Artículo de investigación científico tecnológica

Cómo citar: Ventura Ramos, P. E., Memije Alarcón, N. Y., & Zaragoza Martínez, J. (2026). Falsos positivos de Turnitin y su gestión en la Licenciatura en Derecho de la Universidad Autónoma de Guerrero. *Estrategia y Gestión Universitaria*, 14, e9232. https://doi.org/10.5281/zenodo.23671197

Recibido: 01/09/2026
Aceptado: 16/10/2026
Publicado: 02/11/2026

Autor para correspondencia: ventura-eliza31@hotmail.com

Conflicto de intereses: los autores declaran no tener ningún conflicto de intereses, que puedan haber influido en los resultados obtenidos o las interpretaciones propuestas.

Perla Elisabeth Ventura Ramos [1]
Universidad Autónoma de Guerrero
https://orcid.org/0000-0001-8680-1703
ventura-eliza31@hotmail.com
México

Norma Yadira Memije Alarcón [2]
Universidad Autónoma de Guerrero
https://orcid.org/0000-0002-3402-1112
profesmx2023@gmail.com
México

Jesús Zaragoza Martínez [3]
Universidad Autónoma de Guerrero
https://orcid.org/0000-0003-3685-7578
jesusmar63@msn.com
México

Falsos positivos de Turnitin y su gestión en la Licenciatura en Derecho de la Universidad Autónoma de Guerrero

Turnitin false positives and their management in the Law Degree Program at the Autonomous University of Guerrero

Falsos positivos do Turnitin e sua gestão no Curso de Direito da Universidade Autônoma de Guerrero

Resumen

Introducción: el uso de detectores de inteligencia artificial como Turnitin se ha vuelto común en instituciones de educación superior para evaluar la originalidad de trabajos académicos. Sin embargo, estas herramientas presentan limitaciones importantes, especialmente en la identificación de textos escritos por humanos. **Objetivo:** analizar los falsos positivos en la detección de IA con Turnitin en la Licenciatura en Derecho de la Universidad Autónoma de Guerrero. **Método:** enfoque mixto que incluye encuestas, análisis documental de informes de similitud y entrevistas a docentes. **Resultados:** entre el 10% y 20% de los textos 100% humanos fueron marcados erróneamente como generados por IA. Algunas dependencias educativas utilizan estos porcentajes como único criterio para sancionar estudiantes o rechazar trabajos. **Conclusión:** los detectores de IA no deben usarse como única evidencia de mala conducta académica y se propone un marco de evaluación integral que combine revisión humana, procesos de apelación y evaluación alternativa. Este estudio aporta evidencia empírica desde el contexto mexicano y latinoamericano para informar políticas institucionales más justas y efectivas.

Palabras clave: Turnitin, inteligencia artificial, falsos positivos, educación superior, integridad académica

Abstract

Introduction: the use of artificial intelligence detection tools such as Turnitin has become common in higher education institutions to assess the originality of academic work. However, these tools present significant limitations, particularly in identifying texts written by humans.

Este es un artículo Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Objective: to analyze false positives in AI detection by Turnitin in the Law Degree Program at the Autonomous University of Guerrero. **Method:** a mixed-methods approach was employed, including surveys, documentary analysis of similarity reports, and interviews with faculty members. **Results:** between 10% and 20% of entirely human-written texts were incorrectly flagged as AI-generated. Some educational institutions use these percentages as the sole criterion for sanctioning students or rejecting academic work. **Conclusion:** AI detectors should not be used as the sole evidence of academic misconduct. A comprehensive evaluation framework is proposed that combines human review, appeal procedures, and alternative assessment methods. This study provides empirical evidence from the Mexican and Latin American context to inform fairer and more effective institutional policies.

Keywords: Turnitin, artificial intelligence, false positives, higher education, academic integrity

Resumo

Introdução: o uso de detectores de inteligência artificial, como o Turnitin, tornou-se comum nas instituições de ensino superior para avaliar a originalidade de trabalhos acadêmicos. No entanto, essas ferramentas apresentam limitações importantes, especialmente na identificação de textos escritos por seres humanos. **Objetivo:** analisar os falsos positivos na detecção de IA pelo Turnitin no curso de Direito da Universidade Autônoma de Guerrero. **Método:** foi adotada uma abordagem de métodos mistos, incluindo questionários, análise documental de relatórios de similaridade e entrevistas com docentes. **Resultados:** entre 10% e 20% dos textos totalmente produzidos por humanos foram incorretamente classificados como gerados por IA. Algumas unidades educacionais utilizam esses percentuais como único critério para sancionar estudantes ou rejeitar trabalhos acadêmicos. **Conclusão:** os detectores de IA não devem ser utilizados como única evidência de má conduta acadêmica. Propõe-se um modelo de avaliação abrangente que combine revisão humana, processos de recurso e formas alternativas de avaliação. Este estudo fornece evidências empíricas a partir do contexto mexicano e latino-americano para subsidiar políticas institucionais mais justas e eficazes.

Palavras-chave: Turnitin, inteligência artificial, falsos positivos, ensino superior, integridade acadêmica



Introducción

Los procesos de gestión universitaria afrontan en el día de hoy el reto de integrar las nuevas tecnologías de la inteligencia artificial (IA) a los procesos de evaluación docente y científica en cada comunidad universitaria. Turnitin es un instrumento elaborado para avalar la singularidad de los trabajos y los estudios superiores desde hace un tiempo los han tomado para la revisión de los mismos. De acuerdo con la visión de un educador, según Arteaga (2023), la inteligencia artificial generativa se utiliza en la tutoría personalizada, en la evaluación automática, en la adaptación de materiales y en la identificación de plagio, siendo este último un uso significativo. Todo esto ocurre en un contexto donde existen desigualdades como: falta de infraestructura, niveles de alfabetización, manejo de datos y consideraciones éticas. Esta situación se observa como un desafío frecuente en las instituciones educativas.

La veloz incorporación de herramientas de IA en la educación superior de América Latina ha transformado los métodos de enseñanza y aprendizaje de los alumnos. En las instituciones educativas, ha beneficiado ámbitos académicos como el rendimiento y el desarrollo de habilidades digitales (González Torres et al., 2024; Tiglla, 2025; Cándano Pérez et al., 2025). En ocasiones se producen los llamados falsos positivos que afectan la seriedad y credibilidad de los resultados investigativos de estudiantes y profesores. Desde la gestión de procesos académicos la situación es mucho peor cuando las instituciones utilizan los porcentajes de detección de IA o de similitud como el criterio para castigar a los autores o simplemente rechazar el trabajo. La utilización de detectores de inteligencia artificial, como Turnitin, se incrementa continuamente en el ámbito de la administración universitaria, con el objetivo de asegurar la autenticidad de los documentos académicos. No obstante, estas herramientas tienen limitaciones significativas, sobre todo en cuanto a la detección de contenidos elaborados por personas. Para Alnaimat et al. (2025) diversos equipos de editores desde hace algún tiempo implementan intervenciones de carácter éticos interno basados en inteligencia artificial para proveer toda la documentación inicial, contar con más evaluadores y sobre todo descubrir potenciales plagios mucho antes de que se revisen por pares los manuscritos, con el objetivo de que sean mirados por las personas.

A partir de dos investigaciones de tipo experimental con dos grupos de profesores, tanto de experiencia vasta como noveles, ninguno logró detectar de manera segura textos que estuvieran copiados desde la IA (Fleckenstein et al., 2024). Estos autores exponen con muy pequeñas instrucciones la IA origina contenidos difíciles de identificar poniendo en tela de juicio evaluaciones en cualquier contexto educativo.

Turnitin (2023c) señala que la evaluación de producción escrita mediante IA tiene como objetivo asistir a los docentes en la detección de textos que podrían haber sido creados por una herramienta generativa de IA y no se debe considerar como la única base para tomar decisiones negativas sobre un alumno. Es esencial un análisis más profundo y el juicio humano, junto con la implementación de las normativas académicas específicas de una institución, para establecer si hubo alguna falta de conducta académica.

