

Rubrics and Generative AI: Towards Process-Based Assessment.

Sandra Oviedo^{1,2} [0009-0004-1172-6955], Daniel Díaz^{1,2} [0009-0000-1671-5826],

Alejandra Otazu² [0009-0000-3418-0714], Juan Cúneo² [0009-0006-6728-5591],

María Becerra² [0009-0002-8913-9107],

¹ Instituto de Informática, ² Departamento de Informática,
Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de San Juan
San Juan, Argentina
{soviedo, ddiaz}@iinfo.unsj.edu.ar; {aotazu, jcuneo,
mbecerra}@unsj-cuim.edu.ar

Abstract. The emergence of Generative Artificial Intelligence (GAI) in higher education has triggered a crisis in traditional assessment models, whose validity historically focused on the final product. Through the use of GAI, the outcome of a teaching-learning task is, in many cases, guaranteed in terms of quality, given the ability of algorithms to automate complex intellectual tasks. This paper proposes a paradigm shift toward a process-based pedagogy. The objective is to present a set of guidelines for designing assessment rubrics that capture the cognitive interaction between the student and technology. Based on Biggs' 3P Model, the proposal identifies three critical dimensions for grading: human agency and critical supervision, dialogic scaffolding (linked to prompt engineering), and cognitive sovereignty based on integrity and transparency. Theoretical results suggest that by evaluating the student's journey and curation capacity rather than just the static outcome, it is possible to mitigate the risk of cognitive externalization. It is concluded that next-generation rubrics must act as navigation maps that foster a productive human-AI symbiosis, ensuring that technology enhances the development of high-level professional competencies and secures academic honesty in the new university digital ecosystem.

Keywords: Generative AI, Assessment Rubrics, 3P Model.

Rúbricas e IA Generativa: Hacia una Evaluación por Procesos.

Resumen. La irrupción de la Inteligencia Artificial Generativa (IAG) en la educación superior ha generado una crisis en los modelos de evaluación tradicionales, cuya validez se centraba históricamente en el producto final. Mediante el uso de IAG el resultado de una tarea del proceso de enseñanza-aprendizaje, en muchos casos, está asegurado en calidad, ante la capacidad de

los algoritmos para automatizar tareas intelectuales complejas. Este artículo propone un cambio de paradigma hacia una pedagogía del proceso. El objetivo es presentar un conjunto de pautas para el diseño de rúbricas de evaluación que permitan capturar la interacción cognitiva entre el estudiante y la tecnología. Fundamentada en el Modelo 3P de Biggs, la propuesta identifica tres dimensiones críticas para la calificación: la agencia humana y supervisión crítica, el andamiaje dialógico (vinculado a la ingeniería de prompts) y la soberanía cognitiva basada en la integridad y transparencia. Los resultados teóricos sugieren que, al evaluar el trayecto y la capacidad de curaduría del alumno en lugar de solo el resultado estático, es posible mitigar el riesgo de externalización cognitiva. Se concluye que las rúbricas de nueva generación deben actuar como mapas de navegación que fomenten una simbiosis productiva humano-IA, garantizando que la tecnología potencie el desarrollo de competencias profesionales de alto nivel y asegure la honestidad académica en el nuevo ecosistema digital universitario.

Palabras clave: IA Generativa, Rúbricas de Evaluación, Modelo 3P.

1 Introducción

La irrupción masiva de la Inteligencia Artificial Generativa (IAG) ha marcado un punto de inflexión sin precedentes en la Educación Superior. A diferencia de otras innovaciones tecnológicas que optimizaron la búsqueda o el almacenamiento de información, la IAG —y específicamente los Modelos de Lenguaje de Gran Escala (LLM)— posee la capacidad de realizar tareas que hasta hace poco se consideraban exclusivas del intelecto humano: redactar ensayos, resolver problemas matemáticos complejos, generar código de programación y sintetizar literatura científica. Esta capacidad de producción automatizada ha puesto en jaque los sistemas de evaluación tradicionales, los cuales, históricamente, han centrado su validez en el producto final entregado por el estudiante.

En este nuevo ecosistema digital, la evaluación basada únicamente en resultados estáticos resulta insuficiente y vulnerable. El riesgo de la "externalización cognitiva" es latente; si el sistema educativo continúa premiando el producto sin considerar el trayecto, se corre el peligro de erosionar la autonomía intelectual del aprendiz. Por lo tanto, surge la necesidad imperativa de transitar desde una pedagogía del resultado hacia una pedagogía del proceso, donde el foco de la calificación se desplace hacia la interacción, la toma de decisiones y el pensamiento crítico.

