

Identidad moral y uso ético de la inteligencia artificial generativa en la educación superior

Moral Identity and Ethical Use of Generative Artificial Intelligence in Higher Education

César Augusto García Avitia

Universidad de Colima

<https://orcid.org/0000-0003-1841-3033>

garciaavitia@ucol.mx

México

Resumen:

Este artículo analiza cómo la teoría de la identidad moral puede contribuir a afrontar los desafíos éticos asociados con el uso de la inteligencia artificial generativa (IAG) en la educación superior. A partir de evidencia reciente que muestra una adopción masiva de estas tecnologías por parte del estudiantado universitario, el trabajo examina las tensiones que surgen entre innovación tecnológica, integridad académica y formación profesional. Mediante un enfoque analítico-argumentativo basado en revisión de literatura interdisciplinaria en psicología moral, ética aplicada y educación superior, el artículo examina primero el uso de la IAG como problema ético-educativo, considerando riesgos relacionados con la deshonestidad académica, la dependencia cognitiva y los problemas epistémicos asociados a sesgos y errores en sistemas de IAG. Posteriormente se desarrolla el marco conceptual de la identidad moral, destacando su función como mecanismo autorregulatorio que vincula valores internalizados con el comportamiento moral. Se argumenta que fomentar la internalización de valores como honestidad, responsabilidad epistémica y justicia puede fortalecer el uso ético de la IAG más allá de mecanismos de vigilancia o sanción. Finalmente, se contrasta este enfoque formativo con estrategias institucionales de prohibición o hipervigilancia, señalando limitaciones y proponiendo integrar el desarrollo de la identidad moral en la formación universitaria.

Palabras clave: identidad moral, inteligencia artificial generativa, integridad académica, ética en la educación superior, comportamiento moral

Abstract:

This article analyzes how moral identity theory can help address the ethical challenges associated with the use of generative artificial intelligence (GAI) in higher education. Drawing on recent evidence showing the widespread adoption of these technologies among university students, the paper examines tensions between technological innovation, academic integrity, and professional education. Using an analytical and argumentative approach based on an interdisciplinary literature review in moral psychology, applied ethics, and higher education studies, the article first conceptualizes the use of GAI as an ethical-educational issue. It discusses risks related to academic dishonesty, cognitive dependency, and epistemic problems such as bias and hallucinations in AI systems. The paper then develops the conceptual framework of moral identity, highlighting its role as a self-regulatory mechanism that connects internalized moral values with behavior. Building on this framework, the article argues that fostering values such as honesty, epistemic responsibility, and fairness can support ethical uses of GAI beyond external monitoring or sanctioning mechanisms. Finally, the study contrasts this educational perspective with

institutional responses based on prohibition or technological surveillance, pointing out their practical and pedagogical limitations and proposing the integration of moral identity development into university education.

Keywords: *moral identity, generative artificial intelligence, academic integrity, ethics in higher education, moral behavior*

Recibido: 23/03/2026 | **Aceptado:** 26/06/2026 | **Publicado:** 01/07/2026 | pp. 1 – 15

DOI: 10.19136/etie.v9n17.6391

Identidad moral y uso ético de la inteligencia artificial generativa en la educación superior

| **Introducción**

La irrupción de la inteligencia artificial generativa (IAG) en la educación superior ocurrió con una velocidad que pocas innovaciones tecnológicas han mostrado en la historia reciente. En un contexto donde herramientas capaces de producir texto, resumir lecturas, proponer ideas de investigación, traducir y redactar cualquier tipo de documento se han vuelto de uso cotidiano, la discusión ética no puede ser periférica: se ubica en el centro de la vida académica y de la formación profesional. El uso masificado de esta tecnología implica toda clase de efectos que deben reflexionarse y las universidades deben mostrar una participación plena en este debate, tanto por sus implicaciones en la vida académica como en la sociedad en general.

La relevancia que ha tomado el uso de la IAG en la educación superior es enorme. Algunas cifras disponibles confirman una penetración masiva. Un ejemplo es el caso del Reino Unido, donde el informe Student Generative AI Survey 2025 elaborado por Freeman (2025) para la Higher Education Policy Institute y Kortext reportó que 92% del estudiantado universitario encuestado dijo haber usado IA en alguna forma (frente a 66% en 2024) y que 88% había utilizado IAG en relación con evaluaciones (frente a 53% el año previo). Además, 18% afirmó haber incluido texto generado por IA directamente en su trabajo (en la categoría de AI-generated and edited text o uso directo), lo cual coloca el desafío de la integridad académica como un problema estructural.

En paralelo, la encuesta global de la Digital Education Council (2024) con 3,839 estudiantes de 16 países indicó que 86% usa IA regularmente en sus estudios y que 54% la utiliza semanalmente y 24% que la usa a diario. También encontró baja preparación percibida: 58% declaró no tener conocimientos y habilidades suficientes en IA, y solo 5% reportó plena comprensión de directrices o guías institucionales. La categoría de uso que más participantes reportaron fue la de búsqueda de información, con un 69%, pero el 24% ya reporta usarlo para crear borradores de documentos. Entre los países donde se recabó información para esta encuesta se encuentra México, en donde el Tecnológico de Monterrey fue la institución responsable.

En el mismo año que el estudio de la Digital Education Council, una investigación en la Universidad de Harvard con 360 participantes estudiantes de pregrado y posgrado halló que cerca de 65% de estudiantes había usado o planeaba usar IAG para trabajo académico (Deschenes y McMahon, 2024). Sin embargo, también el 65% no consideraba sus resultados lo suficientemente confiables para uso académico, lo que apunta a un paisaje de adopción de esta tecnología a pesar de las altas dudas sobre su calidad. Esto denota una carencia de reflexión ética sobre el uso de esta tecnología que no debe pasarse por alto, incluso en estudiantes de instituciones de alto reconocimiento por su excelencia académica.