Cuando se va a elegir qué detector utilizar es muy importante reconocer que

el uso de múltiples al unísono puede conducir a la toma de decisiones más serias y concisas, especialmente cuando se trata de la educación superior donde el lenguajes es muchas veces técnico y puede conducir a falsos positivos (Hyatt et al., 2025). Los falsos positivos provocan conflictos en el ámbito académico y ponen en duda la efectividad de las políticas institucionales que dependen únicamente de estos sistemas. El procedimiento de evaluación en el ámbito académico científico presenta el dilema de cómo medir la originalidad en los trabajos académicos, dado que tanto estudiantes como profesores cuentan con acceso a herramientas de inteligencia artificial que son capaces de generar textos completos en cuestión de segundos.

Annie Chechitelli, Directora de Producto en Turnitin, indica que se ha observado que en aquellos documentos donde se identifica menos del 20% de contenido generado por inteligencia artificial, se produce una mayor frecuencia de falsos positivos (Turnitin, 2023a).

En México, no se puede pasar por alto el caso de organizaciones como la Universidad Autónoma de Guerrero. Específicamente, la Facultad de Derecho ha adoptado el uso de Turnitin como un recurso para identificar potenciales casos de plagio y, más recientemente, para reconocer contenidos producidos por inteligencia artificial. Este es un sistema que analiza los trabajos presentados por alumnos y profesores en relación con una extensa base de datos compuesta de documentos académicos, sitios web y artículos.

Para Pellegrina & Helmy (2025) en el campo de la honestidad científica, existe un resumen de las habilidades y limitaciones de los detectores de texto generados por inteligencia artificial. Se identifican debilidades en su rendimiento, sensibilidad a distintos tipos de texto y susceptibilidad a ataques de reescritura o situaciones adversas. Esto se da en el contexto de recientes esfuerzos editoriales enfocados en identificar "fábricas de artículos" y plataformas colaborativas, así como servicios específicos para la modificación de imágenes.

Desde el año 2023, esta herramienta ha añadido un nuevo indicador que revela el porcentaje del texto que "posiblemente" fue creado por inteligencia artificial. De acuerdo con información proporcionada por la misma empresa, este detector es capaz de reconocer más del 98% de los textos producidos por ChatGPT y cuenta con menos del 1% de resultados falsos positivos. Aunque Turnitin mismo señala que sus informes deben ser entendidos como señales y no como evidencias definitivas (Turnitin, 2024).

Según Navarrete-Cazales y Cruz-Victoria (2026), la incorporación de inteligencia artificial en el ámbito educativo aborda la tensión entre progreso y deshonestidad. Por un lado, el uso de IA puede ser visto como similar a la utilización de otras herramientas permitidas (tutores online, software de corrección, calculadoras avanzadas); por otro lado, si la inteligencia artificial se encarga de realizar aspectos importantes del trabajo intelectual, podría ser considerada una forma de engaño. En los años 2023 y 2024 se realizó una investigación sobre estudios de 100 textos sobre neurocirugía (Schneider et al., 2025) que fueron publicados por revistas técnicas, encontrándose que uno de cada cinco presentaba secciones identificadas como elaboradas por inteligencia artificial. Los apartados que más frecuentemente presentaban esta peculiaridad son las de resúmenes y métodos.

La circunstancia ha puesto de manifiesto zonas ambiguas: por ejemplo, ¿es incorrecto que un alumno recurra a ChatGPT para obtener aclaraciones adicionales sobre un tema complicado? Seguramente no, de hecho, podría considerarse beneficioso como un apoyo al aprendizaje. ¿Y qué hay de emplearlo para producir la versión inicial de un ensayo, que posteriormente el alumno modifique y enriquezca con citas? Algunos lo considerarían como una forma de colaboración o ayuda legítima, mientras que otros lo criticarían por no derivar completamente del estudiante. El uso de la inteligencia artificial no implica automáticamente un engaño o una copia. Estas situaciones ambiguas demuestran que el juicio depende, ante todo, del propósito y del método aplicado.

Explican Murdock y Tadinada (2025) que, a pesar de la creciente confianza en la tecnología, las pruebas recientes indican que, si bien las herramientas de detección de IA demuestran una precisión moderada a alta en la distinción entre el texto generado por IA y la escritura humana, siguen siendo poco fiables y frecuentemente producen falsos positivos que afectan negativamente tanto a los estudiantes como a los educadores. En el área de las publicaciones científicas, ciertos detectores que se toman por más “fieles” revelan un sesgo más alto contra determinados grupos y en disciplinas específicas, lo que influye decididamente en el proceso de toma de decisiones en entornos educativos y de publicación (Pratama, 2025).

Plantea Khan et al. (2025) que un elemento que dificulta la detección precisa y exacta de los porcientos de similitud y la utilización de IA son las herramientas de parafraseo y de humanización, al igual que el tamaño del grupo de datos del texto en cuestión. Por su parte otros estudiosos como Hua y Yao (2024) aseguran que hay ciertos modelos diseñados específicamente para la identificación y atribución, los cuales exhiben un rendimiento notable en conjuntos de datos sometidos a control, pero que las herramientas de reescritura y personalización reducen de manera significativa la eficacia en la detección.

Para Weber et al. (2023), estos instrumentos generan restricciones concluyentes, tales como textos humanos traducidos por la máquina, o textos hechos con IA y editados por personas, así como parafraseados (híbridos).

¿Qué implicaciones para la gestión universitaria traen consigo los falsos positivos detectados por Turnitin?

Tras revelar esto, la universidad tiene la obligación de implementar protocolos de apelación para proteger a su gente. Es necesario que siempre haya una revisión humana y que quede claro que ningún detector puede servir como prueba única de deshonestidad académica. Hay que asegurar que exista una revisión colegiada y mecanismos de apelación donde tanto alumnos como profesores puedan aportar pruebas extras. Es necesario que los profesores aprendan hasta dónde llegan los programas de detección de plagio. Si solo confían en estas herramientas, el resultado es impreciso. En su lugar, hay que apostar por otras formas de evaluar, como las entrevistas, el examen de documentos o el uso de rúbricas detalladas. Con este cambio, los estudiantes y los académicos tendrán un entorno de mayor confianza para trabajar sin vivir bajo la tensión de un algoritmo, lo que garantiza tanto la integridad de las investigaciones como el bienestar dentro de la universidad.

En 2026, varios estudios realizados en universidades mexicanas demostraron que a casi uno de cada cuatro ensayos escritos por personas se les etiquetó mal como si fueran obra de una IA. Incluso Turnitin, que es de lo más fiable que hay para detectar esto, tuvo un 11% de falsos positivos en documentos formales. Chávez-Martínez (2026) señala que, desde un punto de vista ético, hay que tener claro que la IA no puede ser autora de un trabajo académico. Básicamente, porque le falta responsabilidad moral, intención cognitiva y esa capacidad de juicio crítico que sí tienen los seres humanos. Al final, la autoría científica va de la mano con la verdad, saber interpretar las cosas y hacerse responsable de lo que se escribe, y esas son cualidades que solo los seres humanos tienen.

Un estudio publicado en Education and Information Technologies (Springer, 2026) deja claro que Turnitin todavía tiene problemas para distinguir bien entre lo que escribe una persona y lo que genera una IA. A veces, aunque el texto sea casi todo de autoría humana, el sistema marca trozos como si fueran artificiales. Esto provoca bastantes falsos positivos y hace que muchos se pregunten si realmente es fiable usar este detector en la universidad.

En varias universidades de México se vio que casi uno de cada cuatro ensayos escritos por personas fue marcado por error como si fueran IA. Por otro lado, Turnitin es algo más preciso, pero aun así llega a tener un 11 por ciento de falsos positivos en trabajos académicos. Todo esto demuestra lo difícil que es detectar textos en español, sobre todo cuando se trata de lenguaje jurídico o técnico (Trejo, 2026).

Al igual que en los estudios hechos en Estados Unidos y Europa, también se han reportado tasas de falsos positivos. Las investigaciones en contextos anglosajones indican que Turnitin tuvo entre un 8 y un 12% de errores en trabajos universitarios. En el caso de las universidades chinas, los detectores llegaron a marcar hasta un 20% de falsos positivos en textos académicos que habían sido traducidos al inglés. Ciertas disciplinas, como el derecho, sufren este inconveniente con mayor intensidad. El lenguaje empleado allí es técnico, rígido y sigue estándares muy marcados. Resulta casi inevitable repetir las menciones a leyes, códigos y artículos de la Constitución.