Este artículo propone un marco de pautas para el diseño de rúbricas de evaluación adaptadas a entornos mediados por IAG. La propuesta se fundamenta teóricamente en el Modelo 3P de Biggs, que permite analizar el aprendizaje como un sistema integrado de factores de Presagio, Proceso y Producto. Bajo esta óptica, la rúbrica deja de ser un mero instrumento de control para transformarse en un mapa de navegación cognitiva.

El objetivo central de este trabajo es definir dimensiones de evaluación que capturen la agencia humana del estudiante. Se introducen conceptos clave como el andamiaje dialógico, que describe la interacción iterativa con la IA, y la soberanía cognitiva, que refiere a la capacidad del alumno para mantener el mando intelectual y la

responsabilidad ética sobre los contenidos generados. A través de estas pautas, se busca empoderar a los docentes con herramientas que permitan distinguir entre el uso sustitutivo de la tecnología y una simbiosis productiva que potencie el desarrollo de competencias profesionales de alto nivel. Se trata de una investigación en curso.

2 Marco teórico

2.1 Inteligencia Artificial Generativa

La Inteligencia Artificial Generativa (IAG) representa una rama de la inteligencia artificial capaz de producir contenido nuevo y original, como texto, imágenes, audio, video e incluso código, a partir de los datos con los que fue entrenada. A diferencia de las IA que solo analizan o clasifican información existente, la IAG "crea" algo desde cero, haciendo que parezca generado por un humano. La IA Generativa no es solo una herramienta para automatizar, sino un motor para la creatividad y la innovación. Sin embargo, su rápido avance también plantea desafíos importantes en términos de ética, originalidad y su impacto en el empleo y las habilidades humanas (Sengar, Hasan, Kumar, & Carroll, 2025).

Los Modelos de Lenguaje Grandes (LLM) son aplicaciones específicas que hacen posible el procesamiento de lenguaje natural. Estos se apoyan en una arquitectura particular llamada Transformadores la que permite a los LLM procesar y relacionar grandes cantidades de información textual de manera muy eficiente, capturando el contexto y las sutilezas del lenguaje. Sin los Transformadores, los LLM no serían tan potentes (Alier, García-Peñalvo, & Camba, 2024), (Blank, 2023).

Además, la IAG también crea otros tipos de contenido. Aquí es donde entran los Modelos de Difusión y las Redes Generativas Antagónicas. Estos son métodos o enfoques diferentes dentro de la IAG, específicamente diseñados para generar imágenes, audio o video, en lugar de texto (Sengar et al., 2025). Por ejemplo, DALL-E y Midjourney usan Modelos de Difusión para convertir texto en imágenes.

Más allá del tipo de modelo o método de IAG se use (ya sea un LLM para texto o un Modelo de Difusión para imágenes), la forma en que se interactúa con ellos es preponderante. Aquí es donde el "Prompt Engineering" se vuelve vital. La ingeniería de prompt es la habilidad de escribir las instrucciones o preguntas correctas (los "prompts") para guiar a la IAG a generar el resultado específico y preciso que se requiere (Knoth, Tolzin, Janson, & Leimeister, 2024).

2.2 El Modelo Presagio-Proceso-Producto (Modelo 3P)

El Modelo Presagio-Proceso-Producto, o Modelo 3P, desarrollado por John Biggs, ofrece un marco para comprender la perspectiva del estudiante en el proceso de enseñanza-aprendizaje. Este enfoque sistémico se compone de tres factores principales interconectados: Presagio, Proceso y Producto, en (Ka Yuk Chan & Hu, 2023), (Biggs, 1999) y (Sarzoza Herrera, 2013) se describen las variables a considerar en cada una de las tres etapas.

Factores de Presagio. Los factores de presagio son elementos previos al aprendizaje que influyen en cómo el estudiante aborda una tarea.

El Proceso de Aprendizaje. Cada enfoque de aprendizaje se distingue por un motivo (por qué el alumno desea aprender) y una estrategia (cómo actúa durante el aprendizaje). El proceso de aprendizaje describe cómo los estudiantes abordan y dan sentido al contexto de enseñanza, considerando sus conocimientos previos, motivaciones personales y la naturaleza de la tarea. La combinación de motivo y estrategia define los tres enfoques de aprendizaje propuestos por Biggs: Enfoque profundo, Enfoque superficial o de reproducción y Enfoque de alto rendimiento o logro.

