*Recibido 23 Ene. 2024* *ReCIBE, Año 13 No. 2, Sep. 2024*
*Aceptado 15 Sep. 2024*

# Inteligencia Artificial en la Academia: Oportunidades y Desafíos

# Artificial Intelligence in Academia: Opportunities and Challenges

**Fernando Wario Vazquez[1,2]**
**Ricardo Ramírez Romero[3]**

[1] Centro Universitario de Ciencias Exactas e Ingenierías, Universidad de Guadalajara
[2] Correspondencia : fernando.wario@academicos.udg.mx
[3] Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara

# RESUMEN

Aunque en los últimos años el uso de la inteligencia artificial (IA) se ha vuelto más común y accesible, especialmente con herramientas como ChatGPT, sus raíces se remontan a más de 50 años. Desde la década de 1950, cuando Turing propuso la primera prueba para evaluar la capacidad de las máquinas de emular el comportamiento humano, la IA ha tenido un desarrollo con altibajos. No obstante, en la última década, ha experimentado un crecimiento significativo, manifestándose en múltiples aspectos de la vida cotidiana, incluida la academia. El uso de la IA en la academia se ha popularizado recientemente, particularmente entre estudiantes de nivel medio y superior, con herramientas como ChatGPT. Sin embargo, hay otras aplicaciones destacadas, como Inciteful, Litmaps, Jenni, Wisio y Elicit, que facilitan la recopilación y el análisis de grandes volúmenes de información. Además, herramientas como Grammarly, Quillbot y Jarvis son auxiliares valiosos en la redacción y revisión de textos académicos. Dado este contexto, es crucial analizar tanto las oportunidades que la IA ofrece en el ámbito académico como las áreas que requieren una atención especial para asegurar un uso ético y responsable de estas herramientas. Este artículo examina los orígenes de la IA, su recorrido histórico, y presenta diversas herramientas aplicables en la academia. Se exploran los aspectos que demandan mayor cuidado para un uso ético y responsable y se discuten las posibles vías futuras en el desarrollo de estas herramientas tecnológicas.



**Palabras clave**: Inteligencia Artificial, Academia, ChatGPT, aplicaciones, regulación



# ABSTRACT

Although artificial intelligence (AI) has become increasingly common and accessible in recent years, particularly with tools like ChatGPT, its origins date back over 50 years. Since the 1950s, when Alan Turing proposed the first test to evaluate a machine's ability to mimic human behavior, AI has experienced a series of advancements and setbacks. However, AI has grown substantially in the last decade, integrating into various aspects of daily life, including academia. The adoption of AI in academia has recently surged, particularly among high school and university students, through tools like ChatGPT. In addition to ChatGPT, other AI applications such as Inciteful, Litmaps, Jenni, Wisio, and Elicit have facilitated the efficient collection and analysis of large datasets. Tools like Grammarly, Quillbot, and Jarvis have also become valuable aids for drafting academic texts. Given this context, it is essential to examine both the opportunities that AI presents in the academic field and the critical areas that require attention to ensure these technologies' ethical and responsible use. This article explores AI's origins and historical evolution, highlights various AI tools applicable in academic settings, and examines the aspects requiring careful consideration for ethical and responsible use. It concludes by discussing potential future directions for developing these technological tools.



**Keywords**: Artificial Intelligence, Academia, ChatGPT, applications, regulation

# 1. INTRODUCCIÓN

Las nuevas herramientas basadas en tecnologías de inteligencia artificial (IA) parecieran cada vez acercarnos más a un futuro como los descritos en novelas de ciencia ficción. Para algunos usuarios, esto más que alentador es inquietante, pues más allá de problemáticas ya vividas con otros avances tecnológicos como la pérdida de puestos de trabajo y la concentración de poder por parte de las potencias mundiales, se encuentran nuevas preocupaciones, como la falta de comprensión global por parte del usuario. Particularmente, sobre el funcionamiento de muchos de los algoritmos, la privacidad de los datos con los que son entrenados los modelos detrás de las herramientas de IA, los sesgos mostrados por algunos modelos, reflejo de los datos con los cuales han sido entrenados, e incluso la potencial capacidad de la IA de autogestionarse y salirse de nuestro control. Estas preocupaciones han llevado a un amplio debate sobre las medidas que deberían tomarse para regular el uso y desarrollo de nuevas tecnologías de IA. Sin desestimar estas preocupaciones, es una realidad que las tecnologías de IA han impactado ya en menor o mayor medida a diversos sectores económicos y sociales, siendo la academia particularmente alcanzada con el reciente boom de la IA generativa.