Los razonamientos se levantan casi siempre sobre la misma base: premisa, argumento y conclusión. Por eso, los alumnos recién llegados a la carrera escriben con tal pesadez que sus trabajos saltan a la vista de cualquier filtro automático. Están aprendiendo a encajar las piezas en un molde demasiado estrecho.

En la Universidad Autónoma de Guerrero, los estudiantes de la carrera de Derecho envían regularmente sus trabajos a la herramienta Turnitin. Los profesores e investigadores universitarios envían artículos y capítulos de libros a diversas revistas y editores para su publicación y envían muestras aleatorias de estos trabajos a la herramienta Turnitin. Si la herramienta califica estos trabajos como escritos parcialmente por IA, algunos educadores utilizan esos porcentajes para castigar a los estudiantes, rechazar trabajos o cuestionar la reputación de los docentes.

Esta investigación busca documentar qué está pasando en la Facultad de Derecho de la Universidad Autónoma de Guerrero. Entre otros cuestionamientos se busca responder: ¿Cuántos estudiantes de primer grado han sido marcados

erróneamente como usuarios de IA? ¿Qué características tienen los textos que generan falsos positivos? ¿Cómo están usando los docentes y las dependencias educativas los resultados de Turnitin?

El objetivo general de esta investigación es analizar los falsos positivos en la detección de IA con Turnitin en estudiantes y docentes de la Licenciatura en Derecho de la Universidad Autónoma de Guerrero. Los objetivos específicos son: (1) identificar la prevalencia de falsos positivos en los 58 estudiantes de primer grado; (2) documentar casos de docentes investigadores cuyos trabajos han sido marcados erróneamente; (3) analizar las características de los textos que generan falsos positivos; (4) evaluar cómo están usando los docentes y dependencias educativas los resultados de Turnitin.

Materiales y métodos

Esta investigación utilizó un enfoque mixto, combinando métodos cuantitativos y cualitativos para obtener una visión completa del problema. El diseño fue descriptivo, ya que se describe la prevalencia de falsos positivos en la detección de IA con Turnitin y se detalla cómo se relaciona este fenómeno con características específicas de los textos. El enfoque mixto permite, por un lado, cuantificar la frecuencia de falsos positivos mediante el análisis de informes de similitud de Turnitin. Por otro, comprender las experiencias, percepciones y prácticas de estudiantes y docentes a través de encuestas y entrevistas.

El estudio se realizó en la Facultad de Derecho de la Universidad Autónoma de Guerrero, ubicada en Chilpancingo (UAGro), Guerrero, México. La UAGro es una institución pública de educación superior que atiende a estudiantes en diversas disciplinas, incluyendo las especialidades de derecho, medicina, ingeniería, ciencias sociales y humanidades. La Facultad de Derecho ofrece la Licenciatura en Derecho, que tiene una duración de nueve semestres y forma profesionales capaces de ejercer en diversas áreas del ámbito jurídico. Los estudiantes de primer grado (primer semestre) están cursando materias introductorias como Introducción al Derecho, Derecho Romano, Historia del Derecho, y materias de formación general.

En cuanto al uso de Turnitin, la Universidad Autónoma de Guerrero cuenta con licencia institucional para esta plataforma. Los docentes pueden subir trabajos de estudiantes para generar reportes de similitud, y algunas dependencias educativas requieren este reporte como parte del proceso de evaluación de tesis, artículos y trabajos de grado. Sin embargo, no hay información pública detallada sobre las políticas específicas de uso de Turnitin para detección de IA en la Facultad de Derecho.

La población de estudio estuvo conformada por dos grupos:

Tabla 1

Población de estudio

GRUPO 1	POBLACIÒN	MUESTRA	CRITERIO DE INCLUSIÒN
Estudiantes de primer grado	58	58	Estar inscrito en primer grado de la Licenciatura en Derecho durante el ciclo escolar 2024-2026 y haber presentado al menos un trabajo evaluado con Turnitin.
GRUPO 2			
Docentes investigadores	25	15	Ser docente de tiempo completo o medio tiempo en la Facultad de Derecho, realizar investigación activa y haber presentado al menos un trabajo (artículo, capítulo de libro, ponencia) evaluado con Turnitin entre 2024-2026.

Fuente: Elaboración propia.

La decisión de incluir a todos los 58 estudiantes de primer grado se tomó porque es una población pequeña y accesible, lo que permitió obtener datos completos sin necesidad de muestreo. En el caso de los docentes, se utilizó un muestreo intencional por conveniencia, seleccionando a aquellos que estaban disponibles y dispuestos a participar en el estudio.

Instrumentos de recolección de datos

- Encuesta para estudiantes

Se diseñó una encuesta estructurada con 12 preguntas cerradas y 2 preguntas abiertas. Las preguntas cerradas incluían opciones de respuesta múltiple, escala de Likert (de "muy en desacuerdo" a "muy de acuerdo") y preguntas de sí/no. Las preguntas abiertas permitían a los estudiantes describir sus experiencias con Turnitin con sus propias palabras.

Contenido de la encuesta:

- Datos demográficos (edad, sexo, municipio de origen)
- Experiencia previa con herramientas de IA (ChatGPT, Gemini, Copilot, etc.)
- Número de trabajos presentados con Turnitin en el semestre
- Porcentajes de IA reportados por Turnitin en sus trabajos
- Autoría real de los trabajos (100% humano, uso parcial de IA, uso total de IA)

- Reacciones de los docentes ante los porcentajes de IA
- Conocimiento sobre las limitaciones de Turnitin
- Experiencias de cuestionamiento o sanción por porcentajes de IA
- Conocimiento sobre procesos de apelación
- Percepciones sobre la justicia del sistema actual
- Preguntas abiertas sobre experiencias específicas y sugerencias de mejora

La encuesta se aplicó de manera presencial en el aula, con autorización de los docentes y consentimiento informado de los estudiantes. El tiempo promedio de respuesta fue de 15-20 minutos.

- Análisis documental de informes de Turnitin

Se solicitaron a los estudiantes y docentes copias de los informes de similitud de Turnitin de sus trabajos recientes. Se recopilaron un total de 127 informes: 89 de estudiantes y 38 de docentes.

Para cada informe, se registraron los siguientes datos:

- Tipo de trabajo (ensayo, examen, proyecto parcial, tesis, artículo, etc.)
- Longitud del documento (número de palabras)
- Porcentaje de similitud
- Porcentaje de IA
- Fecha de entrega
- Materia o contexto de elaboración
- Calificación obtenida (si estaba disponible)
- Comentarios del docente (si estaban disponibles)

Además, para verificar la autoría real, se solicitó a los participantes evidencia de su proceso de escritura: borradores previos, historial de edición en Google Docs o Word, notas manuscritas, o cualquier otro documento que demostrara el proceso de elaboración del trabajo. Esto fue crucial para identificar casos de falsos positivos con certeza.

Se realizó una revisión documental de las políticas, reglamentos y lineamientos de la Universidad Autónoma de Guerrero y de la Facultad de Derecho y el procedimiento de búsqueda está relacionado con:

- Integridad académica
- Uso de herramientas de detección de plagio
- Evaluación de trabajos estudiantiles
- Procesos de sanción por mala conducta académica
- Uso de inteligencia artificial en trabajos académicos.

Se consultaron los siguientes documentos:

- Reglamento General de la Universidad Autónoma de Guerrero
- Reglamento de Integridad Académica (si existe)
- Lineamientos de la Facultad de Derecho para evaluación de trabajos
- Políticas de uso de Turnitin (si existen)
- Comunicados oficiales sobre uso de IA (si existen)

También se consultaron documentos similares de otras universidades mexicanas para comparar y contextualizar los hallazgos.

- Entrevistas semiestructuradas a docentes

Se realizaron entrevistas semiestructuradas a 15 docentes investigadores de la Facultad de Derecho. Las entrevistas duraron entre 30 y 45 minutos cada una y se grabaron con consentimiento de los participantes para su posterior transcripción y análisis.