Más recientemente, la Digital Education Council (2026) realizó un estudio en América Latina, incluyendo México, sobre el uso de la inteligencia artificial entre estudiantes universitarios, el cual se ha consolidado como una práctica habitual: el 92% emplea estas herramientas en sus actividades académicas y el 67% las utiliza de manera diaria o semanal. ChatGPT es la aplicación predominante (88%), principalmente para buscar información (65%), generar ideas o borradores (61%), apoyar la redacción (47%) y resumir documentos (42%). Además, el 73% considera que la IA será una

competencia relevante para su futuro profesional. Sin embargo, aunque el 68% mantiene una actitud positiva hacia su incorporación en la educación, persisten preocupaciones importantes, ya que el 65% considera que puede propiciar un aprendizaje superficial y afectar el pensamiento crítico y la creatividad, mientras que el 56% manifiesta inquietud por la privacidad de los datos personales, lo que refleja una adopción amplia acompañada de desafíos éticos y educativos que las instituciones de educación superior deben atender.

En el caso de México, la Encuesta Nacional sobre Usos y Percepciones de la Inteligencia Artificial Generativa en la Educación Superior en México (ENIAG 2025) confirma que la IA se ha integrado ampliamente a la vida académica del estudiantado mexicano (Secretaría de Educación Pública [SEP], 2025). El estudio analizó las respuestas de 1,143,451 estudiantes de instituciones de educación superior de todo el país, encontrando que el 93% conoce el concepto de inteligencia artificial generativa y el 66% la utiliza de manera cotidiana (al menos una vez por semana). Asimismo, el 79% emplea estas herramientas para generar textos y el 82.3% las considera un recurso útil para apoyar tareas complejas como analizar, razonar, reflexionar y crear. En cuanto a su impacto, el 69% percibe que la IA ha mejorado su desempeño académico, el 80% considera que transformará su área de estudio y el 76.2% manifiesta la necesidad de recibir formación específica en esta tecnología. No obstante, también persisten preocupaciones sobre sus implicaciones, pues uno de cada tres estudiantes expresa inquietud por la desinformación generada por la IA y la mayoría afirma utilizar estas herramientas con cautela y verificar sus respuestas, lo que evidencia una adopción masiva acompañada de una actitud crítica frente a sus riesgos y limitaciones.

Con base en lo anteriormente descrito, la incorporación masiva de la inteligencia artificial generativa en la educación superior plantea desafíos éticos que van más allá de la tecnología e impactan directamente los procesos formativos. Entre ellos destaca la integridad académica, al cuestionarse la autoría, la originalidad y el esfuerzo intelectual cuando se utilizan contenidos generados por IA. Asimismo, preocupa su posible efecto sobre la calidad del aprendizaje, ya que un uso acrítico puede favorecer aprendizajes superficiales, dependencia tecnológica y debilitamiento del pensamiento crítico y la creatividad. También existen riesgos asociados a la confiabilidad de la información, debido a la generación de errores, sesgos y desinformación, así como a la privacidad y protección de los datos personales. A ello se suma la insuficiente alfabetización en IA y la falta de lineamientos institucionales claros para orientar su uso responsable. En este contexto, las instituciones de educación superior enfrentan el reto de promover una formación ética, fortalecer las competencias digitales y establecer políticas que permitan aprovechar el potencial educativo de la IA sin comprometer los principios fundamentales de la educación universitaria.

Este panorama revela una tensión central: si el uso de IAG se vuelve ubicuo, ¿qué enfoque educativo puede sostener la integridad académica sin caer en respuestas que dañen la confianza, como la hipervigilancia, o resulten impracticables, como la prohibición total? Así, para tratar de responder a esta pregunta, este artículo defiende que la teoría de la identidad moral (Blasi, 1984; Boegershausen et al., 2015; Narvaez y Lapsley, 2009) ofrece una vía robusta para afrontar el problema: desplaza el foco desde la supervisión externa hacia la autorregulación y la coherencia del yo moral, sin abandonar el diseño institucional y pedagógico. En otras palabras, el uso ético de IAG no puede depender solo de reglas y sanciones; requiere también que las y los estudiantes incorporen, como parte de su autoconcepto, disposiciones morales que hagan razonable actuar con honestidad, responsabilidad y justicia sin necesidad de supervisión externa. Cabe resaltar que esta perspectiva es relevante para ser aplicada de

manera universal, pero requiere adaptaciones de acuerdo a las condiciones de cada contexto sociocultural y las oportunidades y obstáculos que estos brindan.

El objetivo de este artículo es analizar cómo los aportes de la teoría de la identidad moral pueden aplicarse para afrontar los desafíos éticos que plantea el uso de la inteligencia artificial generativa en la educación superior. Se argumenta que el desarrollo de una identidad moral en estudiantes y docentes constituye una estrategia formativa clave para promover un uso responsable de estas tecnologías. Asimismo, el texto busca examinar la importancia y los desafíos de integrar el desarrollo de la identidad moral en la formación universitaria, proponer estrategias educativas para fomentarla y contrastar este enfoque formativo con modelos de regulación basados en la prohibición, la hipervigilancia o el control tecnológico, evaluando sus limitaciones frente a una perspectiva centrada en la formación ética del estudiantado.