El Producto del aprendizaje. El producto del aprendizaje se refiere a los resultados y consecuencias del proceso. Se puede evaluar de varias maneras: Cuantitativamente, Cualitativamente e Institucionalmente. Además, se consideran los resultados afectivos, que describen cómo los estudiantes se sienten respecto a su propio aprendizaje.

2.3 Rúbricas de evaluación

Una rúbrica es un marco de evaluación esencial en la educación superior que alinea la enseñanza, el aprendizaje y la evaluación para lograr los resultados educativos deseados. Una rúbrica es una herramienta de evaluación sistemática que describe criterios específicos y niveles de desempeño para medir los logros y habilidades de los estudiantes (Okafor, 2024), (Taylor, Kisby, & Reedy, 2024). Se definen fundamentalmente como documentos que articulan las expectativas de una tarea mediante el listado de lo que es importante y la descripción de los niveles de calidad en una escala que va desde lo excelente hasta lo insuficiente .

Características y funciones principales de una rúbrica. Se describen componentes y condiciones que deben cumplir las rúbricas de evaluación, según (Okafor, 2024) .

- **Componentes estructurales:** Generalmente consta de tres elementos clave: criterios de evaluación, una estrategia de calificación y definiciones de estándares o niveles de calidad. Estos niveles permiten desglosar habilidades complejas en componentes medibles.
- **Transparencia y Claridad:** Su propósito es ofrecer una guía clara tanto para educadores como para alumnos, definiendo qué se espera en cada nivel de logro. Esto ayuda a reducir la ansiedad de los estudiantes al proporcionarles los parámetros de éxito antes de comenzar una actividad.
- **Consistencia y Equidad:** Las rúbricas funcionan como un marco de referencia que asegura la objetividad, consistencia y justicia en la calificación, especialmente cuando hay múltiples evaluadores para una misma tarea.
- **Herramienta de aprendizaje:** Más allá de la calificación, las rúbricas fomentan la autorreflexión y el aprendizaje activo, ya que permiten a los estudiantes identificar sus fortalezas y áreas de mejora. También sirven como herramientas de planificación y redacción durante el proceso de elaboración de un trabajo.

Tipos y aplicaciones de rúbricas. Las rúbricas se pueden clasificar de diversas maneras según su estructura, su nivel de especificidad y su aplicación tecnológica. A continuación, se explican los tipos de rúbricas de mayor interés.

Rúbricas Analíticas: Se destacan por su capacidad para ofrecer retroalimentación detallada y constructiva (Okafor, 2024), (Oroh, Ali, Pelenkahu, Usman2, & Rorintulus, 2025).

Características: Desglosan habilidades complejas (como la escritura o el pensamiento crítico) en componentes medibles o criterios específicos.

- Utilidad: Permiten a los estudiantes identificar con precisión sus fortalezas y áreas de mejora, guiándolos en el perfeccionamiento de su desempeño. Son muy utilizadas para evaluar portafolios y tareas de escritura extensas.

Rúbricas Genéricas vs. Rúbricas Específicas. Las rúbricas se pueden clasificar según el alcance de su aplicación en los módulos académicos (Taylor et al., 2024) Mientras que las rúbricas analíticas son fundamentales para el aprendizaje formativo por su detalle, la elección entre una rúbrica genérica o específica depende de si se busca estandarización institucional o una guía precisa para el estudiante en una tarea determinada.

- Rúbricas Genéricas: Son marcos de evaluación que se aplican de manera uniforme en diferentes módulos o asignaturas. Aunque promueven la consistencia, algunos estudiantes las perciben como "básicas" o menos informativas.
- Rúbricas Específicas o Personalizadas: Están alineadas a una tarea o módulo particular. Los estudiantes suelen valorarlas más porque son más informativas y detalladas respecto a lo que se espera en un trabajo concreto.