Temas abordados en las entrevistas:

- Experiencia personal con Turnitin (como docente y como autor)
- Criterios utilizados para evaluar la originalidad de los trabajos estudiantiles
- Uso que hacen de los porcentajes de similitud y de IA de Turnitin
- Casos conocidos de falsos positivos o falsos negativos
- Conocimiento sobre las limitaciones de Turnitin
- Políticas o lineamientos de su dependencia sobre uso de Turnitin
- Procesos de sanción o apelación cuando hay sospecha de uso indebido de IA
- Percepciones sobre la justicia y efectividad del sistema actual
- Sugerencias de mejora

Las entrevistas se realizaron de manera presencial o virtual, según la preferencia de cada docente. Todas las transcripciones fueron anonimizadas para proteger la identidad de los participantes.

Procedimiento

El trabajo de campo se realizó entre enero y junio de 2026, siguiendo las siguientes etapas:

Etapa 1: Preparación (enero 2026)

- Diseño de instrumentos (encuesta, guía de entrevista, formato de análisis documental)
- Validación de instrumentos con expertos en metodología de la investigación

- Solicitud de permisos institucionales a la Facultad de Derecho
- Elaboración de formatos de consentimiento informado

Etapa 2: Aplicación de encuestas (febrero-marzo 2026)

- Visita a las aulas de primer grado de la Licenciatura en Derecho
- Explicación del propósito del estudio a los estudiantes
- Aplicación de la encuesta de manera presencial
- Recolección de consentimientos informados firmados

Etapa 3: Recolección de informes de Turnitin (marzo-abril 2026)

- Solicitud a estudiantes y docentes de copias de informes de Turnitin
- Verificación de autoría mediante revisión de borradores y evidencias de proceso
- Registro de datos en base de datos Excel

Etapa 4: Entrevistas a docentes (abril-mayo 2026)

- Contacto y agendamiento de entrevistas con docentes
- Realización de entrevistas (presenciales o virtuales)
- Transcripción y anonimización de entrevistas

Etapa 5: Revisión documental (mayo 2026)

- Búsqueda y análisis de políticas institucionales
- Comparación con políticas de otras universidades mexicanas

Etapa 6: Análisis de datos (junio 2026)

- Análisis estadístico descriptivo de datos cuantitativos
- Análisis de contenido de datos cualitativos (entrevistas y preguntas abiertas)
- Triangulación de datos cuantitativos y cualitativos

Análisis de datos

Datos cuantitativos: Los datos de las encuestas y de los informes de Turnitin se analizaron utilizando estadística descriptiva. Se calcularon frecuencias, porcentajes, medias y desviaciones estándar para las variables principales. Se utilizó el software SPSS versión 28 para el análisis estadístico.

Las variables principales analizadas fueron:

- Porcentaje de IA reportado por Turnitin
- Autoría real verificada (humano/IA/híbrido)
- Tipo de trabajo (ensayo, examen, proyecto, tesis, artículo)
- Longitud del documento

- Calificación obtenida
- Reacción del docente (cuestionamiento, sanción, sin acción)

Datos cualitativos: Las entrevistas y las preguntas abiertas de las encuestas se analizaron mediante análisis de contenido temático. Se identificaron temas recurrentes, patrones y categorías emergentes. Se utilizó el software Atlas.ti versión 9 para la codificación y organización de los datos cualitativos.

Los temas principales identificados fueron:

- Experiencias con falsos positivos
- Percepciones sobre la justicia del sistema
- Criterios de evaluación de los docentes
- Sugerencias de mejora

Se realizó triangulación de datos comparando los resultados cuantitativos con los cualitativos, y comparando las perspectivas de estudiantes con las de docentes. Esto permitió obtener una visión más completa y validar los hallazgos desde múltiples fuentes. Esta investigación siguió los principios éticos básicos para la investigación con seres humanos: Consentimiento informado, Confidencialidad, Uso de datos, Aprobación institucional y No maleficencia.

Resultados y discusión

Características de la muestra

Estudiantes:

De los 58 estudiantes de primer grado, 34 eran mujeres (58.6%) y 24 eran hombres (41.4%). La edad promedio fue de 19.2 años (rango: 17-24 años). La mayoría provenía de Chilpancingo y municipios cercanos (75.9%), mientras que el resto de otras regiones de Guerrero (24.1%).

En cuanto al acceso a tecnología, 52 estudiantes (89.7%) reportaron tener computadora personal en casa, y 56 (96.6%) reportaron tener acceso a internet. Sin embargo, solo 23 estudiantes (39.7%) reportaron haber usado herramientas de IA como ChatGPT antes de ingresar a la universidad.

Docentes:

De los 15 docentes entrevistados, 9 eran hombres (60%) y 6 eran mujeres (40%). La experiencia docente promedio fue de 12.4 años (rango: 3-28 años). Todos los docentes tenían al menos una publicación en revista o libro, y 11 de ellos (73.3%) tenían grado de doctorado.

Prevalencia de falsos positivos

Estudiantes: de los 89 informes de Turnitin analizados de estudiantes, 67 correspondían a trabajos que fueron verificados como 100% humanos (mediante

revisión de borradores, historiales de edición y entrevistas). De estos 67 trabajos 100% humanos, 12 (17.9%) tuvieron un porcentaje de IA superior al 20% en Turnitin, lo que se considera un falso positivo según los criterios de este estudio.

Distribución de porcentajes de IA en trabajos 100% humanos:

- 0-5%: 32 trabajos (47.8%)
- 6-10%: 15 trabajos (22.4%)
- 11-20%: 8 trabajos (11.9%)
- 21-30%: 6 trabajos (9.0%)
- 31-40%: 4 trabajos (6.0%)
- 41-50%: 2 trabajos (3.0%)
- Más de 50%: 0 trabajos (0%)

El porcentaje promedio de IA en trabajos 100% humanos fue de 11.3% (desviación estándar: 12.8%). El porcentaje máximo fue de 47%, correspondiente a un ensayo de Introducción al Derecho que fue escrito completamente por la estudiante sin uso de IA. Los estudiantes manifestaron sentimientos de frustración y desconfianza hacia el sistema, además de temor a sanciones injustas.

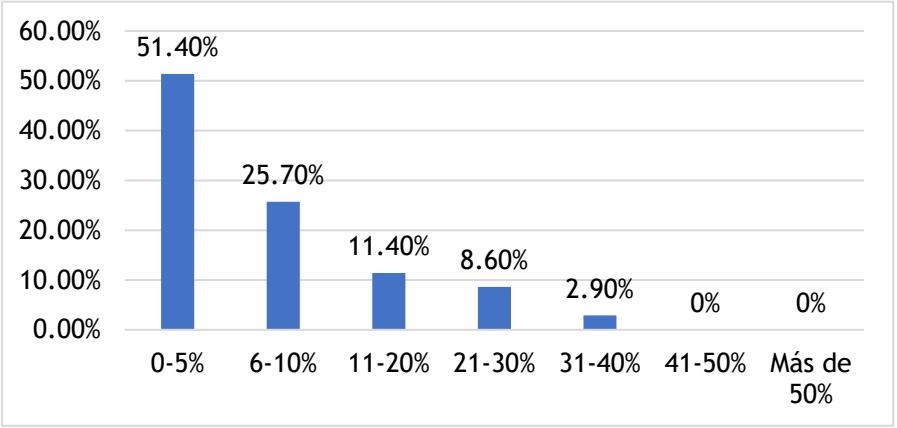
Docentes: de los 38 informes de Turnitin analizados de docentes, 35 correspondían a trabajos verificados como 100% humanos. De estos 35 trabajos, 4 (11.4%) tuvieron un porcentaje de IA superior al 20% en Turnitin.

Distribución de porcentajes de IA en trabajos 100% humanos de docentes:

- 0-5%: 18 trabajos (51.4%)
- 6-10%: 9 trabajos (25.7%)
- 11-20%: 4 trabajos (11.4%)
- 21-30%: 3 trabajos (8.6%)
- 31-40%: 1 trabajo (2.9%)
- 41-50%: 0 trabajos (0%)
- Más de 50%: 0 trabajos (0%)

Figura 1

Distribución de porcentajes de IA en trabajos 100% humanos de docentes



Fuente: Elaboración propia.

