https://raeia.org/books/la-gran-enciclopedia-recursos-interactivos/

27. Guerra, M. (2025). IA para la productividad docente: Herramientas multifuncionales. En *EDU BOOK: Inteligencia artificial generativa enfoques prácticos para docentes* (pp. 8–14). Instituto para el Futuro de la Educación. https://observatorio.tec.mx/edu-reads/edu-book-inteligencia-artificial-generativa-enfoques-practicos-para-docentes/

28. Guzmán Martínez, M. I., & Marín Uribe, R. E. (2020). Innovación educativa y su aplicación en los procesos de enseñanza-aprendizaje. *RIDE Revista Iberoamericana para la Investigación y el Desarrollo Educativo, 10*(20), 1–17. https://doi.org/10.23913/ride.v10i20.621

29. Hamodi, C., López Pastor, V., & López Pastor, A. (2015). Medios, técnicas e instrumentos de evaluación formativa y compartida del aprendizaje en educación superior. *Perfiles Educativos, 37*(147), 146–161.

30. Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

31. Educational Testing Service (ETS) & Khan Academy. (2022). *AI and the future of learning: Aligning competencies, instruction, and assessment*. ETS & Khan Academy. https://es.khanacademy.org/

32. Kyoto University. (2022). *AI-enhanced OSCE assessment project report*. Kyoto University School of Medicine.

33. León, G., & Apolo, D. (2025). Pedagogías emergentes y educación: Una revisión sistemática de estudios en Redalyc. *Revista Social Fronteriza, 5*(5).

34. Llanos Torrico, B. A. (2025). La evaluación educativa en la era de la inteligencia artificial. *Revista Científica, 12*(1), 73–88. https://doi.org/10.53287/undf7848pz65o

35. Luna de la Luz, V., & González-Flores, P. (2020). Transformaciones en educación médica: Innovaciones en la evaluación de los aprendizajes y avances tecnológicos (parte 2). *Investigación en Educación Médica, 9*(34), 87–99. https://doi.org/10.22201/facmed.20075057e.2020.34.20220

36. Montoya Ramírez, N. M., Andachi Chango, E. E., Defaz Toapanta, V. E., & Guilcapi Lunavictoria, D. O. (2024). Evaluación del aprendizaje en la era de la inteligencia artificial. *Pol. Con., 9*(3), 1977–1998. https://doi.org/10.23857/pc.v9i3.6770

37. Mora Naranjo, B. M., Aroca Izurieta, C. E., Tiban Leica, L. R., Sánchez Morrillo, C. F., & Jiménez Salazar, A. (2023). Ética y responsabilidad en la implementación de la inteligencia artificial en la educación. *Ciencia Latina Revista Científica Multidisciplinar, 7*(6), 2054–2076. https://doi.org/10.37811/cl_rcm.v7i6.8833

38. Moreno, T. (2016). *Evaluación del aprendizaje y para el aprendizaje: Reinventar la evaluación en el aula*. Universidad Autónoma Metropolitana.

39. Mujica-Sequera, R. (2024). Clasificación de las herramientas de la inteligencia artificial en la educación. *Revista Tecnológica-Educativa Docentes 2.0, 17*(1), 31–40. https://doi.org/10.37843/rted.v17i1.513

40. NBME. (2020). *Advances in test development and validation using AI and data analytics* (Technical Report). National Board of Medical Examiners.

41. NBME. (2022). *USMLE test evolution and AI applications* (Research Series). National Board of Medical Examiners.

42. NBME. (s. f.). *AI in assessment: Ethics, innovation and research* [Webinar]. https://www.nbme.org/events/on-demand/ai-in-assessment

43. OECD. (2021). *21st-century readers: Developing literacy skills in a digital world*. OECD Publishing.

44. OECD. (2023). *Artificial intelligence in education: Challenges and opportunities*. OECD Publishing.

45. Pearson VUE. (2021). *AI-supported item generation and validation: Efficiency gains report*

46. Pedraza, E. (2022). La analítica del aprendizaje y las expectativas de los estudiantes universitarios. *ACADEMO Revista de Investigación en Ciencias Sociales y Humanidades, 9*(2), 151–164.