Rúbricas según la Tarea Evaluada. Las rúbricas también se categorizan por el tipo de producto académico que miden (Oroh et al., 2025):

- Rúbricas de Borradores: Utilizadas para evaluar las versiones preliminares de un trabajo, enfocándose en la estructura del esquema, la introducción y los puntos principales.
- Rúbricas de Escritura Expositiva: Específicamente diseñadas para ensayos de solución de problemas, donde se evalúa la identificación del problema, el análisis de causa-efecto y la propuesta de soluciones lógicas.
- Rúbricas de Evaluación de Pensamiento Crítico: Estas rúbricas incluyen criterios como el análisis, el razonamiento lógico, la síntesis de ideas y la profundidad de la comprensión (Okafor, 2024).

Rúbricas en el Contexto de la Inteligencia Artificial (IA). Con la transformación digital, han surgido nuevas formas de interactuar con estas herramientas (Ilieva, Yankova, Ruseva, & Kabaivanov, 2025).

- Rúbricas Generadas por IA: Herramientas como ChatGPT o Eduaide.ai permiten a los instructores crear borradores de rúbricas alineados con objetivos de aprendizaje o taxonomías como la de Bloom.

- Rúbricas para Calificación Automatizada: Sistemas como Gradescope emplean rúbricas para estandarizar la evaluación y agrupar respuestas similares, asegurando la consistencia en clases con gran número de alumnos.

3 Hacia una evaluación centrada en el proceso

3.1 La percepción de los estudiantes.

En (Oviedo et al., 2025) se presenta un estudio acerca de la percepción de los estudiantes respecto a la integración de la IAG en el proceso de aprendizaje, dicho estudio se organizó según las etapas del modelo 3P y los hallazgos más destacados se exponen a continuación.

Etapas de Presagio: Motivación y Fiabilidad. Los datos revelan una actitud mayoritariamente positiva hacia el potencial de la IAG en la fase inicial del aprendizaje. *Motivación y Eficiencia:* Cerca del 60% de los estudiantes percibe la IAG como un factor motivador para iniciar proyectos académicos. Además, un 72% cree que usar IAG ahorra tiempo y esfuerzo en la investigación preliminar.

El Desafío de la Fiabilidad: Sin embargo, la desconfianza es alta. Un 48% desconfía de la precisión y fiabilidad de la información inicial proporcionada por la IAG. En palabras de los estudiantes: “Me ayuda a entender temas difíciles, pero no siempre la información es correcta”. Esto subraya la necesidad de un pensamiento crítico robusto desde el inicio del proceso.

Etapas de Proceso: Profundización vs. Superficialidad. La IAG se consolida como una herramienta de productividad y apoyo cognitivo. *Apoyo y Productividad:* Aproximadamente el 70% usa la IAG activamente para la generación de ideas, esquemas o resúmenes. *Enfoque Profundo:* Un alto porcentaje, el 92%, indica que la IAG les ayuda a profundizar en los temas. Esto sugiere que la usan para obtener explicaciones alternativas, facilitando así un enfoque profundo de aprendizaje. *La Tensión Ética y Cognitiva:* Esta ayuda viene con riesgos: Un 44% de los estudiantes se preocupa porque el uso frecuente de IAG disminuya sus habilidades de pensamiento crítico o escritura original.

El riesgo del enfoque superficial es real, 28% reconoce que, al usar la IAG, a veces se centra más en cumplir con la tarea que en comprender realmente el contenido. Un dato preocupante es que el 42% ha utilizado la IAG para generar respuestas o secciones completas de trabajos.

Etapas de Producto: Autoría y Orgullo Afectivo. En la etapa final de producto, la IAG mejora los resultados formales, pero genera dilemas éticos y emocionales. *Calidad Formal:* El 68% (sumando "De acuerdo" y "Totalmente de acuerdo") percibe que la IAG ayuda a mejorar la calidad formal de los trabajos (redacción, estructura). *El Miedo a la Indistinción:* Un 46% teme que sus docentes no puedan diferenciar entre un trabajo original y uno asistido por IAG. *La Dimensión Afectiva:* La ambivalencia emocional es palpable: el 42% de los estudiantes se sentiría menos orgulloso de un trabajo asistido por IA que de uno hecho solo. *El Dilema de la Autoría:* “La IA mejora mis informes, pero ya no sé si el texto es mío”.

La conciencia ética existe, pues cerca del 49% reconoce el plagio de trabajos enteramente generados por IAG como un problema real.