El porcentaje promedio de IA en trabajos 100% humanos de docentes fue de 8.7% (desviación estándar: 9.4%). El porcentaje máximo fue de 36%, correspondiente a un artículo enviado a una revista indexada sobre derecho constitucional.

Características de los textos que generan falsos positivos

El análisis de los trabajos con falsos positivos reveló varios patrones comunes:

- Longitud del documento

Los trabajos más cortos tuvieron mayor probabilidad de falsos positivos. El promedio de palabras en trabajos con falsos positivos fue de 1,420 palabras, mientras que en trabajos sin falsos positivos fue de 2,680.

Relación entre longitud y porcentaje de IA:

- Menos de 1,000 palabras: promedio 24.3% de IA
- 1,000-2,000 palabras: promedio 16.8% de IA
- 2,000-3,000 palabras: promedio 11.2% de IA
- Más de 3,000 palabras: promedio 7.4% de IA

Esto coincide con lo reportado por Turnitin, que advierte que los documentos cortos tienen mayor incidencia de falsos positivos.

- Tipo de trabajo

Los ensayos y exámenes parciales tuvieron mayor probabilidad de falsos positivos que los proyectos o tesis.

Porcentaje de trabajos con falsos positivos por tipo:

- Ensayos: 22.4% (11 de 49 ensayos)
- Exámenes parciales: 19.2% (5 de 26 exámenes)

- Proyectos parciales: 12.5% (3 de 24 proyectos)
- Tesis/trabajos finales: 8.3% (1 de 12 tesis)
- Artículos/docentes: 11.4% (4 de 35 artículos)

Esto sugiere que los trabajos más cortos y menos estructurados (ensayos, exámenes) son más propensos a falsos positivos que los trabajos largos y bien desarrollados (tesis, proyectos).

- Características lingüísticas

Los trabajos con falsos positivos tendían a tener ciertas características lingüísticas comunes:

Uso de frases formales estandarizadas:

- "En conclusión..." (presente en 83% de trabajos con falsos positivos vs. 45% en trabajos sin falsos positivos)
- "De acuerdo con..." (76% vs. 52%)
- "Según la teoría de..." (71% vs. 38%)
- "Por lo tanto..." (68% vs. 41%)
- "Es importante destacar que..." (64% vs. 35%)
- Citas normativas estandarizadas:

Los trabajos que citaban artículos constitucionales, códigos o leyes textualmente tuvieron mayor probabilidad de falsos positivos. Por ejemplo, citar el artículo 1° constitucional ("En los Estados Unidos Mexicanos todas las personas gozarán de los derechos humanos reconocidos en esta Constitución...") genera coincidencias espurias porque ese texto está en millones de documentos.

Estructura rígida:

Los trabajos con estructura muy predecible (introducción, desarrollo, conclusión, cada uno con frases estandarizadas) tuvieron mayor probabilidad de falsos positivos. Esto es particularmente común en estudiantes de primer grado que apenas están aprendiendo a escribir en estilo académico.

Idioma y estilo

Como era de esperarse, todos los trabajos estaban en español. Sin embargo, se observó que los trabajos con mayor uso de tecnicismos jurídicos en latín (habeas corpus, amparo, res judicata, etc.) o con citas en inglés tuvieron porcentajes de IA ligeramente más altos. También se observó que los trabajos de estudiantes cuyo primer idioma no era el español (por ejemplo, estudiantes de comunidades indígenas) tuvieron porcentajes de IA más altos en promedio, aunque esto no fue estadísticamente significativo debido al pequeño tamaño de la muestra.

Sin embargo, en la investigación de Sajid et al. (2025), no se encuentran diferencias importantes entre las personas que utilizan la IA y hablan inglés y los que no; excepto en la redacción de correos. Un aspecto importante en el ámbito de la detección de plagio recalca la utilización de la paráfrasis y la reescritura,

aprovechando los desarrollos en el análisis semántico, así como la combinación de métodos para características textuales y no textuales.

La revisión de los reportes mostró que los textos jurídicos con abundantes citas normativas y estructuras repetitivas eran los más propensos a ser clasificados como sospechosos. En varios casos, Turnitin marcó como "IA generada" párrafos que consistían en transcripciones de artículos constitucionales o leyes, lo cual evidencia una limitación técnica del detector.

Reacciones de los docentes ante el uso de la IA por estudiantes:

Cuando un trabajo tenía un porcentaje de IA superior al 20%, las reacciones de los docentes variaron:

Distribución de reacciones (n=16 trabajos con >20% IA):

- Cuestionamiento verbal al estudiante: 10 casos (62.5%)
- Solicitud de reelaboración del trabajo: 6 casos (37.5%)
- Reducción de calificación: 4 casos (25.0%)
- Rechazo del trabajo: 2 casos (12.5%)
- Sin acción (aceptación sin cuestionamiento): 3 casos (18.8%)
- Revisión de evidencias de autoría: 5 casos (31.3%)

Es notable que en 3 de los 16 casos (18.8%), el docente no tomó ninguna acción a pesar del porcentaje alto de IA. En entrevistas, estos docentes comentaron que "conocen el estilo de escritura del estudiante" o "el contenido es muy bueno para ser de IA", lo que sugiere que usan su juicio profesional más allá del porcentaje de Turnitin.

Uso de Turnitin por parte de los docentes

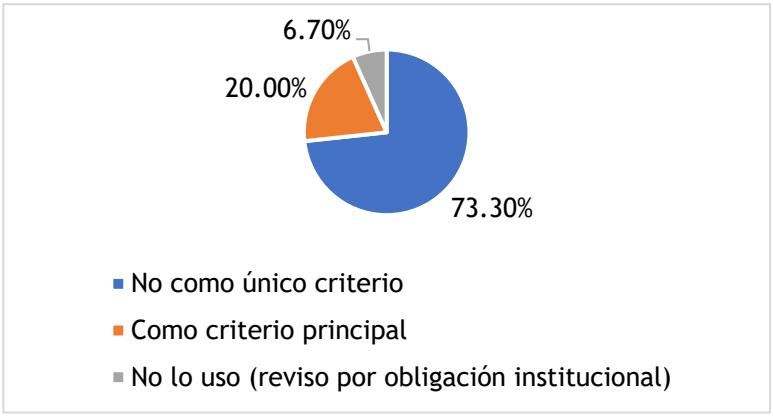
En las entrevistas, los 15 docentes describieron cómo usan Turnitin en sus evaluaciones:

Uso del porcentaje de similitud:

- Lo uso como referencia, pero no como único criterio: 11 docentes (73.3%)
- Lo uso como criterio principal de evaluación: 3 docentes (20.0%)
- No lo uso, solo lo reviso por obligación institucional: 1 docente (6.7%)

Figura 2

Uso de Turnitin (porcentaje de similitud) como criterio de evaluación por los profesores



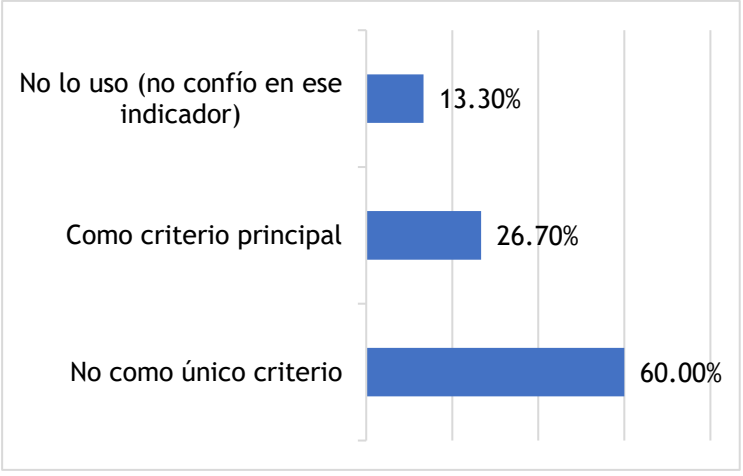
Fuente: Elaboración propia.

