



De la opacidad algorítmica a la automatización del aprendizaje Cuestionamientos sobre la agencia en el uso de IAGen en la educación superior

Rodrigo Alonso Cardoso González

INSTITUTO DE INVESTIGACIONES SOCIALES, UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Resumen. Este ensayo explora cómo la integración acelerada de la inteligencia artificial generativa (IAGen) en la educación superior puede propiciar una automatización del aprendizaje cuando se le atribuye una autoridad epistémica desde condiciones de opacidad algorítmica, lo que se traduce en un distanciamiento entre la intención y el resultado obtenido. El texto nos conduce desde debates sociológicos y filosóficos sobre agencia humana y artificial, hacia estudios recientes sobre prácticas de estudiantes. Centrado en esta perspectiva, se examinan tensiones entre asistencia y reemplazo, eficiencia contra trazabilidad, y un exceso de confianza en los algoritmos que puede poner en riesgo la voz propia. Finalmente, se sostiene que la responsabilidad se debe enseñar y debe ser distribuida, no individual, y que los riesgos son un reflejo de desafíos pedagógicos y estructurales de la educación que trascienden lo estrictamente técnico.

Abstract. This essay explores how the accelerated integration of Generative Artificial Intelligence (GenAI) in higher education can lead to an automation of learning when it is attributed epistemic authority under conditions of algorithmic opacity, resulting in a growing distance between intention and outcome. The text takes us from sociological and philosophical debates on human and artificial agency to recent studies on students' practices with GenAI. From this perspective, it examines tensions between assistance and replacement, efficiency versus traceability, and an overreliance on algorithms that may put one's own voice at risk. Finally, it is argued that responsibility should be taught and distributed rather than individual, and that these risks reflect pedagogical and structural challenges in education that transcend the strictly technical.

En los últimos años hemos observado un acelerado interés y desarrollo en sistemas de inteligencia artificial presentados como tecnologías capaces de extender e intensificar actividades que antes requerían más tiempo y recursos. Nos centraremos en la inteligencia artificial generativa (IAGen): un tipo que, con instrucciones breves, también denominadas prompts, facilita la creación de contenido nuevo como texto, imágenes, audio o video. Estos sistemas operan con modelos de lenguaje de gran tamaño entrenados con cantidades masivas de datos para aprender patrones. Su carácter revolucionario destaca

por el cambio en nuestra relación con la información, ya no sólo pasiva y de consulta, sino interactiva y con la simulación de una interlocución continua. Podemos explicar su impacto por su rápida adopción asociada a su accesibilidad, a las narrativas sobre sus posibilidades y a una interfaz conversacional que facilita la interacción basada en lenguaje natural.

Sostengo que cuando la IAGen se percibe como una autoridad epistémica en un contexto de opacidad algorítmica, puede conducir a una automatización de la toma de decisiones y a la delegación de la agencia en el ámbito educativo. Esto plantea desafíos para la integridad académica y el propio aprendizaje. Sin embargo, este texto se presenta como una invitación para entenderla no como una tecnología determinante, sino como un espejo que refleja y amplifica fenómenos y problemáticas, paralelas y precedentes. Para orientar buenas prácticas tenemos que dar un paso atrás y hacer preguntas en torno a diseños pedagógicos, protocolos institucionales o capacitación sobre su implementación, y el reconocimiento de condiciones diferenciadas como punto de partida en su uso.

La tensión central que exploramos gira en torno a la agencia en riesgo y cómo se reconfigura en condiciones de opacidad algorítmica y de automatización de actividades de aprendizaje. Pensemos en la agencia como la capacidad de una persona de reconocer y reflexionar sobre sus acciones, qué las motiva y qué consecuencias produce. Giddens (1997) la concibe como un proyecto reflexivo del yo, como un proceso en curso, que reconoce la participación de la propia persona y no como un atributo determinado. Permite pensar la agencia desde su carácter relacional.