Para orientar el desarrollo del artículo se adoptó un enfoque analítico-argumentativo basado en revisión de literatura académica. En primer lugar, la introducción presenta el contexto actual del uso de la inteligencia artificial generativa en la educación superior, apoyándose en estadísticas recientes provenientes de informes institucionales y estudios académicos que documentan su creciente adopción entre estudiantes universitarios. Posteriormente, el texto se organiza en varios momentos argumentativos. En la primera sección se expone cómo el uso de la IAG es un problema ético en el ámbito universitario. Después se desarrolla la teoría de la identidad moral, revisando sus principales aportes desde la psicología moral y la ética aplicada. En la tercera sección se analiza cómo los principios de esta teoría permiten comprender los desafíos éticos asociados al uso de la inteligencia artificial generativa en el ámbito universitario. En la cuarta sección se argumenta la importancia de incorporar el desarrollo de la identidad moral como objetivo formativo en la educación superior, identificando también los desafíos institucionales y pedagógicos que ello implica. En la quinta sección se presentan propuestas educativas concretas orientadas a promover una identidad moral congruente con el uso ético de estas tecnologías. Finalmente, el artículo contrasta esta perspectiva formativa con enfoques centrados en la prohibición, la hipervigilancia o el control tecnológico del estudiantado, con el fin de argumentar que la construcción de una identidad moral sólida constituye una estrategia más sostenible y coherente con los fines educativos de la universidad.

| *La IAG como problema ético-educativo en la universidad*

En educación superior, la IAG se presenta tanto como infraestructura cognitiva o una extensión de capacidades para buscar, resumir, redactar; así como un atajo que puede sustituir procesos formativos, como el pensamiento crítico, la escritura, la argumentación, el diseño metodológico, entre otros. La guía de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO, 2023) sobre IAG en educación e investigación subraya controversias asociadas con estos sistemas: privacidad, derechos de autor, manipulación de contenido, sesgos y riesgos de dependencia o interacción pedagógicamente inapropiada; la recomendación central es un enfoque centrado en la agencia humana y en el diseño pedagógico responsable.

Se han identificado varias categorías de riesgo con respecto al uso de la IAG en educación superior, pero para este artículo enfatizaremos en dos de ellos, comenzando por las implicaciones sobre integridad académica, entendida como un compromiso, incluso ante la adversidad, con seis valores fundamentales: honestidad, confianza, equidad, respeto, responsabilidad y valentía aplicados a las actividades

académicas (International Center for Academic Integrity, 2025). En una revisión sistemática enfocada en educación superior (Bittle y El-Gayar, 2025) se identifica que la IAG crea oportunidades de apoyo, como tutoría, accesibilidad y retroalimentación; pero también incentivos y medios para nuevas formas de deshonestidad como trabajos completos generados por esta tecnología o parafraseo automatizado que no permiten el aprendizaje proporcional.

La segunda categoría de riesgos son los epistémicos, pues se han documentado ampliamente fenómenos como las alucinaciones, sesgos y falsa confianza (Emsley, 2023; Maleki et al., 2024; Salvagno et al., 2023). Esto representa que la IAG puede presentar de manera muy convincente información falsa o, en otros casos, omitir cierta información y enfatizar en otra debido a sesgos. La anteriormente citada encuesta de la Higher Education Policy Institute (Freeman, 2025) reporta que los resultados falsos y las alucinaciones, así como los sesgos aparecen como factores que inhiben el uso, lo cual es una señal de que estudiantes reconocen, al menos parcialmente, que la IAG no es una fuente confiable por defecto.

Frente a estos riesgos, una respuesta común en el corto plazo ha sido reforzar controles: prohibición en ciertos cursos, vigilancia tecnológica, detectores de texto generado por IA y procedimientos disciplinarios. Sin embargo, este camino tiene límites empíricos y éticos. En primer lugar, la evidencia técnica y organizacional sobre detección muestra incertidumbre y posibilidad de falsos positivos. La propia OpenAI (2023), compañía creadora de ChatGPT, primera plataforma de acceso público con IAG de lenguaje natural, advirtió desde el lanzamiento de su clasificador de texto que podía etiquetar erróneamente textos humanos como generados por IA y que era muy poco fiable en textos cortos; posteriormente, la discusión pública sobre detectores se consolidó en torno a su baja confiabilidad en escenarios reales.

Asimismo, documentación técnica de Turnitin (2026), una reconocida plataforma de identificación de plagio y uso de IA en textos, reconoce explícitamente la posibilidad de falsos positivos y describe rangos donde la interpretación puede ser menos confiable. En investigación académica aplicada a educación superior (Bittle y El-Gayar, 2025), también se han discutido problemas de confiabilidad y efectos éticos de la vigilancia basada en detección, la cual incluso puede generar tensiones entre estudiantes y docentes, afectando la percepción de justicia en las evaluaciones.

En este punto emerge una pregunta central: si la vigilancia es imperfecta y la adopción de la IAG en la educación superior es ya masiva, ¿cómo puede construirse una integridad académica respecto de su uso que sea verdaderamente sostenible? La respuesta no parece encontrarse únicamente en el aumento de controles externos, en la amenaza de sanciones o en la sofisticación de los mecanismos de detección, sino en la formación de sujetos capaces de autorregular su conducta incluso en ausencia de supervisión. Aquí es donde la teoría de la identidad moral aporta una lente distinta y particularmente fecunda: permite comprender el uso de la IAG no solo como un problema técnico, normativo o instrumental, sino como una situación moral que exige deliberación, autoevaluación y coherencia con valores internalizados (Boegershausen et al., 2015).

Desde esta perspectiva, la pregunta relevante deja de ser solamente ¿qué tanto puede vigilar la institución? para desplazarse hacia otra más profunda: ¿qué tipo de persona y de profesional está formando la universidad frente al uso de estas tecnologías? Cuando la honestidad, la responsabilidad, la justicia y el compromiso con el propio aprendizaje forman parte del autoconcepto moral del estudiantado, el uso ético de la IAG deja de depender exclusivamente de la posibilidad de ser descubierto y pasa a sostenerse en convicciones personales. Así, la integridad académica se fortalece no solo como obediencia a reglas externas, sino como expresión de una identidad moral que orienta las decisiones

incluso en contextos ambiguos, de baja supervisión o alta tentación de delegar indebidamente el trabajo intelectual.

| ***Teoría de la identidad moral y evidencia de su relevancia***

El concepto de identidad ha ocupado un lugar central en la psicología del desarrollo, la psicología social y la psicología moral, evolucionando desde perspectivas centradas en la formación del yo hasta enfoques contemporáneos que destacan su carácter narrativo, social y moral. Uno de los primeros trabajos sistemáticos del concepto proviene de la obra de Erik Erikson, quien introdujo la noción de identidad del yo en el marco de su teoría psicosocial del desarrollo. Para Erikson (1968), la identidad constituye un sentido relativamente coherente de continuidad personal que se consolida especialmente durante la adolescencia, etapa en la cual los individuos enfrentan la crisis de identidad versus confusión de roles. En este sentido, la identidad se refiere a la integración de experiencias personales, valores y compromisos que permiten a la persona reconocerse como la misma a lo largo del tiempo y actuar con cierta consistencia en distintos contextos.