47. Porter, A. (2025, 14 de julio). *Responsible AI governance: Frameworks, risks, and real-world benefits*. BigID Blog. https://bigid.com/es/blog/responsible-ai-governance/

48. ProctorU. (s.f.). *How online proctoring works.* https://www.proctoru.com

49. Quinde-Moncerrate, H. G., Quinde-Zambrano, L. F., & Franco-Arroyo, P. P. (2025). Incidencia de la inteligencia artificial en la educación contemporánea: Revisión sistemática. *Sociedad & Tecnología, 8*, 358–368. https://doi.org/10.51247/st.v8i2.505

50. Rincón Macías, M. A. (2022). La inteligencia artificial: ¿aliada o amenaza en la evaluación educativa? *Desarrollo Profesoral Universitario, 10*(1), 37–48. https://doi.org/10.26852/2357593X.691

51. Royal College of Physicians and Surgeons of Canada. (2024). *Artificial intelligence, medicine and education: Balancing innovation, opportunity and risk*. https://www.royalcollege.ca

52. Sánchez, M. (2022). Evaluación del, para y como aprendizaje. En M. Sánchez & A. Martínez (Eds.), *Evaluación y aprendizaje en educación universitaria: Estrategias e instrumentos* (pp. 17–35). CUAIEED, UNAM.

53. Santana-Soriano, E., & Baez-Vizcaino, K. (2025). Inteligencia artificial, gamificación y realidad virtual en la educación secundaria dominicana: Un análisis descriptivo. *Edutec. Revista Electrónica de Tecnología Educativa*, 196–215.

DOI: https://doi.org/10.21556/edutec.2025.92.3833

54. Santos, R., González-Flores, P., & Sánchez, M. (2022). *Glosario de innovación educativa*. CUAIEED, UNAM.

55. Sereño, F. (2024). La formación docente en la era digital: Práctica reflexiva, aprendizaje situado e inteligencia artificial. *Revista Ensayos Pedagógicos, 19*(2), 1–15. https://dialnet.unirioja.es/servlet/articulo?codigo=9846484

56. Shute, V., & Ventura, M. (2022). *Stealth assessment and AI: Integrating measurement into digital environments*. MIT Press.

57. Suárez Estavillo, U. (2025). La inteligencia artificial en la educación: ¿transformación o infoxicación? Un análisis crítico de la nueva frontera educativa. *Sintaxis, 14*, 69–88. https://doi.org/10.36105/stx.2025n14.05

58. Sung, Y.-T., Chang, K.-E., & Liu, T.-C. (2023). *The effects of integrating ChatGPT into teaching and learning: Opportunities and challenges for educational practice. Computers & Education: Artificial Intelligence*, 4, 100140.
https://doi.org/10.1016/j.caeai.2023.100140

59. López, A.T., & Pedraza, N. Y. (2017). La objetividad en las pruebas estandarizadas. *Revista Iberoamericana de Evaluación Educativa*, *10*(1), 11–31.
https://doi.org/10.15366/riee2017.10.1.001

60. UNAM (2025). *Recomendaciones para el uso educativo de la inteligencia artificial generativa en la UNAM*. Grupo Académico de Inteligencia Artificial Generativa en Educación. https://iagenedu.unam.mx/recomendaciones

61. UNESCO. (2023a). *Guía de la UNESCO para el uso de la IA generativa en la educación y la investigación*. UNESCO https://doi.org/10.54675/EWZM9535

62. UNESCO. (2023b). *Guía de uso de la inteligencia artificial en educación: Principios, oportunidades y orientaciones para docentes*. UNESCO Publishing.