Uso del porcentaje de IA:

- Lo uso como referencia, pero no como único criterio: 9 docentes (60.0%)
- Lo uso como criterio principal de evaluación: 4 docentes (26.7%)
- No lo uso, no confío en ese indicador: 2 docentes (13.3%)

Figura 3

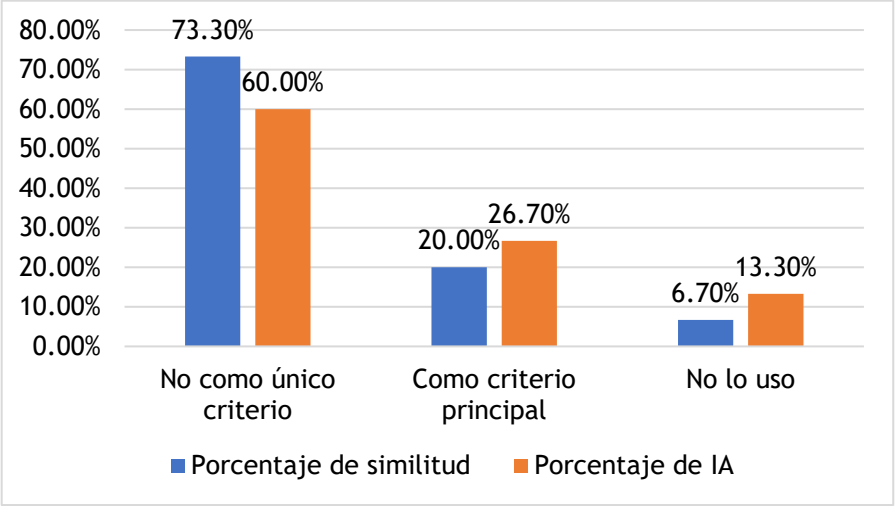
Uso de Turnitin (porcentaje de IA) como criterio de evaluación por los profesores



Fuente: Elaboración propia.

Figura 4

Comparación entre el uso de Turnitin como criterio de evaluación para los porcientos de similitud y el uso de IA



Fuente: Elaboración propia.

Umbrales considerados "aceptables":

- Menos del 20% de IA: 12 docentes (80.0%)
- Menos del 30% de IA: 2 docentes (13.3%)
- Menos del 10% de IA: 1 docente (6.7%)

Estos datos muestran que la mayoría de los docentes (73.3% para similitud, 60.0% para IA) usan Turnitin como referencia, pero no como único criterio. Sin embargo, 4 docentes (26.7%) sí usan el porcentaje de IA como criterio principal, lo que es problemático dada la evidencia de falsos positivos. Los profesores señalaron que los informes de Turnitin se utilizan en ocasiones como criterio único para rechazar trabajos, lo que genera conflictos éticos y académicos. Algunos docentes reconocieron que carecen de formación específica para interpretar los resultados, lo que aumenta la probabilidad de decisiones erróneas.

Políticas institucionales

La revisión documental reveló que:

- No existe un reglamento específico de integridad académica en la UAGro que mencione explícitamente el uso de IA.
- No se identificó en los documentos consultados una política institucional sobre umbrales aceptables de similitud o IA en Turnitin.
- No existe un proceso formal de apelación para estudiantes cuestionados por porcentajes de Turnitin.
- Algunas dependencias tienen lineamientos internos no oficiales (por ejemplo, "menos del 20% de similitud es aceptable"), pero estos no están documentados ni son uniformes.

Un docente comentó: "Cada profesor usa Turnitin como quiere. No hay lineamientos claros. Algunos son muy estrictos, otros son más flexibles. Esto genera mucha inconsistencia y, a veces, injusticias." Desde el punto de vista de la dirección de la universidad, el aspecto principal de este problema radica en las formas concretas de aplicación. El uso de estos detectores puede conducir a conclusiones injustas sobre el plagio, lo que influirá negativamente en los resultados académicos de los estudiantes y en la reputación de los autores docentes.

En la actualidad, es innegable que la Inteligencia Artificial forma parte integral de la sociedad. Gracias a los recientes avances en la IA generativa, ha ingresado en el campo de la Educación a una velocidad impresionante, aunque es fundamental considerar los beneficios que la IA podría ofrecer en este ámbito en las Instituciones de Educación Superior, así como los peligros que puede acarrear (Ortiz et al., 2025). Se coincide con Merlo y Rinaldi (2025) que realizan una investigación que tuvo como objetivo analizar el impacto del uso de Turnitin en trabajos de una maestría. Hicieron hincapié, tanto en el área de porciento de similitud como en el del uso de la IA. El presente estudio coincide con este en la reducción de su uso, así como la disminución de la utilización de textos no originales.

Los descubrimientos son coherentes con estudios recientes en diferentes entornos. Por ejemplo, para Alonso (2024) la utilización de inteligencia artificial para investigar bases de datos de estudios y después parafrasear levemente sin reconocer a los autores originales es algo habitual. Esto sugiere que las diferencias entre el entendimiento del autor respecto al tema y la calidad del material expuesto revelan una posible dependencia de la inteligencia artificial para apropiarse de forma inapropiada de las ideas ajenas.

Además, estudios realizados en Estados Unidos señalan que depender únicamente de estos sistemas puede llevar a penalizaciones injustas y a una disminución de la confianza en las instituciones (Deep, 2026). Para Erol et al. (2025) la confianza de los docentes en las herramientas tecnológicas crece para el análisis de textos evaluativos de los estudiantes. Pero al mismo tiempo la evidencia sobre herramientas que detectan la IA aún no alcanza la fiabilidad total y aparecen los falsos positivos que tanto predisponen a estudiantes y profesores; hecho que demuestra la imposibilidad de depender exclusivamente de estos instrumentos automatizados.

Se concuerda con Ruiz-Tirado et al. (2024) cuando plantean que la comprobación de la veracidad de los datos, sumada al fomento de una ética robusta en la aplicación de las tecnologías, se transforma en una labor esencial para asegurar que el saber obtenido sea genuino y significativo. Los hallazgos de esta investigación validan que la detección de inteligencia artificial con Turnitin presenta falsos positivos, constituyendo un inconveniente real y considerable en la Facultad de Derecho de la Universidad Autónoma de Guerrero. De un total de 67 trabajos completamente elaborados por estudiantes, 12 (17. 9%) mostraron tasas de IA que superaron el 20%, lo que es definido como un falso positivo según los parámetros de esta investigación. En lo que respecta a los profesores, 4 de 35 trabajos (11. 4%) arrojaron falsos positivos.

Es fundamental señalar que los datos presentados en el presente estudio son más elevados que el 1% de falsos positivos que menciona Turnitin (Turnitin, 2023a).

Esta diferencia puede deberse a varias causas. En primer lugar, Turnitin indica tasas de falsos positivos a nivel de documento cuando el porcentaje de inteligencia artificial supera el 20%, mientras que la investigación actual abarcó todos los trabajos que son completamente humanos, sin importar el porcentaje indicado. En segundo lugar, Turnitin fue desarrollado mayormente con información en inglés, y diversas investigaciones han mostrado que las tasas de error son entre 15 y 20% más altas en español. Por último, el ámbito específico del derecho, con su terminología técnica y referencias normativas estandarizadas, podría incrementar la posibilidad de encontrar falsos positivos.

Los hallazgos de esta investigación indican que los falsos positivos en la identificación de inteligencia artificial mediante Turnitin no son casos aislados, sino un fenómeno frecuente que impacta la confianza en los procesos de evaluación académica. Este descubrimiento coincide con estudios a nivel internacional que han registrado limitaciones parecidas.

Por ejemplo, Fernández y Ruiz (2023) en España identificaron que los documentos con terminología técnica específica mostraban proporciones de falsos positivos que superaban el 15%, lo cual se alinea con los resultados encontrados en la Licenciatura en Derecho de la Universidad Autónoma de Guerrero. De forma análoga, Soto (2024) indicó en Chile que los alumnos de derecho y ciencias sociales mostraban una mayor inclinación a ser identificados por los sistemas de IA, esto se debía a la estricta normativa y a la repetición de enfoques legales.