En el presente artículo exploramos el camino que ha llevado a la IA a convertirse en el que probablemente sea el campo de estudio más relevante en la actualidad, y cómo diversas tecnologías a las que ha dado lugar están siendo utilizadas en la academia, evaluando sus potenciales beneficios, pero también los desafíos que plantean su uso responsable. Destacamos como con el paso del tiempo se ha vuelto crucial garantizar un desarrollo ético y supervisado de estas nuevas tecnologías para evitar dependencias perniciosas y permitir que la IA sea una herramienta segura y efectiva tanto en la academia como en otros sectores.

# 2. ORÍGENES E HISTORIA DE LA IA

Aunque el auge de las herramientas de Inteligencia Artificial (IA) es un fenómeno reciente, y cada vez es más frecuente encontrarlas en diversos aspectos de la vida cotidiana, la IA como campo de estudio tiene más de siete décadas de historia. En 1950, cuando las computadoras eran aún utilizadas principalmente como simples calculadoras a gran escala, Alan Turing abordó en su artículo “Computing Machinery and Intelligence” publicado en la revista Mind el problema de la IA (Turing, 1950). En este artículo, Turing buscaba establecer un estándar para indicar si una computadora puede ser considerada “inteligente”; para este fin propuso un juego llamado “El Juego de la Imitación” el cual posteriormente sería conocido como el “Test de Turing”. En este juego, un evaluador humano juzga las conversaciones en lenguaje natural entre un humano y una máquina, si el evaluador no es capaz de distinguir de manera confiable la máquina del humano, la máquina ha pasado la prueba. De esta manera Turing evita la pregunta de si las máquinas son capaces de pensar, la cual plantea el problema de definir la palabra pensar; en su lugar, la prueba propuesta por Turing respondería a la pregunta: ¿Pueden las máquinas actuar como entidades pensantes?

Después, durante el verano de 1956, en la universidad Dartmouth College, ubicada en Hanover, New Hampshire (Estados Unidos), tuvo lugar la conferencia organizada por John McCarthy, Marvin Minsky, Nathaniel Rochester y Claude Shannon (Moor, 2006). Esta conferencia es popularmente considerada como el evento seminal de la IA como campo de estudio, haciendo uso por primera vez del término “Inteligencia Artificial” para referirse a lo que hasta entonces era el campo de las “maquinas pensantes” y recibía nombres como cibernética, teoría de los autómatas y procesamiento complejo de la información.

El interés por las ideas discutidas durante la conferencia de Dartmouth, llevó a la IA a gozar de un primer periodo de apogeo durante las décadas de 1960 y 1970. Durante este periodo se dieron algunos desarrollos prometedores como el de ELIZA (Weizenbaum, 1966), una versión temprana de procesador de lenguaje natural (referido al habla y escritura humana y abreviado como NLP por sus siglas en inglés) (Nadkarni et al., 2011) desarrollado por Joseph Weizenbaum del MIT y el cual, es considerado como uno de los primeros chatbots de la historia. A finales de los años 70s surgieron los primeros esfuerzos por establecer un organismo que permitiera la comunicación continua entre las diferentes instituciones

trabajando en temas de IA, fue así como se llevó a cabo la conferencia conjunta en IA de 1977 y 1979. Posteriormente, se formaría la Asociación Americana de Inteligencia Artificial, la cual finalmente daría lugar a la actual Asociación para el Avance de la Inteligencia Artificial (AAAI por sus siglas en inglés).