63. UNESCO. (2025). *AI and the future of education: Disruptions, dilemmas and directions*. UNESCO. https://doi.org/10.54675/KECK1261

64. Universidad Internacional de La Rioja. (s. f.). *Proyecto de innovación educativa*. Revista UNIR Educación.

65. USMLE. (2023). *Data-informed evolution of exam content* (Annual Report). United States Medical Licensing Examination.

66. Vivas Urías, M. D., & Ruiz Rosillo, M. A. (2025). *Inteligencia artificial generativa. Buenas prácticas docentes en educación superior*. Octaedro / Horizontes-Universidad. https://www.researchgate.net/publication/391700547_Inteligencia_artificial_generativa_Buenas_practicas_docentes_en_educacion_superior

67. Walker, P. B., Haase, J. J., Mehalick, M. L., Steele, C. T., Russell, D. W., & Davidson, I. N. (2025). Harnessing metacognition for safe and responsible AI. *Technologies, 13*(3), 107. https://doi.org/10.3390/technologies13030107

68. Wargo, K., & Anderson, B. (2024, 5 de diciembre). *Striking a balance: Navigating the ethical dilemmas of AI in higher education*. EDUCAUSE Review. https://er.educause.edu/articles/2024/12/striking-a-balance-navigating-the-ethical-dilemmas-of-ai-in-higher-education

69. Weidinger, L., Uesato, J., Rauh, M., Griffin, C., Huang, P., Mellor, J., Glaese, A., Cheng, M., Balle, B., Kasirzadeh, A., Biles, C., Brown, S., Kenton, Z., Hawkins, W., Stepleton, T., Birhane, A., Hendricks, L. A., Rimell, L., Isaac, W., … Gabriel, I. (2022). Taxonomy of risks posed by language models. En *FAccT '22: Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 214–229). Association for Computing Machinery. https://doi.org/10.1145/3531146.3533088

70. Zakrajsek, T. (2025, 10 de marzo). *Teaching students AI strategies to enhance metacognitive processing*. The Scholarly Teacher. https://www.scholarlyteacher.com/post/teaching-students-ai-strategies-to-enhance-metacognitive-processing

71. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education, 16*(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0

Otros recursos

1. Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university* (4.ª ed.). Open University Press.

2. Boyle, P. (2025, 27 de febrero). *AI in medical education: 5 ways schools are employing new tools*. AAMCNews. https://www.aamc.org/news/ai-medical-education-5-ways-schools-are-employing-new-tools

3. Fatima, S. S., Sheikh, N. A., & Osama, A. (2024). Authentic assessment in medical ducation: Exploring AI integration and student-as-partners collaboration. *Postgraduate Medical Journal, 100*(1190), 959–967. https://doi.org/10.1093/postmj/qgae088

4. Floridi, L. (2021). *Ethics of artificial intelligence*. Oxford University Press.

5. Harden, R. M. (2016). Construct educational objectives for teaching and learning. *Medical Teacher, 38*(6), 593–598. https://doi.org/10.3109/0142159X.2016.1173661

6. Hicke, Y., Geathers, J., Rajashekar, N., Chan, C., Jack, A. G., Sewell, J., Preston, M., Cornes, S., Shung, D., & Kizilcec, R. (2025, 1 de marzo). *MedSimAI: Simulation and formative feedback generation to enhance deliberate practice in medical education* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2503.05793

7. Hodges, B. (2024, 19 de noviembre). *Artificial intelligence, medicine and education: Balancing innovation, opportunity and risk*. Royal College of Physicians and Surgeons of Canada. https://www.royalcollege.ca/en/news/artificial-intelligence--medicine-and-education--balancing-innov.html

8. ISTE. (2023). *ISTE guidance for educators on artificial intelligence in learning*. International Society for Technology in Education.

9. Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys, 55*(12), 1–38. https://doi.org/10.1145/3571730

10. Kaya, M., Sonmez, E., Halici, A., Yildirim, H., & Coskun, A. (2025). Comparison of AI-generated and clinician-designed multiple-choice questions in emergency medicine exam: A psychometric analysis. *BMC Medical Education, 25*(949). https://doi.org/10.1186/s12909-025-07528-6

11. Law, A. K. K., So, J., Lui, C. T., Choi, Y. F., Cheung, K. H., Hung, K. K., & Graham, C. A. (2025). AI versus human-generated multiple-choice questions for medical education: A cohort study in a high-stakes examination. *BMC Medical Education, 25*(208). https://doi.org/10.1186/s12909-025-06796-6

12. Lucchese, M. S. M., Lizzio, S., Calneggia, M. I., Güizzo, M. J., Tomatis, M. C., Gandini, B., & Novella, M. de L. (2023). Evaluación en especializaciones médicas: Un estudio en una universidad argentina. *Educación Médica, 24*(2). https://doi.org/10.1016/j.edumed.2023.100796

13. Masters, K., MacNeil, H., Benjamin, J., Carver, T., Nemethy, K., Valanci-Aroesty, S., Taylor, D. C. M., Thoma, B., & Thesen, T. (2025). Artificial Intelligence in Health Professions Education assessment: AMEE Guide No. 178. *Medical Teacher, 47*(9), 1410–1424. https://doi.org/10.1080/0142159X.2024.2445037

14. Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura (OEI). (2024). *Inteligencia artificial y educación: Insumos para su abordaje desde Iberoamérica*. OEI

15. Ramineni, C., & Williamson, D. M. (2018). Understanding mean score differences between the e-rater automated scoring engine and humans for demographically based groups in the GRE General Test. *Educational Testing Service*. https://doi.org/10.1002/ets2.12192

16. UNAM. (2019). *Modelo educativo para la UNAM*. UNAM

Universidad Nacional Autónoma de México

Dr. Leonardo Lomelí Vanegas
Rector

Dra. Patricia Dolores Dávila Aranda
Secretaria General

Mtro. Hugo Concha Cantú
Abogado General

Mtro. Tomás Humberto Rubio Pérez
Secretario Administrativo

Dra. Diana Tamara Martínez Ruiz
Secretaria de Desarrollo Institucional

M.I. Fernando Macedo Chagolla
Secretario de Servicio y Atención a la Comunidad Universitaria

Dr. Manuel Palma Rangel
Secretario de Prevención y Apoyo a la Movilidad y Seguridad Universitaria

Dra. Norma Blazquez Graf
Coordinadora para la Igualdad de Género

Autoras y autores

Integrantes de la Red Colaborativa de la Comisión Permanente de Licenciatura, Consejo de Evaluación Educativa

Licenciatura

Mtro. Antonio Cerritos
Facultad de Medicina

Mtro. Reyes Farfán Sánchez
Escuela Nacional de Ciencias Forenses

Lic. Yazmín Alejandra Lara Gutiérrez
Facultad de Enfermería y Obstetricia

Mtra. Sandra Paola Muñoz García
Escuela Nacional de Estudios Superiores León

Dra. Liliana Valladares Riveroll
Facultad de Filosofía y Letras

Coordinación de Evaluación, Innovación y Desarrollo Educativos

Dr. Melchor Sánchez Mendiola
Coordinador

Lic. Enrique Ricardo Buzo Casanova
Director

Mtra. Elibidú Ortega Sánchez
Subdirectora

Lic. Lourdes Rocío Muñoz Alonso

Mtra. Adriana Durán Luna

Mtro. José Alberto Vargas Castillo

Lic. Selene Cruz Juárez

SECRETARÍA GENERAL
Universidad Nacional Autónoma de México

CEIDE
Coordinación de Evaluación, Innovación y Desarrollo Educativos

CONSEJO DE EVALUACIÓN EDUCATIVA DE LA UNAM