A partir de los resultados de la encuesta, se puede generalizar que los estudiantes encuentran en el uso de la IAG una fuente de motivación y coraje para iniciarse en desafíos de nuevos aprendizajes. Asimismo, se pone en evidencia la necesidad de replantear las formas de evaluación a fin de que la misma resulte justa, transparente y reduzca temores de invisibilidad, desvalorización y dilemas éticos.

3.2 Rúbricas de evaluación en las etapas del Modelo 3P del proceso de aprendizaje mediado por IAG.

En un contexto de proceso de aprendizaje mediado por IAG, la rúbrica debe cumplir funciones claves en las etapas del Modelo 3P.

En la etapa Presagio, las rúbricas deben dar claridad al estudiante sobre qué se espera de él antes de empezar. Reducir la ansiedad y establecer las "reglas del juego" respecto a la ética y la IA.

En la etapa de Proceso, las rúbricas deben actuar como un andamiaje. El estudiante debería poder autoevaluarse mientras interactúa con la IA, verificando si está cumpliendo con la "curaduría de contenidos" o si solo está copiando.

Finalmente, en la etapa de Producto, las rúbricas deberían permitir al docente calificar de manera justa, valorando el valor agregado humano por encima del resultado generado por el algoritmo.

Las rúbricas tradicionales (que solo miran el resultado final) son vulnerables ante la IAG. Las nuevas rúbricas deben capturar lo que sucede en la fase de Proceso. La rúbrica deja de ser una simple lista de cotejo para convertirse en una hoja de ruta que guía la agencia humana del alumno.

3.3. Una nueva generación de Rúbricas de Evaluación. Dimensiones.

La IA puede resolver el producto final fácilmente, esto rompe las rúbricas tradicionales, dando lugar a una rúbrica de nueva generación, de manera tal que se vea la rúbrica como el contrato pedagógico que asegure que el estudiante mantenga su soberanía cognitiva mientras aprovecha los beneficios de la inteligencia artificial. En este sentido, la rúbrica debe evaluar:

1. **La trazabilidad:** ¿Cómo llegó el estudiante a ese resultado?
2. **La reflexión:** ¿Qué cambios hizo el estudiante a lo que entregó la IA?
3. **La autoría:** ¿Se percibe aporte personal y la agencia humana en el trabajo?

Para que la rúbrica sea efectiva en la era de la IA, al definir una rúbrica de evaluación, se deberá tener en cuenta las siguientes dimensiones:

Dimensión de "Agencia Humana" y Supervisión Crítica. Dimensión de "Agencia Humana" y Supervisión Crítica es el eje que garantiza que el estudiante no sea un espectador pasivo, sino el verdadero autor y responsable de su aprendizaje. A diferencia de una rúbrica tradicional que solo corregiría si el resultado es "correcto", esta

dimensión evalúa cómo el estudiante se posicionó frente a la máquina. Tiene dos componentes:

La Agencia Humana. El estudiante al mando. La agencia es la capacidad de actuar con intención. En una rúbrica, evaluar la agencia significa medir si el estudiante:

- Mantiene el control: El alumno decide qué partes del trabajo resuelve la IA y qué partes resuelve él, basándose en sus propios objetivos de aprendizaje.
- Aporta valor agregado: No se limita a copiar lo que la IA entrega, sino que añade su propia voz, contexto local, o juicios de valor que un algoritmo no puede generar.
- Toma decisiones: Si la IA sugiere un camino, el estudiante tiene la potestad de rechazarlo si no se ajusta al rigor académico esperado.

La Supervisión Crítica. El estudiante como curador. Este es el ejercicio de "desconfianza operativa". Evalúa la capacidad de aplicar pensamiento crítico, mediante:

- Detección de alucinaciones: Identificar cuando la IA inventa datos o referencias bibliográficas (un problema común en los LLM).
- Validación de fuentes: El estudiante debe contrastar lo que dice la IA con material bibliográfico oficial (en la encuesta, el 79.8% de los alumnos ya hace esto; la rúbrica ahora le daría puntaje por ese esfuerzo).
- Identificación de sesgos: Reconocer si la respuesta de la IA tiene una visión sesgada o limitada y corregirla.

Esta dimensión es la respuesta al temor de la "dependencia cognitiva". Si el docente evalúa la supervisión crítica, está obligando al alumno a pensar. En términos del Modelo 3P, esta dimensión pertenece a la fase de Proceso, asegurando que el estudiante mantenga su Soberanía Cognitiva. En lugar de calificar si la respuesta es correcta, calificar la capacidad de validación.