En el ámbito anglosajón, (Deep, 2026) señala que depender únicamente de estas herramientas podría debilitar la confianza en las instituciones y provocar castigos injustos, lo que subraya la urgencia de desarrollar políticas en las universidades que fusione la tecnología con la supervisión humana. Estudios recientes en el campo de la ética educativa indican que una dependencia excesiva en herramientas automatizadas puede crear un ambiente de desconfianza y vigilancia desmedida, lo que va en contra de los principios de formación crítica y autonomía de los estudiantes (Pérez, 2021). En este contexto, la administración universitaria debe tomar una postura proactiva en la regulación del uso de estas tecnologías, implementando protocolos específicos que impidan tomar decisiones al azar.

Se coincide con Montesinos y Guaña-Moya (2026) que en reciente estudio señalan que en situaciones en las que se utilizan documentos que combinan esfuerzo humano y automatización, la creación de contenido junto con la intervención humana durante el proceso de elaboración de estos textos, disminuye la proporción de detectabilidad respecto a su procedencia, en comparación con entornos de inteligencia artificial exclusivamente. Los hallazgos de esta investigación coinciden con estudios recientes que han sido divulgados en publicaciones indexadas en Scopus, los cuales alertan sobre los peligros de depender únicamente de detectores de inteligencia artificial.

Tsigaris y Teixeira (2026), en un artículo publicado en ScienceDirect, destacan que la mayor parte de los resultados obtenidos por los detectores de IA son erróneos y que estos sistemas tienden a castigar de manera injusta a alumnos que elaboran escritos con formas estructurales y repetitivas características de la

redacción académica.

No se coincide con los resultados hallados por Aquino et al. (2026) quien evaluó la confianza que se puede tener en la utilización de los algunos sistemas de detección de texto generado por IA y porcientos de similitud en instituciones educativas de Ecuador. Los resultados encontrados muestran altas tasas de falsos negativos y casi nula la posibilidad de disciplinar a los que los llevan a cabo. La investigación conduce a la desconfianza que generan situaciones como estas y la real posibilidad de regresar completamente a los escritos de papel y lápiz y sobre todo a defensas orales.

Angelier (2026), en AI and Ethics (Springer), enfatiza que los sistemas de detección de inteligencia artificial presentan sesgos y producen errores de identificación que perjudican la justicia en la evaluación, sugiriendo la necesidad de un marco que integre la tecnología y la revisión por parte de humanos. Este estudio ofrece evidencia procedente de México y América Latina, donde el uso del lenguaje legal, la estandarización de las citas normativas, las prácticas jurídicas y la rigidez de la estructura académica aumentan la probabilidad de cometer errores. Esto indica que los falsos positivos no constituyen únicamente un desafío técnico, sino que también son un asunto cultural y de disciplina, lo que implica la necesidad de implementar estrategias de gestión universitaria que se adapten al contexto.

La discusión indica que la cuestión no se limita solo a la tecnología, sino que también se relaciona con la administración universitaria: de qué manera las entidades comprenden y utilizan los resultados que ofrecen estas herramientas. Yanangómez et al. (2026) llevaron a cabo un estudio reciente cuyo objetivo fue evaluar la fiabilidad de varios detectores de Inteligencia Artificial aplicados a trabajos académicos en entornos educativos. Ellos concluyen que ciertos detectores de IA muestran falsos positivos y tienen limitaciones significativas en su exactitud, por lo que no se deberían usar como evidencia definitiva en los procesos de evaluación académica. Estos autores sugieren que estas herramientas deben ser complementadas con una revisión humana, análisis del contexto y criterios pedagógicos antes de formarse opiniones sobre la veracidad de un trabajo realizado por un estudiante.

El estudio actual se alinea con las recomendaciones de Chechitelli (Turnitin, 2023b) para tratar con falsos positivos:

- Estar informado antes de tomar decisiones: es crucial contemplar de antemano la probabilidad de un falso positivo y crear un enfoque sobre cómo proceder al evaluar el resultado.
- Adopta una mentalidad optimista: en este ámbito repleto de novedades y aspectos no familiarizados, busca otorgar a tus alumnos la presunción de buena fe.
- Sé sincero y transparente: es esencial admitir de antemano que pueden ocurrir falsos positivos, por lo que tanto educadores como alumnos necesitan estar listos para mantener una comunicación clara y honesta.

Conclusiones

Existen características comunes en los textos que generan falsos positivos y son detectados por Turnitin en la Licenciatura en Derecho de la Universidad Autónoma de Guerrero. Algunas de ellas son la longitud del documento (menos de 1,000 palabras), los ensayos como tipo de trabajo, el uso de frases formales y citas normativas estandarizadas. Los trabajos con estructura muy predecible (introducción, desarrollo, conclusión) y los textos con mayor uso de tecnicismos jurídicos en latín o con citas en inglés. En este sentido, los datos reunidos en este estudio permiten plantear que los detectores de IA no deben usarse como única evidencia de mala conducta académica por lo que se propone un marco de evaluación integral que combine revisión humana, procesos de apelación y evaluación alternativa.

Referencias

Alnaimat, F., Feras AlSamhori, A. R., Hamdan, O., Seiil, B., & Qumar, A. B. (2025). Perspectives of Artificial Intelligence Use for In-House Ethics Checks of Journal Submissions. *Journal of Korean Medical Science*, 40(21), 1–11. https://doi.org/10.3346/JKMS.2025.40.E170

Alonso, R. A. (2024). Hacia un marco ético de la inteligencia artificial en la educación. Teoría de la Educación. *Revista Interuniversitaria, 36*(2), 79-98. https://revistas.usal.es/tres/index.php/1130-3743/article/view/31821

Angelier, V. (2026). *The pitfalls of AI detection in academic writing: Bias, false positives, and the need for inclusive assessment*. Aland Ethics. Springer. https://doi.org/10.1007/s43681-026-01376-w

Aquino, R. M., Estrella, N. P., Iza, G. P., & Pérez, M. I. (2026). Evaluación de sistemas de detección de IA y transformación del aprendizaje en la educación secundaria de Cotopaxi. (2026). *Revista Científica Y Tecnológica VICTEC*, 7(13), 237-247. http://server.istvicenteleon.edu.ec/victec/index.php/revista/article/view/251

Arteaga, J. E. (2023). Perspectivas de los docentes sobre la IA generativa: implicaciones para el aprendizaje personalizado en América Latina. *Ethos Scientific Journal, 1*(1), 56–69. https://doi.org/10.63380/ESJ.V1N1.2023.22

Cándano Pérez, M., Del Solar-Duarte, N., Remeseiro Reguero, R., Bracho-Fuenmayor, P. L., Baeza Ugarte, C. G., & Guzmán Sánchez, M. (2025). Aprendizaje basado en el uso de herramientas de IA: experiencia de realidad aumentada en la asignatura de “Nociones de Economía” con estudiantes de derecho de la Universidad Tecnológica Metropolitana del Estado de Chile. *Revista De Educación Y Derecho*, (2-Extraordinario), 331–353. https://doi.org/10.1344/REYD2024.2-Extraordinario.49196

Chávez-Martínez, O. (2026). ¿Cómo detectar textos científicos generados con

inteligencia artificial? *Revista Médica del Instituto Mexicano del Seguro Social*, *64*(1), e6952. https://doi.org/10.5281/zenodo.17477984

Deep, P. D. (2026). Artificial intelligence and academic integrity in higher education: Student perceptions, institutional responses, and the limits of AI detection. *International Journal of Studies in Education and Science (IJSES)*, *7*(3), 397-420. https://doi.org/10.46328/ijses.7882

Erol, G., Ergen, A., Gülşen Erol, B., Kaya Ergen, Ş., Bora, T. S., Çölgeçen, A. D., Araz, B., Şahin, C., Bostancı, G., Kılıç, İ., Macit, Z. B., Sevgi, U. T., & Güngör, A. (2025). Can we trust academic AI detective? Accuracy and limitations of AI-output detectors. *Acta Neurochirurgica*, *167*(214), 1–12. https://doi.org/10.1007/S00701-025-066224/FIGURES/