Posteriormente, la investigación empírica sobre identidad fue ampliada por James Marcia (1966), quien operacionalizó las ideas de Erikson mediante el modelo de estatus de identidad, basado en los procesos de exploración y compromiso. Marcia propuso cuatro estados posibles en el desarrollo de la identidad, que son logro, moratoria, exclusión y difusión. Estos describen diferentes configuraciones del proceso mediante el cual las personas exploran alternativas y asumen compromisos respecto a valores, creencias y proyectos de vida. Este modelo permitió estudiar la identidad de manera más sistemática en contextos educativos y sociales, destacando su carácter dinámico y evolutivo.

A partir de las décadas siguientes, el concepto de identidad se amplió hacia perspectivas más sociales y constructivistas. Dan McAdams (1993) ha señalado que la identidad puede entenderse también como una narrativa personal, es decir, una historia que las personas construyen sobre sí mismas para dar sentido a su pasado, presente y futuro. Desde esta perspectiva, la identidad no es únicamente una estructura interna estable, sino un proceso continuo de interpretación de la experiencia que se configura en interacción con contextos culturales y sociales.

En las últimas décadas, la investigación ha incorporado además la dimensión ética del autoconcepto a través del concepto de identidad moral. Investigadores como Augusto Blasi (1984, 1993) han argumentado que los valores morales pueden integrarse en el núcleo de la identidad personal, de modo que actuar de acuerdo con ellos se convierte en una forma de mantener la coherencia con el propio sentido de sí mismo. En esta línea, la identidad moral se refiere al grado en que los rasgos y valores morales son centrales para el autoconcepto y, por tanto, influyen en la motivación para comportarse éticamente. Lapsley (2015) subraya que identidad y moralidad están profundamente implicadas y que la identidad moral puede entenderse como parte de procesos de integración del yo y regulación interna, con implicaciones para cómo se narra y comprende la responsabilidad personal.

La teoría de la identidad moral surge, en parte, como respuesta a la brecha entre juicio moral y acción: personas pueden saber qué es lo correcto y aun así no hacerlo (Narvaez y Lapsley, 2009). Un aspecto clave es que el comportamiento moral depende no solo de razonar normas, sino de que la moralidad sea importante para la persona y su perspectiva sobre sí mismo, es decir, su autoconcepto. Aquino y Reed (2002) definen la identidad moral como una autoconcepción organizada en torno a un conjunto de rasgos

morales. Algunos ejemplos de estos rasgos podrían ser la honestidad, la compasión, la justicia u otros criterios, valores o principios éticos y morales que pueden convertirse en adjetivos y definir a la persona, como sería el caso de considerarla honesta, compasiva o justa. Los autores proponen que su influencia en el comportamiento depende de cuán importante o central es esa identidad moral para la persona. En su marco, la identidad moral no es un rasgo fijo, pues puede activarse o suprimirse por el contexto, y su prominencia varía con experiencias y ámbitos de vida.

Un aporte influyente de Aquino y Reed (2002) es distinguir dos dimensiones operativas: a) internalización, que es el grado en que los rasgos morales son centrales para el autoconcepto y b) simbolización, que es el grado en que esos rasgos se expresan públicamente mediante acciones y señales sociales. Esta distinción es crucial para educación superior, porque permite entender por qué algunas conductas éticas pueden ser meramente reputacionales, como mostrar ciertas virtudes y no necesariamente autorregulatorias, como ser virtuoso en privado. En general, la teoría y la evidencia científica sugieren que la identidad moral importa porque funciona como un mecanismo autorregulatorio, pues cuando la moralidad es central para el yo, actuar inmoralmemente amenaza la coherencia personal y produce costos internos como la culpa, la vergüenza o la disonancia, lo que aumenta la probabilidad de elegir el comportamiento correcto incluso bajo presión.

Diversos estudios empíricos han mostrado que la identidad moral se asocia de manera significativa con la conducta prosocial, el comportamiento ético y la evitación de acciones antisociales, ya que fortalece la disposición de las personas a actuar de acuerdo con sus principios morales incluso en situaciones de ambigüedad o presión social (Hertz y Krettenauer, 2016; Jennings et al., 2015). En este sentido, la identidad moral no solo constituye un componente del desarrollo moral, sino también un mecanismo motivacional clave que conecta el juicio moral con la acción, al fomentar la autorregulación y la coherencia entre valores, decisiones y comportamiento.

El metaanálisis realizado por Hertz y Krettenauer (2016), con 111 estudios en múltiples campos, estimó una asociación significativa entre identidad moral y conducta moral (modelo de efectos aleatorios, $r = .22$, $p < .01$, IC 95% [.19, .25]). Los tamaños del efecto fueron similares entre distintos tipos de resultados conductuales, incluyendo conductas prosociales, evitación de comportamientos antisociales y comportamientos éticos. Sin embargo, los estudios basados exclusivamente en autoinformes presentaron tamaños de efecto mayores, mientras que los más pequeños se observaron en investigaciones que emplearon medidas implícitas o técnicas de priming (influencia inconsciente) para activar la identidad moral. Asimismo, se encontró un efecto cultural marginal: los estudios realizados en culturas colectivistas tendieron a reportar tamaños de efecto menores que aquellos realizados en culturas individualistas.