La autoridad epistémica se produce cuando se reconoce en la otra persona (o entidad en este caso) una mayor disposición y fiabilidad por su conocimiento experto en un tema. Si lo planteamos en el escenario de la interacción con estos modelos, entramos en un terreno polémico. ¿Una inteligencia artificial generativa puede ser una autoridad epistémica? Para Murcia (2024) no es posible porque no es un igual epistémico, sino que propone que se trata de un igual doxástico; es decir, que mediante su interacción puede influir en las creencias de las y los usuarios. No puede ser un igual epistémico porque carece de voluntad; está configurado por algoritmos cuyos efectos pueden ser imperceptibles y un comportamiento por aprendizaje automatizado. En esta aproximación, retomando la agencia desde Giddens (1997), asumimos que la IAGen no puede tenerla dado que no posee voluntad ni reflexividad. Sin embargo, ¿qué ocurre si un sujeto, como usuario, le atribuye agencia mediante la interacción e interpretación?



Los modelos de lenguaje de gran tamaño continúan evolucionando rápidamente y hacen que la interacción humano-máquina sea cada vez más fluida y personalizada. Defiendo que, incluso cuando no pueden tener agencia, la simulan. Si la o el usuario la reconoce, aunque no exista, la agencia se torna real en los efectos de la interacción y en cómo influye en el uso que le dé a la información. Sobre este aspecto, las narrativas en torno a lo que puede lograr la IAGen pueden incidir en su uso y los imaginarios que produce (Coeckelbergh, 2023), una situación que requiere atención particular en el contexto de la educación superior. La atribución de autoridad epistémica no es una situación aislada o voluntaria, sino resultado de las condiciones que subyacen su acceso y uso.

Para analizar, entiendo la opacidad algorítmica como el efecto consecuente de una caja negra (Pasquale, 2015), particularmente como una alfabetización técnica asimétrica entre diseñadores y usuarios, y como una opacidad funcional-semántica que se evidencia en los sesgos e interpretación de los contenidos generados artificialmente (Burrell, 2016). De manera complementaria, desde una perspectiva sociotécnica de la teoría del actor-red (Latour, 2005), la IAGen puede entenderse como un actante que incide en las acciones humanas. Aquí debemos poner todavía más atención si se le atribuye un sentido de autoridad y se utiliza sin comprender cómo funciona.

La opacidad algorítmica también es pedagógica

En el ámbito educativo encontramos una clave adicional cuando pensamos en la metáfora de la caja negra sobre el uso de IAGen. Mi hipótesis es que la opacidad algorítmica también es pedagógica, consecuencia de un desconocimiento del propósito de las actividades de aprendizaje asignadas. El objetivo de la educación superior se convierte en otra caja negra. Esta provocación nos permite reflexionar sobre la motivación de la educación superior como una oportunidad de formación y no como un trámite. Cuando se comprende qué puede aportar una actividad a la formación académica o profesional, el compromiso puede ser mayor; en cambio, si se percibe como una tarea sin propósito, la asistencia de la IAGen puede tornarse una invitación a simular su realización. La aparente pérdida de agencia es también un cuestionamiento pedagógico, y no es un desafío menor.

La relación entre la IAGen como autoridad epistémica y la opacidad algorítmica, como escenario desde el cual se produce, configuran condiciones en las que el riesgo de automatización del aprendizaje se incrementa. Se evidencia mediante una distancia

epistémica, funcional y simbólica entre la intención de uso y el resultado obtenido: ¿entendiendo cómo y por qué obtuve esta información y no otra, y cuál fue mi rol en motivar esta respuesta? Si lo queremos examinar de otra manera, al señalar la agencia en riesgo, podemos ilustrarlo como la voz propia, la expresión de un posicionamiento reflexivo.

No se puede culpar estrictamente a la herramienta, pero tampoco a la o el usuario. Las prácticas emergen en un contexto caracterizado por condiciones estructurales que favorecen o restringen su exploración. Por sí misma, la tecnología no resuelve estas problemáticas. Esta premisa tampoco significa que en el ámbito educativo las y los estudiantes carezcan de agencia o que sean actores pasivos, sino que la ejercen al decidir cuándo, cómo y para qué usan la IAGen; un proceso que se debe acompañar e incentivar de manera informada. Se requieren condiciones para ejercer la agencia con libertad e información, y también con un marco ético en su uso. La solución no es prohibir tajantemente ni tampoco es factible. No basta con exigir la ética como si fuera un adorno, sino que se debe diseñar y construir colectivamente (Lugo, 2023; Pérez Salazar, 2025).