Este temprano auge de la IA comenzaría a desacelerar a finales de los años 70s y continuaría a la baja hasta inicios de los 90s en lo que es conocido como “Invierno de la IA”. Aunque la historia de la IA cuenta con varios periodos de gran entusiasmo seguidos por la desilusión de algunos ante los resultados de sus desarrollos en escenarios de la vida real y de fuertes críticas a la confiabilidad de sus tecnologías, este periodo fue particularmente difícil para los grupos de investigación trabajando en IA, los cuales sufrieron un severo recorte en el financiamiento de sus proyectos. Dicho recorte se debió principalmente a la evaluación sobre el estado de la investigación en IA que realizó el reconocido matemático Sir James Lighthill a petición del Parlamento del Reino Unido. En dicho reporte, Lighthill criticaba severamente los algoritmos de la IA, los cuales consideraba solo adecuados para resolver versiones de “juguete” de los problemas del mundo real (McCarthy, 1974).

Probablemente el evento que motivó la reactivación de grandes fondos para la investigación y desarrollo en IA fue el encuentro entre la supercomputadora desarrollada por IBM® “Deep Blue” y el maestro ajedrecista Gary Kaspárov. El 10 de febrero de 1996, Deep Blue venció a Kasparov en la primera partida de seis pactadas. Aunque Kasparov terminó ganando el encuentro con un resultado de 4 - 2, la victoria de las partidas de Deep Blue sobre el campeón mundial de ajedrez vigente representó un hecho histórico en el campo de la IA.

Muchos años después del éxito de Deep Blue, IBM® desarrolló “Watson DeepQA”, un sistema computacional capaz de responder a preguntas formuladas en lenguaje natural. En febrero de 2011 Watson DeepQA participó en el popular concurso de televisión estadounidense “Jeopardy!”, resultando vencedor frente a dos destacados participantes del concurso. En una muestra más de los importantes avances logrados en temas de NLP (procesamiento de lenguaje natural), los gigantes de la industria Apple® y Amazon® presentaron sus asistentes virtuales en 2011 y 2014, respectivamente. A pesar de la sorprendente capacidad de estos asistentes virtuales para comprender y responder en lenguaje natural, sus limitaciones son evidentes, siendo incapaces de responder a preguntas o comandos diferentes a aquellos para los cuales han sido programados.

El siguiente gran avance en IA se dió como resultado de mejoras tanto en hardware como en los modelos de redes neuronales. En particular, los avances fueron evidentes en aplicaciones de visión por computadora, donde se encontró una enorme área de oportunidad al implementar los modelos de redes neuronales convolucionales (CNN por sus siglas en inglés) en unidades de procesamiento gráfico (GPU por sus siglas en inglés), reduciendo de manera drástica el tiempo requerido para entrenar las CNNs y permitiendo a su vez la implementación de modelos mucho más complejos (LeCun et al., 2015). Como resultado de esta convergencia de factores, la CNN nombrada AlexNet (Krizhevsky et al., 2017), resultó vencedora en el reto de clasificación de imágenes de la base de datos ImageNet. AlexNet, desarrollada por Alex Krizhevsky, Ilya Sutskever y Geoffrey Hinton de la Universidad de Toronto, obtuvo un porcentaje de error de tan solo el 15.3%, colocándose así a 10.8 puntos de distancia de su seguidor más cercano.

Dos décadas después de la victoria de Deep Blue sobre Kasparov, en marzo de 2016, el programa computacional AlphaGo (Silver et al., 2016) marcó otro hito en la historia de la IA al vencer al jugador profesional de Go nivel 9-dan, Lee Sedol, siendo esta la primera vez que un programa de Go vencía a un profesional del nivel más alto y sin piedras de ventaja (Figura 1). El juego de Go, a pesar de ser considerado fácil de aprender, resulta particularmente difícil para los programas computacionales por la gran cantidad de posiciones posibles en el tablero. Por este motivo, los desarrolladores de AlphaGo utilizaron un enfoque completamente distinto al de Deep Blue, el cual era un sistema experto (Jackson, 1998) operando a base de reglas y variables previamente ajustadas por expertos ajedrecistas. Por su parte, AlphaGo es un conjunto de redes neuronales y algoritmos de búsqueda entrenados a través de métodos de aprendizaje reforzado, incluyendo millones de partidas jugadas contra expertos y contra si mismo. La victoria de AlphaGo al nivel más alto en un juego previamente considerado fuera del alcance de los sistemas computacionales renovó el entusiasmo en las nuevas tecnologías de IA.