En paralelo, el metaanálisis de Lefebvre y Krettenauer (2019) encontró una asociación robusta entre identidad moral y emociones morales ($r = .32$, $p < .01$, IC 95% [.27, .36]). La magnitud de esta relación varió según el tipo de emoción moral evaluada: los efectos más altos se observaron en emociones orientadas hacia otras personas, como la empatía, la simpatía y la compasión ($r = .41$), mientras que los más bajos correspondieron a emociones negativas evaluativas hacia otros, como la ira moral, el desprecio y el asco ($r = .16$). Esto es relevante para el uso ético de IAG porque muchas decisiones problemáticas, como entregar texto generado completamente con esa tecnología, implican no solo reglas, sino sensibilidad ante daños a otros, como la percepción de injusticia hacia los pares, el engaño a docentes y la erosión de confianza ante las instituciones.

Desde una perspectiva de desarrollo y contexto, un metaanálisis reciente sobre antecedentes de identidad moral resalta que factores del entorno social y organizacional, como los estilos de liderazgo y clima ético, influyen en su formación (Xu et al., 2024). El análisis de los tamaños del efecto en 110 estudios que involucraron a 44,441 participantes muestra que el género, los rasgos de personalidad y el contexto organizacional están fuertemente asociados con la identidad moral. Aunque las investigaciones revisadas se han dado en el marco de la psicología organizacional, su transferencia a educación superior es directa en diversas variables, pues una universidad, facultad o programa es también una organización con normas, incentivos, modelos a seguir y señales de estatus.

| *Aplicación de la identidad moral para el uso ético de la IAG en educación superior*

La IAG convierte muchas tareas académicas en decisiones morales de baja fricción, es decir, que no implican un juicio moral. Copiar y pegar un texto generado puede ser inmediato, difícil de detectar y, en ciertos entornos estudiantiles, incluso podría estar socialmente normalizado. Cuando el costo externo de la deshonestidad se reduce, el peso relativo del costo interno, relacionado con la autoimagen y la coherencia con valores, aumenta. Esta es precisamente la zona donde la identidad moral ofrece capacidad explicativa y también potencial pedagógico.

La investigación sobre trampas académicas sugiere que el comportamiento deshonesto no se explica solo por falta de conocimiento de reglas, sino que intervienen presiones de desempeño, integración social y racionalizaciones como la minimización del daño y justificaciones instrumentales (Miles et al., 2022). En el caso de IAG, la encuesta HEPI (Freeman, 2025) muestra que una razón importante para evitar usar IA es el temor a ser acusado de mala conducta académica, mientras que también aparece la percepción de ahorro de tiempo y mejora de calidad como motivadores; este patrón es consistente con un escenario donde hay incentivos fuertes para lograr la eficiencia y, simultáneamente, ambigüedad moral situacional.

Un estudio publicado en *Computers and Education: Artificial Intelligence* con 25 estudiantes universitarios en Estados Unidos concluyó que la integración de IAG se percibe como inevitable, que las fronteras entre trampa evidente y zonas grises generan desacuerdos, y que la influencia de pares y ética personal pesa fuertemente (Huang et al., 2025). Además, políticas de clase aplicadas por docentes pueden influir más que políticas institucionales amplias. Esto sugiere que la intervención no puede reducirse a reglas establecidas en los niveles altos de las universidades, sino que debe operar en el nivel de identidad, normas de grupo y prácticas docentes concretas.

Por otra parte, la distinción entre internalización y simbolización (Aquino y Reed, 2002) en la identidad moral permite reinterpretar fenómenos frecuentes en el debate universitario. Primeramente, un estudiante puede defender públicamente la integridad y a la vez usar IAG para entregar un ensayo casi completo, justificándolo como solo edición. Eso sería alta simbolización con baja internalización. Inversamente, un estudiante con alta internalización podría usar IAG para mejorar claridad o explorar fuentes, pero registrando su proceso, citando apoyos y rechazando la entrega de texto generado como propio, incluso si no hay vigilancia. Usando esa perspectiva, la internalización es más cercana a la coherencia del yo, mientras la simbolización capta, en parte, componentes de autopresentación. En contextos de IAG, donde la reputación puede gestionarse, por ejemplo humanizando un texto generado por esta tecnología para evitar detectores, una ética centrada solo en simbolización es vulnerable.

Cabe resaltar que el uso ético de IAG no se reduce a evitar plagiar; incluye lo que puede llamarse responsabilidad epistémica, que representa verificar, reconocer límites, distinguir entre evidencia y plausibilidad, y evitar la propagación de errores. La UNESCO (2023) enfatiza que se requiere alfabetización en dimensiones humanas y técnicas de la IA, y un enfoque centrado en la interacción pedagógicamente apropiada. Visto desde identidad moral, si ser una persona honesta y responsable es central al yo, entonces entregar información no verificada generada por un modelo no es un mero error técnico, es una falla de responsabilidad con consecuencias para el ecosistema de conocimiento académico (Deep et al., 2025). Esto conecta con marcos de gobernanza que conceptualizan la IA como tecnología sociotécnica con riesgos a gestionar. Por ejemplo, el marco de gestión de riesgos de IA del National Institute of Standards and Technology (2023) enumera características de IA confiable como son la validez, seguridad, transparencia, explicabilidad, privacidad y equidad; y propone funciones para gobernar, mapear, medir y gestionar riesgos. Aunque ese marco está orientado a empresas productivas, ofrece un lenguaje útil para universidades que quieran traducir valores a prácticas.

Otra dimensión más es que la identidad moral también aplica a dilemas distributivos. Si el acceso a herramientas premium o a tutorías basadas en IA amplifica ventajas socioeconómicas, entonces la universidad enfrenta un problema de justicia educativa: ¿cómo evitar que la competencia académica se transforme en competencia de infraestructura o acceso a la tecnología? El informe HEPI (Freeman, 2025) sugiere que muchos estudiantes desean más apoyo institucional y que persiste una brecha de uso por grupo socioeconómico. Un enfoque centrado en identidad moral no niega la necesidad de política institucional; al contrario, la identidad moral puede sostener una demanda ética: las instituciones deberían diseñar condiciones donde la elección moral sea razonable y no un acto heroico excepcional. Esto implica, por ejemplo, acceso equitativo a herramientas, capacitación universal y reglas claras.

| Propuestas para integrar el desarrollo de identidad moral en la educación superior

Integrar identidad moral no significa moralizar la universidad como a través de sermones, significa asumir que la formación profesional y ciudadana incluye la construcción de un yo capaz de actuar con responsabilidad en escenarios nuevos, complejos y tecnológicamente mediados. La evidencia de uso masivo de IAG y de ambigüedad normativa implica que la educación superior debe diseñar experiencias donde el uso ético de IA sea parte del aprendizaje, no un apéndice.