En la educación superior estos desafíos son cada vez más comunes y se encuentran en tensión con la oportunidad, que depende del caso. La IAGen se está integrando como soporte para redactar, sintetizar y estructurar actividades asignadas (Benavides-Lara et al. 2025). Cuando su uso se ve orientado por la optimización del tiempo, aunque parezca algo básico, puede ser un principio que desencadene otras consecuencias cuando la automatización. Además de acelerar el resultado, se predetermina como resultado de un sesgo de anclaje (Choi et al., 2024) al conformarse con las primeras sugerencias de contenido o referencias sin contrastar o cuestionar. Quienes usaron IAGen para una actividad de búsqueda de información se sintieron aliviados por la agilización y simplificación del proceso, pero sacrificaron profundidad, lo que se tradujo en una apropiación limitada (Stadler et. al., 2024).

Con frecuencia, las y los estudiantes son conscientes de la tensión entre asistencia y reemplazo (Vicente-Yagüe-Jara et al., 2023), y de su responsabilidad en la gestión de la información. Esto coincide con una tensión reflejada entre la oportunidad y el temor a la dependencia (Berzunza Criollo, 2024). Eso no implica una preocupación por un uso ético ni que las y los docentes identifiquen si se utiliza (Chao-Rebolledo y Rivera-Navarro, 2024). Podemos ubicar el problema en las condiciones que preceden su uso, en ocasiones marcadas por la ambigüedad de lineamientos y el estigma que afecta la declaración de uso.

Si nos centramos en la generación de información, que es uno de los usos más comunes entre estudiantes (Benavides-Lara et al. 2025; Berzunza Criollo, 2024), un estudio de caso sobre interacción realizado por varios investigadores (Amoozadeh et. al., 2024) distinguió dos principales: el iterativo y el lineal. El iterativo se produce mediante una interacción continua y reflexiva en la que, con base en el resultado obtenido, se hacen ajustes al prompt. Podríamos considerarlo como una meta-pedagogía de la actividad. El lineal se enfoca en el resultado, priorizando la resolución de problemas inmediatos. Para trascender de un uso básico y lineal a uno iterativo y reflexivo se requiere paciencia, diálogo y acompañamiento informado.

Cuando la IAGen se utiliza para acelerar el aprendizaje sin suficiente reflexión, puede favorecer lo que denomino comodidad epistémica; es decir, una disposición a aceptar contenidos generados sin cuestionarlos para resolver una necesidad inmediata. Complementa y se distingue de otras denominaciones como la comodidad cognitiva (Stadler et. al., 2024) y la descarga cognitiva (Gerlich, 2025), a partir de centrar la atención en la confianza y dependencia que se deposita en los sistemas de conocimiento experto en términos cercanos a lo planteado por Giddens (1997); aunque aquí operan en condiciones de opacidad y desiguales en el proceso de generación y evaluación de sus respuestas. Esta perspectiva permite poner atención en la situación del aprendizaje y no sólo en el proceso del aprendizaje, en examinar qué se busca automatizar, por qué y con qué consecuencias.

Considero que se crearán condiciones que propician la automatización y simulación del aprendizaje si se delega la toma de decisiones a la IAGen como suma de los siguientes aspectos: atribución de conocimiento experto; opacidad algorítmica; desconocimiento del propósito de las actividades de aprendizaje asignadas; ausencia de protocolos o alfabetización sobre cómo utilizarla desde principios éticos hacia la información. Esto revela una situación crítica sobre la comunicación del propósito de la educación más que de la tecnología implementada. Aunque los sesgos no son nuevos ni exclusivos de la IAGen, sí tiene un potencial de amplificarlos por la manera en que reconfigura la interacción con la información. Podemos pensar en estos modelos conversacionales como un eco. De acuerdo con los criterios de formulación de la pregunta e interpretación de la respuesta, pueden amplificar la desinformación desde el convencimiento o servir para contrastar o complementar un conocimiento previo. Depende de la actitud hacia la información buscada. Los sesgos de anclaje, cercanos a los de confirmación, se evidencian como una conformidad con los resultados obtenidos en la interacción con la IAGen sin una mayor contrastación, verificación o incluso iteración de la información (Choi et. al., 2024).