Para evaluar esta dimensión específica, se podrían proponer estos descriptores en la rúbrica:

- Nivel Insuficiente: El estudiante entrega el texto de la IA sin cambios. No hay evidencia de verificación de datos ni de una voz personal. La agencia es nula.
- Nivel En Desarrollo: El estudiante edita algunas partes y corrige errores ortográficos de la IA, pero acepta la estructura y los argumentos sin cuestionar su veracidad o profundidad.
- Nivel Competente: El estudiante declara el uso de la IA, identifica al menos un punto donde la IA fue imprecisa y aporta una conclusión personal que complementa lo generado.
- Nivel Excelente (Soberanía Cognitiva): El estudiante demuestra una curaduría total. Presenta el registro de cómo cuestionó a la IA (prompts de refinamiento), valida

cada dato con fuentes externas y la producción final refleja una síntesis original donde la IA fue solo un soporte.

Dimensión de "Andamiaje Dialógico" (Prompt Engineering). La Dimensión de "Andamiaje Dialógico", centrada en el *Prompt Engineering*, es el componente de la rúbrica que evalúa la calidad de la interacción técnica y cognitiva entre el estudiante y la inteligencia artificial.

En lugar de calificar solo el resultado, esta dimensión pone la lupa en la capacidad del estudiante para "dialogar" con la máquina, transformando una simple instrucción en un proceso de pensamiento guiado.

Andamiaje Dialógico. El concepto de andamiaje (procedente de la psicología educativa de Vygotsky) se refiere al apoyo temporal que se le da a un alumno para alcanzar un nivel de aprendizaje que no lograría solo. La IA no hace el trabajo por el alumno, sino que sirve de estructura para que el alumno suba de nivel. A diferencia de una búsqueda en Google, la IAG permite una conversación de ida y vuelta. El "Andamiaje Dialógico" mide la capacidad del estudiante para no quedarse con la primera respuesta.

La rúbrica evalúa si el estudiante usa la IA para superar un bloqueo creativo o para profundizar en un tema, manteniendo siempre la guía del proceso. También se evalúa: El refinamiento. ¿Hubo repreguntas? ¿Hubo ajustes de tono o de parámetros? ¿El estudiante guió a la IA mediante un contexto claro?

El Prompt Engineering como habilidad cognitiva. En esta dimensión, el *prompting* no es visto como un truco técnico, sino como la manifestación del pensamiento lógico del estudiante. Para escribir un buen *prompt*, el alumno debe tener claro:

- Contexto: Quién es la IA (ej. "Actúa como un experto en enfermería clínica").
- Tarea: Qué debe hacer exactamente.
- Restricciones: Qué no debe incluir o qué formato debe seguir.
- Iteración: Cómo mejorar la respuesta anterior.

Esta dimensión pertenece puramente a la fase de Proceso. Es la evidencia tangible de lo que el estudiante "hace" durante el aprendizaje. Si un estudiante tiene un buen desempeño en esta dimensión, es una señal clara de que está ejerciendo su Agencia Humana, ya que no está siendo un usuario pasivo, sino un "arquitecto" de la información.

Niveles de desempeño sugeridos para la Rúbrica. Se pueden proponer estos niveles para evaluar el *Andamiaje Dialógico*:

- Nivel Inicial (Uso Lineal): El estudiante introduce *prompts* breves y genéricos (ej. "Haz un resumen de X"). Acepta la primera respuesta sin iterar. El andamiaje es pobre porque la IA domina la estructura.
- Nivel En Desarrollo (Interacción Básica): El estudiante da instrucciones más claras y pide ajustes (ej. "Hazlo más corto" o "Usa lenguaje técnico"). Hay un intento de diálogo, pero es superficial.

- Nivel Competente (Diseño de Instrucción): El estudiante utiliza técnicas de *prompting* (contexto, rol, ejemplos). Se observa una secuencia de interacción donde el alumno guía a la IA hacia un objetivo específico.
- Nivel Excelente (Simbiosis Cognitiva): El estudiante demuestra un refinamiento complejo. Utiliza la IA para explorar diferentes ángulos de un problema, corregir sus propios sesgos y profundizar en la complejidad. El *prompting* refleja una comprensión total del tema.