Una primera propuesta es tratar la ética del uso de IAG como parte de la identidad profesional de cada disciplina: investigación responsable en ciencias, integridad argumentativa en humanidades, confidencialidad y sesgos en salud, trazabilidad y seguridad en ingeniería, etc. UNESCO (2023) recomienda desarrollar competencias relacionadas con IA para aprendices y ajustar currículos y evaluación para mitigar riesgos. Esta curricularización debe evitar el error de abordar el tema en un solo curso, por ejemplo: una materia optativa de ética, desconectada del trabajo real. Si la identidad moral es una organización del yo en torno a rasgos y compromisos, se construye por repetición situada en las decisiones reales de investigación, escritura, diseño, colaboración, citación y divulgación.

La identidad moral también se fortalece cuando el estudiantado practica deliberación, responsabilidad y coherencia en casos concretos. Aquí se puede aprovechar la misma IAG como objeto de análisis crítico. A continuación se presentan algunas posibilidades para generar deliberación.

- Casos de zona gris: se pueden abordar los casos de zona gris, es decir, aquellos que no es tan claro que representen una deshonestidad o un problema ético, como lo son el parafraseo, uso de IA para estructurar ideas, generación de borradores, traducción, corrección de estilo, creación de referencias y detección de referencias inventadas. El objetivo no es solo dictaminar si esto está permitido o no permitido, sino reconstruir razones, haciendo preguntas como: ¿qué se pretende evaluar?, ¿qué se daña?, ¿qué se aprende?
- Diálogos de responsabilidad: diseñar actividades donde el estudiantado compare su razonamiento moral antes y después de usar IA, explicita riesgos y defina límites autoimpuestos.
- Reflexión narrativa: siguiendo enfoques que enfatizan la integración del yo, pedir a estudiantes elaborar una historia de aprendizaje sobre cómo han negociado presión, eficiencia y honestidad con tecnología. La idea de integración del yo moral y de regulación interna aparece como clave en discusiones contemporáneas del desarrollo de identidad moral (Lapsley, 2015).

Un aspecto más es que, dado que la IAG vuelve vulnerables productos finales como ensayos o reportes, muchas propuestas actuales recomiendan mover el foco hacia procesos, como borradores sucesivos, bitácoras, justificación de decisiones, defensas orales, y triangulación de evidencias de aprendizaje. En este sentido, la literatura sobre integridad académica e IAG destaca la necesidad de estrategias de evaluación más robustas que no dependan únicamente de la detección (Bittle y El-Gayar, 2025).

Probablemente la estrategia más congruente con identidad moral es incorporar declaraciones reflexivas de uso de IA, no con fines punitivos, sino formativos, en donde se explique qué herramienta se usó, para qué, qué partes se aceptaron o rechazaron, cómo se verificó, y cómo se preservó la autoría intelectual. Esta práctica convierte el uso de IA en un acto visible y deliberado, reduciendo la tentación de autoengaño y aumentando la coherencia entre el yo moral y la acción.

Otro aspecto relevante es que la identidad moral se activa o inhibe por señales del entorno. Si profesores prohíben IAG pero la usan sin transparencia, se crea una doble moral que erosiona el clima ético del aula. En HEPI (Freeman, 2025), una proporción relevante de estudiantes reporta ambigüedad sobre qué usos son aceptables, aun cuando declaran que las políticas son claras (80%). En la encuesta global del Digital Education Council (2024), la baja proporción que se siente plenamente consciente de directrices también apunta a un déficit de comunicación y apropiación institucional.

Por ello, una propuesta clave consiste en estructurar la formación docente en tres niveles complementarios. El primero corresponde a la alfabetización técnica y crítica en inteligencia artificial generativa, orientada no solo a desarrollar competencias para utilizar estas herramientas, sino también a comprender su funcionamiento, sesgos, limitaciones, riesgos de alucinaciones, implicaciones éticas y efectos sobre la privacidad, la propiedad intelectual y la integridad académica. El segundo nivel se centra en el rediseño de las estrategias de enseñanza y evaluación, promoviendo la creación de tareas auténticas, colaborativas y contextualizadas que incorporen la IA como un recurso para potenciar el aprendizaje, pero que al mismo tiempo exijan análisis, argumentación, creatividad, reflexión y toma de decisiones, reduciendo la posibilidad de sustituir el esfuerzo intelectual del estudiantado. Finalmente, el tercer nivel enfatiza el desarrollo de prácticas de transparencia y responsabilidad profesional, mediante la declaración explícita de cuándo, cómo y con qué propósito se utilizó la IA para elaborar clases, materiales didácticos, rúbricas, actividades o procesos de evaluación. Esta práctica favorece una cultura de apertura y uso ético de la tecnología, al tiempo que ofrece al estudiantado un referente concreto sobre cómo integrar la IA de manera responsable. Desde una perspectiva formativa, este modelamiento

docente trasciende la adopción de una buena práctica técnica, ya que constituye un mecanismo de socialización moral mediante el cual el profesorado facilita la comprensión de normas, valores y criterios éticos que orientan el comportamiento académico y profesional de las futuras generaciones.

| ***Prohibición e hipervigilancia frente a la propuesta de identidad moral***

Por último, se argumentará cómo la prohibición y la hipervigilancia no representan el camino más éticamente adecuado. Hay contextos en los que limitar usos es razonable, por ejemplo en evaluaciones de certificación profesional o exámenes presenciales; y hay situaciones donde investigar mala conducta es necesario. Pero el punto central es que prohibición y vigilancia, por sí solas, tienden a fallar como estrategia global por cuatro razones: practicidad, confiabilidad, justicia y formación.