Horowitz y Kahn (2024) advierten que el sesgo no es fijo, sino que depende de la limitación de los conocimientos que preceden su uso, de la dificultad de la tarea asignada y de la confianza que se le atribuye; un conjunto que favorece una automatización de actividades.

Cuando hay una limitada profundidad analítica (Stadler et. al., 2024) y operan sesgos de anclaje en las respuestas (Choi et. al., 2024), la forma adquiere un papel central. Lo podemos observar mediante una atribución de autoridad epistémica con base en su fluidez lingüística y no por la validez de la información o trazabilidad del proceso (Paolucci, 2025), una puesta en escena discursiva que resulta convincente y consistente. Es una actuación porque la IAGen, como actante, encarna el rol y la actividad asignados. Sin embargo, a pesar de que no puede tener agencia como la entendemos (Giddens, 1997; Murcia, 2024), al menos desde sus limitaciones técnicas y acuerdos conceptuales, sí puede simularla y sus consecuencias serán resultado de cómo la gestión de dicha información se interpreta (Coeckelbergh, 2025) y de cómo se traduzca en acciones.

La relevancia de las preguntas informadas

No hay respuesta si no hay pregunta, lo cual implica que los contenidos generados no son de la IAGen, ni tampoco estrictamente de las y los usuarios, sino un resultado de su interacción. Pero no siempre hay conciencia de la incidencia que una persona tiene como usuaria en la respuesta generada. Si se actúa desde la opacidad y se privilegia la confianza en la información generada por IAGenen como autoridad epistémica y como brújula de su uso, la automatización de resultados implica que se pierda la trazabilidad de la evidencia de los argumentos presentados y la reflexividad sobre el proceso. Un ejemplo es el estudio realizado por González-Geraldo y Ortega-López (2024), quienes llegaron a conclusiones similares al encontrar que la mayor parte de universitarios no podían identificar el contenido generado por IAGenen, pero confiaban en lo que denominaron melodía del algoritmo, un convencimiento opaco y discursivo, bases que sustentan el riesgo de automatización en el aprendizaje.

Esta perspectiva nos permite examinar la potencial comodidad epistémica no como una cuestión de voluntad de delegar una agencia en el aprendizaje, sino como un problema que es resultado de una opacidad sistémica. En este escenario tecnológicamente desafiante, se requiere proporcionar condiciones para incentivar la agencia; de lo contrario, cuando el resultado de la interacción está encerrado en una caja negra, no sólo se desconoce cómo se generó la información, sino que también que, la instrucción es una herramienta de agencia,



un marco de referencia construido por la o el usuario que puede proporcionarle mayor control y dirección de la interacción, y de resultados obtenidos.

Una mejor pregunta, o una pregunta desde una posición informada, puede producir mejores resultados. No obstante, cuando se formula desde la oscuridad de la caja negra, la melodía del algoritmo (González-Geraldo y Ortega-López, 2024) marca el ritmo del aprendizaje y no el propio estudiante. Para hacer frente a esta situación, Coeckelbergh (2023) defiende que se requiere una responsabilidad narrativa, además de la causal, moral y relacional; en otras palabras, una responsabilidad hermenéutica donde se discuta lo humano y lo técnico, que se cuestionen los roles y cómo se afectan, y se reflexione activamente sobre cómo se usa la IAGen y cuáles son sus repercusiones.

Cuando se carece de una alfabetización sobre las implicaciones de su uso, la información obtenida artificialmente no siempre es cuestionada, verificada y contrastada. Este problema escala en el contexto de las alucinaciones artificiales, respuestas que parecen reales y que son presentadas como un hecho; pero que no existen. Además, este problema no sólo radica en que la fuente no exista, sino en que se use para informar o tomar decisiones. Todavía es insuficiente la conciencia sobre las alucinaciones y el rol de uno mismo para controlar la información generada, sumado a un exceso de confianza en las respuestas generadas (Williamson y Prybutok, 2024).