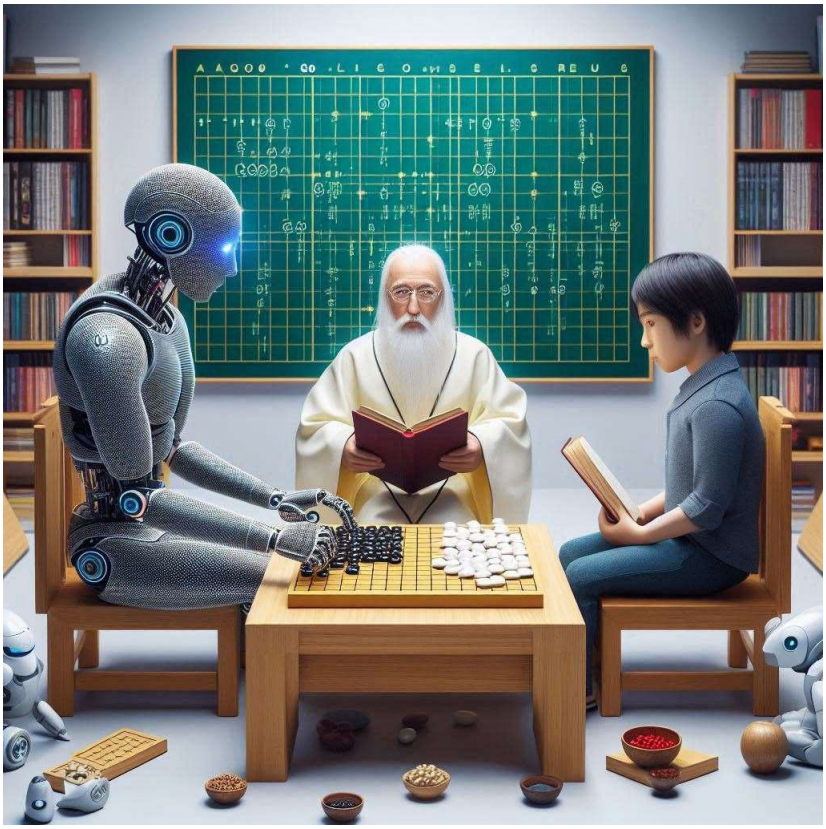

Figura 1. Creada con la herramienta Copilot (DALL-E 3) con IA usando el prompt: “Genera una imagen donde se represente la histórica partida entre la IA y el campeon mundial de Go”

Así como la arquitectura de la CNN AlexNet marcó un antes y un después en el procesamiento de imágenes con redes neuronales, la arquitectura transformer (Vaswani et al., 2017) desarrollada por un grupo de investigadores en Google® permitió avances acelerados en la interpretación y generación de nueva información con redes neuronales. El tiempo de entrenamiento requerido por las redes neuronales basadas en la arquitectura transformer es significativamente menor al de otros modelos previos, pues estas no cuentan con unidades recurrentes. En la actualidad se pueden encontrar variaciones de la arquitectura transformer en modelos de lenguaje grande (LLM por sus siglas en inglés) y en la generación de imágenes a partir de descripciones en lenguaje natural (Radford et al., 2019). La compañía OpenAI® desarrolló los transformadores generativos preentrenados (GPT por sus siglas en inglés), los cuales se han convertido en los modelos de referencia LLM y de la IA generativa. Después de presentar su primer modelo GPT en 2018 (Radford et al., 2018), OpenAI ha lanzado nuevas iteraciones de este modelo, con la más reciente versión – GPT-4o – siendo lanzada en mayo de 2024. GPT-4o es un LLM-multimodal y multilingüe, el cual es capaz de procesar y generar texto, imágenes y audio.