Dimensión de Integridad y Transparencia (Soberanía Cognitiva). La Dimensión de Integridad y Transparencia, bajo el enfoque de la Soberanía Cognitiva, es el componente ético y procedimental de la rúbrica. Su objetivo no es solo evitar el plagio, sino garantizar que el estudiante mantenga el "mando intelectual" sobre su trabajo, siendo honesto sobre qué partes fueron delegadas a la tecnología y por qué. Esta dimensión representa el paso de una ética prohibitiva (castigar el uso) a una ética de la responsabilidad (validar el uso transparente). Se detallan los componentes de esta dimensión:

La Soberanía Cognitiva: El concepto central. La soberanía cognitiva es el derecho y la capacidad del estudiante de seguir siendo el dueño de su pensamiento. En la era de la IAG, esto se ve amenazado por la "externalización del razonamiento". En la rúbrica se evalúa: si el estudiante utiliza la IA para potenciar su cerebro (como una prótesis intelectual) o si ha permitido que la IA sustituya su propio juicio. Un estudiante soberano es aquel que puede explicar y defender cada idea de su trabajo, incluso si la IA ayudó a redactarla.

Integridad Académica Proactiva. Ya no se trata solo de "no copiar". La integridad ahora implica la trazabilidad. En la rúbrica se evalúa la declaración explícita de qué herramientas se usaron, qué prompts se emplearon y en qué secciones hubo asistencia algorítmica. La falta de transparencia se califica como una pérdida de integridad, no necesariamente por usar la IA, sino por intentar ocultar su mediación.

Transparencia y Atribución. Es la capacidad técnica de citar correctamente la interacción con la IA. En la rúbrica se evalúa si el estudiante sigue normas de citación (como APA o IEEE adaptadas a IA) y si es capaz de adjuntar un anexo con el historial de la conversación con el bot para demostrar la evolución del pensamiento.

Al incluir la "Soberanía Cognitiva" en la rúbrica, se transforma la evaluación en un acto de metacognición: el alumno debe pensar sobre cómo pensó junto a la IA. Esto asegura que la institución siga certificando a un profesional capaz de pensar por sí mismo, cumpliendo con la "agencia humana", vista como motor del proceso.

Niveles de Desempeño Sugeridos para esta dimensión. Se podría proponer los siguientes descriptores:

- Nivel Insuficiente (Opacidad): El estudiante presenta el trabajo como producción puramente humana ocultando el uso de IAG. No hay rastro de soberanía cognitiva; el contenido es una "caja negra" donde no se distingue la autoría.
- Nivel En Desarrollo (Transparencia Parcial): El estudiante declara haber usado IA de forma genérica, pero no especifica qué partes fueron generadas y cuáles son propias. La soberanía es compartida pero difusa.
- Nivel Competente (Transparencia Activa): El estudiante cita las secciones asistidas por IA y describe brevemente el rol de la herramienta. Demuestra que el núcleo argumental y las conclusiones finales son de su propia autoría.
- Nivel Excelente (Soberanía Cognitiva Plena): El estudiante presenta un informe de transparencia detallado (anexo de prompts). Justifica éticamente por qué utilizó la IA en ciertos puntos (ej. para resumir fuentes o estructurar datos) y demuestra un dominio total sobre el producto final, evidenciando que la IA fue un subordinado de su intelecto.

Esta dimensión conecta directamente con la fase de Producto del Modelo 3P y da respuesta a los resultados que se obtuvieron en la encuesta, donde los estudiantes expresaron preocupación por la veracidad y la ética.

4 Conclusiones

La irrupción de la IA Generativa en el ámbito académico pone bajo la lupa las estrategias de evaluación tradicionales. Rúbricas centradas en el producto final quedan obsoletas, ya que es un hecho, si los estudiantes usan IAG para sus tareas, los resultados serán de muy buena calidad. Los hallazgos y pautas presentados en este artículo permiten concluir que la única respuesta pedagógica sostenible ante la automatización de contenidos es el desplazamiento del foco evaluativo desde el producto estático hacia la riqueza del proceso dinámico.

El diseño de rúbricas basadas en la agencia humana, el andamiaje dialógico y la soberanía cognitiva permite transformar la relación entre el estudiante y la tecnología. Al evaluar la calidad de los *prompts*, la capacidad de supervisión crítica y la transparencia en la atribución de contenidos, la rúbrica deja de ser un instrumento de fiscalización para convertirse en una herramienta que garantiza que la IA actúe como una prótesis intelectual y no como un sustituto del pensamiento. Esta transición asegura que el aprendizaje profundo se mantenga en el centro de la experiencia educativa, evaluando no solo "lo que el estudiante entrega", sino "lo que el estudiante es capaz de hacer y decidir" en colaboración con la máquina.