Problema de practicidad: la IAG ya está integrada

Las cifras de uso identificadas en las encuestas que se han presentado con anterioridad, 92% en el estudio de la HEPI (Freeman, 2025) y 86% en el estudio de la Digital Education Council (2024), sugieren que la IAG se usa con o sin permiso. Cuando la conducta es masiva, la prohibición total suele desplazar el fenómeno a zonas opacas: aumenta la simulación y reduce la conversación pedagógica sobre límites y fines. Además, el análisis global sobre directrices académicas para IA y modelos de lenguaje realizado por Jiao et al. (2025) señala que, en 80 documentos revisados, no se observan directivas generales que prohíban la IAG en academia, salvo contextos donde el acceso está restringido por otras razones, y que las instituciones están desarrollando marcos balanceados más que vetos absolutos.

Problema de confiabilidad: la detección no es una base suficiente

La hipervigilancia suele apoyarse en detectores. Sin embargo, la evidencia disponible muestra que los proveedores y desarrolladores han reconocido limitaciones y errores de clasificación, incluyendo falsos positivos, en sistemas de detección. Las plataformas de uso masivo en instituciones reconocen que los falsos positivos son posibles y ajustan la presentación de resultados para evitar interpretaciones erróneas; y las investigaciones recientes discuten implicaciones éticas y confiabilidad cuestionable de herramientas de detección en educación superior (Bittle y El-Gayar, 2025; OpenAI, 2023; Turnitin, 2026). Si el sistema de control puede acusar injustamente, el costo moral para estudiantes inocentes es alto, provocando ansiedad, daño reputacional, desigualdad, por ejemplo, por el estilo de escritura o por hablar un idioma diferente al del contexto académico como lengua materna. En términos de clima moral, un régimen de sospecha permanente puede erosionar la confianza y, paradójicamente, reducir incentivos internos para actuar con integridad.

Problema de justicia: vigilancia que castiga más a quienes tienen menos poder

En la encuesta HEPI (Freeman, 2025), la preocupación por ser acusado de trampa aparece como freno relevante en 53% de las personas encuestadas, y las brechas socioeconómicas en uso sugieren que un sistema basado en castigo podría amplificar desigualdad, pues las y los estudiantes con menos capital cultural o recursos pueden estar más expuestos a sospecha o menos capaces de demostrar proceso.

Problema formativo: prohibición no enseña agencia responsable

El objetivo de la educación superior no es solo filtrar trampas; es formar profesionales capaces de actuar éticamente con herramientas poderosas. La IAG ya está entrando en el mercado laboral y en prácticas

profesionales; la universidad que solo prohíbe podría graduar estudiantes no preparados para decisiones reales, como cuándo usar automatización, cómo verificar, cómo declarar coautoría instrumental, cómo proteger datos y cómo anticipar impacto social. En contraste, un enfoque de identidad moral asume que el uso ético es una competencia de agencia que requiere deliberación, responsabilidad y coherencia. Su apuesta es que, si la moralidad se vuelve parte significativa del yo a través de la internalización, entonces el comportamiento ético puede sostenerse aun cuando la policía tecnológica no esté presente. La evidencia metaanalítica que vincula identidad moral con conducta moral y con emociones morales sugiere que esta dirección no es solo normativa: tiene base empírica (Hertz y Krettenauer, 2016).

| Conclusiones

La IAG introdujo una discontinuidad en la educación superior: transformó tareas, evaluación y relaciones de confianza en un lapso muy breve. Las estadísticas de adopción, 92% de uso reportado en un estudio nacional del Reino Unido (Freeman, 2025), 86% de uso regular en una encuesta global (Digital Education Council, 2024), indican que estamos ante un cambio estructural, no transitorio. En este escenario, las respuestas basadas exclusivamente en prohibición o hipervigilancia enfrentan límites prácticos y éticos: la detección no es plenamente confiable, pueden ocurrir falsos positivos, se erosiona el clima de confianza, y se pierde la oportunidad educativa de formar agencia responsable (Bittle y El-Gayar, 2025; OpenAI, 2023; Turnitin, 2026).

La teoría de la identidad moral ofrece un marco para reorientar el problema: el uso ético de IAG depende de que la honestidad, la responsabilidad epistémica, la justicia y el respeto por otros sean parte central de la autodefinición estudiantil. La evidencia de la psicología moral muestra asociaciones consistentes entre identidad moral y comportamiento moral, y entre identidad moral y emociones morales que sostienen compromisos prosociales (Hertz y Krettenauer, 2016; Jennings et al., 2015; Lefebvre y Krettenauer, 2019; Xu et al., 2024).

El desafío para la educación superior, entonces, no es escoger entre dejar usar sin restricciones o restringir, sino construir ecosistemas donde la regulación externa a través de políticas claras, evaluación robusta, gobernanza de datos esté alineada con regulación interna como fundamenta la teoría de la identidad moral. En un mundo donde la IAG reduce el costo de actuar de forma deshonesto, solo una formación que fortalezca la coherencia del yo moral puede sostener una integridad académica verdaderamente durable.

| Referencias

- Aquino, K., y Reed II, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83(6), 1423–1440. <https://doi.org/10.1037/0022-3514.83.6.1423>
- Bittle, K., y El-Gayar, O. (2025). Generative AI and academic integrity in higher education: A systematic review and research agenda. *Information*, 16(4), 296. <https://doi.org/10.3390/info16040296>
- Blasi, A. (1984). Moral identity: Its role in moral functioning. En W. M. Kurtines y J. L. Gewirtz (Eds.), *Morality, moral behavior, and moral development* (pp. 128–139). Wiley.