La integración de la IAGen en la educación nos posiciona frente a distintas tensiones. ¿Es una decisión o una reacción? ¿Cómo fomentar la asistencia y prevenir el reemplazo? El temor de dependencia no se reduce a estudiantes, sino como preocupación sobre el propio rol docente (Trejo-Quintana, 2023). Requerimos diagnósticos y capacitaciones constantes sobre cómo informar y dirigir su acompañamiento de acuerdo con protocolos accesibles (Castañeda et. al., 2025). Son necesarios esfuerzos coordinados para convertir estas herramientas en una oportunidad para construir bases para fomentar la agencia epistémica de estudiantes (Coeckelbergh, 2025). Una vía es el fomento de talleres y espacios para la escucha y el diálogo vinculante entre estudiantes, docentes e instituciones, además de propiciar condiciones que favorezcan una apropiación creativa y orientada por la ética de su uso (Lugo, 2023; Lugo y Cardoso-González, 2024).

La IAGen no va a resolver problemas automáticamente, pero tampoco es en sí misma el problema de la educación. Mi propósito no fue dar respuestas, sino mostrar tensiones como una oportunidad para reflexionar colectivamente. Presento esta posición frente al

ímpetu tecnológico que nos insta a tener precaución de no atribuir responsabilidades individuales cuando no se crean condiciones para su uso: es un proceso paralelo y transversal, individual y colectivo, tanto para estudiantes como docentes e instituciones.

La IAGen puede revolucionar la educación, pero no será una consecuencia de su espíritu tecnológico, sino de su diseño e implementación pedagógica, de lineamientos sobre su gestión y del reconocimiento de cuándo y cómo utilizarla; mas, también, de advertir que su uso no es una condición obligatoria. El tiempo apremia, pero hoy urge la pausa para observar, reflexionar y dialogar. Sólo así se puede construir transparencia frente a la opacidad, reconocer e incentivar la agencia con usos éticos e informados, y prevenir la automatización del aprendizaje al enfatizar el motivo pedagógico mediante un giro para privilegiar el proceso sobre el resultado final.

Fuentes

- Amoozadeh, M., Nam, D., Park, J., Ahn, K., Pan, Y., Huang, S., & Redmiles, E. (2024). Student-AI interaction: A case study of CS1 students. En *Proceedings of the 24th Koli Calling International Conference on Computing Education Research* (Artículo 13, pp. 1-13). Association for Computing Machinery. <https://doi.org/10.1145/3699538.3699567>
- Benavides-Lara, M. A., Rendón Cazales, V. J., Escalante Rivas, N., Martínez Hernández, A. M. del P., & Sánchez Mendiola, M. (2025). Presencia y uso de la inteligencia artificial generativa en la Universidad Nacional Autónoma de México. *Revista Digital Universitaria*, 26(1). <https://doi.org/10.22201/ceide.16076079e.2025.26.1.10>
- Berzunza Criollo, M. (2025). Percepciones y usos de ChatGPT en la construcción del pensamiento crítico en estudiantes universitarios. *Ciencia y Reflexión*, 4(3), 201-235. <https://doi.org/10.70747/cr.v4i3.413>
- Burrell, J. (2016). How the machine "thinks": Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1-12. <https://doi.org/10.1177/2053951715622512>