Asimismo, se concluye que la implementación de estas pautas requiere un compromiso institucional para fomentar la alfabetización en IA tanto en docentes como en alumnos. La evaluación centrada en el proceso no solo protege la integridad académica, sino que empodera al estudiante para desarrollar competencias críticas esenciales en el siglo XXI: la curaduría de información y la responsabilidad epistémica. En última instancia, una rúbrica bien diseñada es el mecanismo que garantiza que el título profesional siga certificando a un sujeto con capacidad de juicio independiente,

creatividad y ética, elementos que constituyen la base de la verdadera excelencia académica en la era algorítmica.

Se trata de un trabajo en desarrollo, que inició con el estudio de la percepción de los estudiantes, continuó con lo que se presenta en esta oportunidad: Pautas para el diseño de rubricas en contexto de IAG. Como trabajo futuro se propone la integración de las rúbricas de nueva generación en la planificación docente a fin de analizar resultados y refinar métodos.

Referencias

- Alier, M., García-Peñalvo, F. J., & Camba, J. D. (2024). Generative Artificial Intelligence in Education: From Deceptive to Disruptive. *International Journal of Interactive Multimedia and Artificial Intelligence, Special Issue on Generative Artificial Intelligence in Education, Vol. 8(5)*. doi: 10.9781/ijimai.2024.02.011
- Biggs, J. (1999). What the student does: Teaching for enhanced learning. *Higher Education Research & Development, 18*(1). doi: 10.1080/0729436990180105
- Blank, I. A. (2023). What are large language models supposed to model? *Trends in Cognitive Sciences, 27*(11), 987–989. doi: 10.1016/j.tics.2023.08.006
- Ilieva, G., Yankova, T., Ruseva, M., & Kabaivanov, S. (2025). A Framework for Generative AI-Driven Assessment in Higher Education. *Information- MDPI Journal, 16*. doi: 10.3390/info16060472
- Ka Yuk Chan, C., & Hu, W. (2023). Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *Springer Open. International Journal of Educational Technology in Higher Education*. doi: 10.1186/s41239-023-00411-8
- Knoth, N., Tolzin, A., Janson, A., & Leimeister, J. M. (2024). AI literacy and its implications for prompt engineering strategies. *Computers and Education: Artificial Intelligence, 6*, 100225. doi: <https://doi.org/10.1016/j.caeai.2024.100225>
- Okafor, J. O. (2024). Evaluation of Rubric-Based Assessment in Developing Critical Thinking Skills in Higher Education. *Jurnal Ilmiah Mahasiswa, 2*(2).
- Oroh, E. Z., Ali, M. I., Pelenkahu, N., Usman, H., & Rorintulus, O. A. (2025). Authentic Assessment in Higher Education to Increase Critical Thinking and Develop Metacognitive Awareness. *Studies in English Language and Education, 12*(2), 827-844. doi: 10.24815/siele.v12i2.45077
- Oviedo, S., Diaz, D., Alessio, C., Otazu, A., Cúneo, J. M., Becerra, M. d. C., & Minguez, M. (2025, 2025). *IA Generativa en la UNSJ: ¿cómo impacta el aprendizaje desde la mirada del estudiante?* Paper presented at the XXXI Congreso Argentino de Ciencias de la Computación (Río Negro, 6 al 10 de octubre de 2025).
- Sarzoza Herrera, S. (2013). Aprendizaje desde la perspectiva del estudiante: Modelo Teórico de Enseñanza Aprendizaje 3P. *Acción Pedagógica, 22*.
- Sengar, S. S., Hasan, A. B., Kumar, S., & Carroll, F. (2025). Generative artificial intelligence: a systematic review and applications. *Multimedia Tools and Applications, 84*(21), 23661–23700. doi: 10.1007/s11042-024-20016-1
- Taylor, B., Kisby, F., & Reedy, A. (2024). Rubrics in higher education: an exploration of undergraduate students' understanding and perspectives, *Assessment & Evaluation in Higher Education. Taylor & Francis Group, 49*(6), 799-809. doi: 10.1080/02602938.2023.2299330