- Blasi, A. (1993). The development of identity: Some implications for moral functioning. En G. Noam y T. Wren (Eds.), *The moral self* (pp. 99–122). MIT Press.
- Boegershausen, J., Aquino, K., y Reed II, A. (2015). Moral identity. *Current Opinion in Psychology*, 6, 162–166. <https://doi.org/10.1016/j.copsyc.2015.07.017>
- Deep, P. D., Edgington, W. D., Ghosh, N., y Rahaman, M. S. (2025). Evaluating the Effectiveness and Ethical Implications of AI Detection Tools in Higher Education. *Information*, 16(10), 905. <https://doi.org/10.3390/info16100905>
- Deschenes, A., y McMahon, M. (2024). A survey on student use of generative AI chatbots for academic research. *Evidence Based Library and Information Practice*, 19(2), 2–22. <https://doi.org/10.18438/eblip30512>
- Digital Education Council. (2024). *Digital Education Council Global AI Student Survey 2024. AI or Not AI: What Students Want*. <https://www.digitaleducationcouncil.com/form/global-ai-student-survey-2024>
- Digital Education Council. (2026). *Encuesta sobre la IA en la educación superior en América Latina 2026*. Digital Education Council, Tecnológico de Monterrey e Institute for the Future of Education. <https://www.digitaleducationcouncil.com/post/ai-in-higher-education-latam-survey-2026>
- Emsley, R. (2023). ChatGPT: These are not hallucinations—they’re fabrications and falsifications. *Schizophrenia*, 9(1), 52. <https://www.nature.com/articles/s41537-023-00379-4>
- Erikson, E. H. (1968). *Identity: Youth and crisis*. W. W. Norton.
- Freeman, J. (2025). *Student Generative AI Survey 2025* (HEPI Policy Note 61). Higher Education Policy Institute. <https://www.hepi.ac.uk/wp-content/uploads/2025/02/HEPI-Kortext-Student-Generative-AI-Survey-2025.pdf>
- Hertz, S. G., y Krettenauer, T. (2016). Does moral identity effectively predict moral behavior? A meta-analysis. *Review of General Psychology*, 20(2), 129–140. https://scholars.wlu.ca/psyc_faculty/112/
- Huang, D., Hash, N., Cummings, J. J., y Prena, K. (2025). Academic cheating with generative AI: Exploring a moral extension of the theory of planned behavior. *Computers and Education: Artificial Intelligence*, 8, 100424. <https://doi.org/10.1016/j.caeai.2025.100424>
- International Center for Academic Integrity. (2025). *Fundamental Values*. <https://academicintegrity.org/aws/ICAI/pt/sp/values>
- Jennings, P. L., Mitchell, M. S., y Hannah, S. T. (2015). The moral self: A review and integration of the literature. *Journal of Organizational Behavior*, 36(S1), S104–S168. <https://doi.org/10.1002/job.1919>
- Jiao, J., Afroogh, S., Chen, K., Atkinson, D., y Dhurandhar, A. (2025). The global landscape of academic guidelines for generative AI and LLMs. *Nature Human Behaviour*, 9, 638–642. <https://doi.org/10.1038/s41562-025-02124-6>

- Lapsley, D. (2015). Moral identity and developmental theory (Commentary on Krettenauer and Hertz). *Human Development*, 58, 164–171. <https://doi.org/10.1159/000435926>
- Lefebvre, J. P., y Krettenauer, T. (2019). Linking moral identity with moral emotions: A meta-analysis. *Review of General Psychology*, 23, 444–457. https://scholars.wlu.ca/bspsy_faculty/4/
- Maleki, N., Padmanabhan, B., y Dutta, K. (2024, junio). AI hallucinations: A misnomer worth clarifying. En *2024 IEEE conference on artificial intelligence (CAI)* (pp. 133–138). IEEE. <https://ieeexplore.ieee.org/abstract/document/10605268>
- Marcia, J. E. (1966). Development and validation of ego-identity status. *Journal of Personality and Social Psychology*, 3(5), 551–558. <https://doi.org/10.1037/h0023281>
- McAdams, D. P. (1993). *The stories we live by: Personal myths and the making of the self*. Guilford Press. <https://archive.org/details/storieswelivebyp0000mcad/page/8/mode/2up>
- Miles, P. J., Campbell, M., y Ruxton, G. D. (2022). Why students cheat and how understanding this can help reduce the frequency of academic misconduct in higher education: A literature review. *Journal of Undergraduate Neuroscience Education: JUNE: A Publication of FUN, Faculty for Undergraduate Neuroscience*, 20(2), A150–A160. <https://doi.org/10.59390/LXMJ2920>
- Narvaez, D., y Lapsley, D. K. (2009). Moral identity, moral functioning, and the development of moral character. *Psychology of Learning and Motivation*, 50, 237–274.
- National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1).
- OpenAI. (2023, enero). *Nuevo clasificador de IA para indicar texto escrito con IA*. <https://openai.com/es-419/index/new-ai-classifier-for-indicating-ai-written-text/>
- Salvagno, M., Taccone, F. S., y Gerli, A. G. (2023). Artificial intelligence hallucinations. *Critical Care*, 27(1), 180. <https://link.springer.com/article/10.1186/s13054-023-04473-y>
- Secretaría de Educación Pública. (2025). *Usos y percepciones sobre la inteligencia artificial generativa en la educación superior mexicana: Encuesta Nacional (ENIAG 2025)*. Secretaría de Educación Pública. https://www.gob.mx/cms/uploads/attachment/file/1071411/Dossier_Resultados_ENIAG2025_SEP_.pdf
- Turnitin. (2026). *Using the AI Writing Report*. <https://guides.turnitin.com/hc/en-us/articles/22774058814093-Using-the-AI-Writing-Report>
- UNESCO. (2023). *Guía para el uso de IA generativa en educación e investigación* (F. Miao y W. Holmes, autores). <https://doi.org/10.54675/EWZM9535>
- Xu, X., Li, M., Kwan, H. K., y Zhang, X. (2024). The antecedents of moral identity: A meta-analytic review. *Asia Pacific Journal of Management*, 41, 1697–1728. <https://doi.org/10.1007/s10490-023-09891-8>