- Castañeda, L. M., Ramírez, A. Y., Castillejos, J. M., & Ventura, M. T. (2025). *Uso y desarrollo de la Inteligencia Artificial en la Universidad: docencia e investigación*. Dirección General de Cómputo y Tecnologías de Información y Comunicación (DGTIC), UNAM.
- Chao-Rebolledo, C., & Rivera-Navarro, M. A. (2024). Usos y percepciones de herramientas digitales de inteligencia artificial en la educación superior en México. *Revista Iberoamericana de Educación*, 95(1), 57-72. <https://doi.org/10.35362/rie9516259>
- Choi, A. S., Akter, S. S., Singh, J. P., & Anastasopoulos, A. (2024, noviembre). The LLM effect: Are humans truly using LLMs, or are they being influenced by them instead? En Y. Al-Onaizan, M. Bansal, & Y.-N. Chen (Eds.), *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing* (pp. 22032-22054). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.emnlp-main.1230>
- Coeckelbergh, M. (2023). Narrative responsibility and artificial intelligence. *AI & Society*, 38(6), 2437-2450. <https://doi.org/10.1007/s00146-021-01375-x>
- Coeckelbergh, M. (2025). AI and epistemic agency: How AI influences belief revision and its normative implications. *Social Epistemology*, 40(1). <https://doi.org/10.1080/02691728.2025.2466164>
- Galindo-Domínguez, H., Delgado, N., Campo, L., & Sainz de la Maza, M. (2024). Uso de ChatGPT en educación superior: Un análisis en función del género, rendimiento académico, año y grado universitario del alumnado. *REDU. Revista de Docencia Universitaria*, 22(2), 16-30. <https://doi.org/10.4995/redu.2024.21647>
- Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(6).
- Giddens, A. (1997). *Modernidad e identidad del yo: el yo y la sociedad en la época contemporánea*. Ediciones Península.
- González-Geraldo, J. L., & Ortega-López, L. (2024). ¿Puede engañarnos una IA? Carencias del estudiantado universitario para detectar ChatGPT. *Education in the Knowledge Society (EKS)*, 25. <https://doi.org/10.14201/eks.31760>
- Horowitz, M. C., & Kahn, A. S. (2024). Bending the automation bias curve: A study of human and AI-based decision making in national security contexts. *International Studies Quarterly*, 68(2), sqae020. <https://doi.org/10.1093/isq/sqae020>
- Latour, B. (2005). *Reensamblar lo social. Una introducción a la teoría del actor-red*. Manantial.

- Lugo, L. J. (2023). *Guía para uso ético de la inteligencia artificial: Una propuesta para la investigación y educación*. Amazon y Tesicafé.
- Lugo, L., & Cardozo, R. (2025). Desafíos para la apropiación disruptiva en el uso de inteligencia artificial en contextos académicos: Análisis de casos en la UNAM. En *Inteligencia artificial en la docencia y la investigación* (pp. 90-107). Zenodo. <https://doi.org/10.5281/zenodo.15837921>
- Aguirre Boza, Á. S., & Giménez Castillo, B. (Eds.). (2024). *Tecnología y sociedad: Vol. 5. Inteligencia artificial en la docencia y la investigación*. Ariadna Ediciones; Universidad de Santiago de Chile, Facultad Tecnológica.
- Murcia, A. (2024). ¿Es la inteligencia artificial un igual epistémico? *Daimon. Revista Internacional de Filosofía*, 93, 119-135. <https://doi.org/10.6018/daimon.611821>
- Paolucci, C. (2025). The myth of meaning: Generative AI as language-endowed machines and the machinic essence of the human being. *Semiotica*, 2025(262), 5-23. <https://doi.org/10.1515/sem-2024-0204>
- Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press. <https://www.jstor.org/stable/j.ctt13x0hch>
- Pérez Salazar, G. (2025). Uso responsable y ético de la inteligencia artificial generativa en los contextos universitarios. Posibilidades y limitaciones. En R. Espinosa & F. Ortiz (Coords.), *Hilos del conocimiento: narrativas de educación e inteligencia artificial* (pp. 82-98). UASLP.
- Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, 160, Artículo 108386. <https://doi.org/10.1016/j.chb.2024.108386>
- Trejo-Quintana, J. (2023). Más preguntas que respuestas: La inteligencia artificial y la educación. *Perfiles Educativos*, 45(especial), 43-55. <https://doi.org/10.22201/iisue.24486167e.2023.Especial.61690>
- Vicente-Yagüe-Jara, M. I., López-Martínez, O., Navarro-Navarro, V., & Cuéllar-Santiago, F. (2023). Escritura, creatividad e inteligencia artificial: ChatGPT en el contexto universitario. *Comunicar*, 77, 47-57. <https://doi.org/10.3916/C77-2023-04>
- Williamson, S. M., & Prybutok, V. (2024). The era of artificial intelligence deception: Unraveling the complexities of false realities and emerging threats of misinformation. *Information*, 15(6), 299. <https://doi.org/10.3390/info15060299>