Fleckenstein, J., Meyer, J., Jansen, T., Keller, S. D., Köller, O., & Möller, J. (2024). Do teachers spot AI? Evaluating the detectability of AI-generated texts among student essays. *Computers and Education: Artificial Intelligence*, 6(100209). https://doi.org/10.1016/J.CAEAI.2024.100209

González Torres, V. H., Bracho-Fuenmayor, P. L., Baldevenites, E. V. L., Carrillo Guerrero, M. V., & Santander Erazo, R. D. (2024). Immersive learning in the metaverse: a review of evidence on pedagogical effectiveness and implementation gaps in Higher Education. *Metaverse Basic and Applied Research*, 3. 97. https://doi.org/10.56294/mr2024.97

Hua, H., & Yao, C. J. (2024). Investigating generative AI models and detection techniques: impacts of tokenization and dataset size on identification of AI-generated text. *Frontiers in Artificial Intelligence*, *7*(1469197). https://doi.org/10.3389/FRAI.2024.1469197/BIBTEX

Hyatt, J. P., Bienenstock, E. J., Firetto, C. M., Woods, E. R., & Comus, R. C. (2025). Using aggregated AI detector outcomes to eliminate false positives in STEMstudent writing. *Advances in Physiology Education*, 49(2), 486–495. https://doi.org/10.1152/ADVAN.00235.2024

Khan, H. U., Naz, A., Alarfaj, F. K., & Almusallam, N. (2025). Identifying artificial intelligence-generated content using the DistilBERT transformer and NLP techniques. *Scientific Reports*, 15(20366), 1–26. https://doi.org/10.1038/S41598-025-082087;SUBJMETA

Merlo, C. E., & Rinaldi, S. M. (2025). El uso de Turnitin para la evaluación de la originalidad académica: un balance del primer año de implementación en una universidad argentina. *Revista Científica RES NON VERBA*, 15(1), 49–68. https://doi.org/10.21855/resnonverba.v15i1.999

Montesinos, L., & Guaña-Moya, J. (2026). Detección de plagio y autoría en la era de los LLM: precisión y efectos pedagógicos. *Ethos Scientific Journal*, 4(1), 284-309. https://doi.org/10.63380/esj.v4n1.2026.352

Navarrete-Cazales, Z., & Cruz-Victoria, J. C. (2026). Autoría e integridad académica ante la inteligencia artificial generativa en educación superior. *Revista Educación Superior y Sociedad (ESS)*, 37(2), 260-283. https://doi.org/10.54674/ess.v37i2.1174

Ortiz, J., Quintana Álvarez, J. H., Granda Dihigo, A., & Gómez León, Y. (2025). La Inteligencia Artificial en Educación Superior: análisis de riesgos y oportunidades. *Revista Cubana De Educación Superior*, *44*(especial 2), 74-84. https://revistas.uh.cu/rces/article/view/12335

Pellegrina, D., & Helmy, M. (2025). AI for scientific integrity: detecting ethical breaches, errors, and misconduct in manuscripts. *Frontiers in Artificial Intelligence*, 8(1644098). https://doi.org/10.3389/FRAI.2025.1644098/BIBTEX

Pratama, A. R. (2025). The accuracy-bias trade-offs in AI text detection tools and their impact on fairness in scholarly publication. *PeerJ Computer Science*, 11(e2953). https://doi.org/10.7717/PEERJ-CS.2953/SUPP-3

Ruiz-Tirado, S. G., Morán-Ortega, S. A., Simental-López, L. M., & Tirado-López, A. B. (2024). Barreras de la Inteligencia Artificial generativa en estudiantes de educación superior. Percepción docente. *RITI Journal*, *12*(25). 26-37 https://doi.org/10.36825/RITI.12.25.003

Sajid, M., Sanaullah, M., Fuzail, M., Malik, T. S., & Shuhidan, S. M. (2025). Comparative analysis of text-based plagiarism detection techniques. *Plos One*, 20(e0319551). https://doi.org/10. 1371/JOURNAL.PONE.0319551

Schneider, D. M., Mishra, A., Gluski, J., Shah, H., Ward, M., Brown, E. D., Sciubba, D. M., & Lo, S.-F. L. (2025). Prevalence of Artificial Intelligence-Generated Text in Neurosurgical Publications: Implications for Academic Integrity and Ethical Authorship. *Cureus*, 17(2). https://doi.org/10.7759/CUREUS.79086

Tiglla, B. (2025). Impacto de la Inteligencia Artificial Generativa (IAG) en la personalización del aprendizaje en universidades latinoamericanas. *Alpha International Journal*, 3(1), 18–30. https://doi.org/10.63380/AIJ.V3N1.2025.55

Trejo, D. (2026, julio 29). *Falsos positivos de IA en examen UNAM abren vacío jurídico*. Global Media. https://www.globalmedia.mx/articles/falsos_positivos_de_ia_en_examen_unam_abren_vacio_juridico

Tsigaris, P., & Teixeira, J. A. (2026). AI detecting AI in academic writing: Why most AI detector findings are false. *Computers in Human Behavior Reports*, *14*, 100354. https://doi.org/10.1016/j.chbr.2026.100354

Turnitin. (2023a). *Actualización sobre la detección de escritura con IA por la Directora de Producto Turnitin*. https://latam.turnitin.com/blog/actualizacion-sobre-deteccion-de-escritura-con-ia-por-la-directora-de-producto-turnitin

Turnitin. (2023b) *¿Qué significan los falsos positivos en nuestra función de detección de escritura con IA?* https://www.turnitin.com.mx/blog/que-significan-los-falsos-positivos-en-funcion-de-deteccion-de-escritura-con-ia

Turnitin. (2023c). *Detección de escritura de IA en la vista de reporte clásica*. https://guides.turnitin.com/hc/es/articles/28457596598925-

Detecci%C3%B3n-de-escritura-de-IA-en-la-vista-de-reporte-cl%C3%A1sica-espa%C3%B1ol

Turnitin. (2024). *AI writing detection: Guidance for educators*. Turnitin Official Documentation. https://www.turnitin.com

Weber, D., Anohina, A., Bjelobaba, S., Foltýnek, T., Guerrero Dib, J., Popoola, O., Šigut, P., & Waddington, L. (2023). Testing of detection tools for Algenerated text. *International Journal for Educational Integrity, 19*(26), 1–39. https://doi.org/10.1007/s40979-023-00146-z

Yanangómez, P. A., Díaz, E. R., Ojeda. L. M., Yaguachi, C. B., & Carrillo, M. C. (2026) Falsos positivos en detectores de Inteligencia Artificial en trabajos académicos: análisis comparativo de plataformas digitales en contextos educativos. *ASCE MAGAZINE 5*(2) 1976-1999 https://doi.org/10.70577/asce.v5i2.861

Sobre el autor principal

Perla Elisabeth Ventura Ramos: se formó en la Universidad Autónoma de Guerrero, México, como ingeniera en Sistemas, Máster en Ciencias Computaciones y Dra. en Pedagogía. Actualmente profesora-investigadora en Tecnología de la Educación en la Universidad. Se ha desempeñado como académica a tiempo completo en la Facultad de Derecho de la Universidad Autónoma de Guerrero, en la que imparte las unidades de Aprendizajes Manejo de las Tecnología de Información y Comunicación, Derecho de la Informática, Derecho, Género, Poder, Estadísticas Aplicadas al Derecho.

Declaración de responsabilidad autoral

Perla Elisabeth Ventura Ramos 1: Conceptualización, Curación de datos, Análisis formal, Investigación, Metodología, Recursos, Software, Supervisión, Validación/Verificación, Visualización, Redacción/borrador original y Redacción.
Norma Yadira Memije Alarcón 2: Curación de datos, Análisis formal, Investigación, Recursos, Supervisión, Validación/ Verificación, Visualización, Redacción/ borrador original, y Redacción, revisión y edición.
Jesús Zaragoza Martínez 3: Curación de datos, Análisis formal, Investigación, Recursos, Validación/ Verificación, Visualización, Redacción/ borrador original.

Agradecimientos:

Financiación:

Recursos propios.