# 3. LA IA EN LA ACADEMIA

El boom de los modelos generativos ha renovado el interés y la confianza en herramientas basadas en IA; como consecuencia, en la actualidad existe una gran variedad de desarrollos, tanto basados en modelos generativos como en otras tecnologías de IA, los cuales han sido rápidamente adoptados en diversos sectores, incluyendo la academia. Mientras que algunas de estas herramientas cumplen propósitos generales como la producción de textos e imágenes, otras han sido desarrolladas para solucionar problemáticas específicas del quehacer académico.

El mundo se encuentra cada vez más interconectado y el sector académico no es una excepción, razón por la cual los avances científicos y tecnológicos suceden cada vez con mayor rapidez. Para ser partícipes activos de estos avances es necesario identificar áreas de oportunidad, ya sea en la forma de temas poco explorados que se verían enriquecidos por nuevas perspectivas, o como desarrollos multidisciplinarios e interdisciplinarios de temas en boga. De cualquier manera, el primer paso para identificar estas áreas de oportunidad es contar con una visión global de los desarrollos existentes relacionados al tema de interés. Con esta finalidad se han creado herramientas como *Inciteful*, *Litmaps*, *Connected Papers* y *Open Knowledge Maps*, las cuales proyectan las conexiones existentes entre diferentes estudios en forma de grafos y mapas de conocimiento interactivos. Este tipo de visualizaciones permite al usuario comprender el contexto en el cual han ocurrido los estudios e identificar potenciales líneas de investigación.

Una vez que se cuenta con un tema de investigación definido, es necesario llevar a cabo un extenso estudio bibliográfico. Esta tarea suele consumir una gran cantidad de tiempo, y conforme el cuerpo de conocimiento aumenta puede tornarse en una tarea abrumadora, incluso para los más experimentados. Algunas plataformas de IA como *Jenni*, *Wisio*, *Elicit* y *OpenRead* permiten automatizar ya sea parte o en su totalidad este proceso de búsqueda extensa y procesamiento de información bibliográfica. Estas herramientas incluyen agentes semánticos que a través de NLP generan resúmenes de artículos científicos, destacando los hallazgos más importantes, encontrando contenido relacionado e indicando tendencias claves. Utilizadas de manera responsable, verificando la información y fuentes cuando es necesario, estas herramientas permiten ahorrar tiempo en la realización de estudios bibliográficos, la conceptualización de estudios o la actualización de información.

En la era de la información, con enormes y complejos bancos de datos a disposición, un problema recurrente es el arreglo y acomodo de estos para su análisis y presentación. Tradicionalmente llevar a cabo estas tareas ha requerido de amplios conocimientos de estadística y programación, sin embargo, algunas herramientas de IA están cambiando este escenario. La transición se está dando de diferentes maneras, en algunos casos, herramientas para análisis de datos como *Microsoft® Fabric* y *Cognos Analytics de IBM®* han ido incorporando módulos de IA para automatizar sus procesos y permitir instrucciones en lenguaje natural. Por su parte, asistentes de IA como *GPT-4o* han habilitado en su plataforma el análisis de datos ya almacenados en la nube. Por último, existen herramientas de IA desarrolladas exclusivamente para el análisis de datos, como *Julius* cuya interfaz puede resultar más intuitiva para el usuario final. Sin importar sus diferencias, todas estas herramientas permiten a los usuarios llevar a cabo sus análisis en una fracción del tiempo usualmente requerido y lograr visualizaciones de alta calidad, ayudándoles a comunicar de manera más efectiva sus resultados.

La producción de textos académicos es una tarea sumamente importante, pues es el medio principal a través del cual la comunidad académica se mantiene informada de los avances en su área. Lograr textos de alta calidad que presenten la información de manera clara y convincente requiere tanto de experiencia como del dominio de los recursos de redacción adecuados. Esta tarea, ya de por si demandante, se vuelve aún más compleja si el texto debe ser escrito en un idioma diferente al nativo del autor. Con el fin de atenuar estas dificultades, herramientas basadas en IA como *Grammarly*, *QuillBot* y *Jarvis* han sido desarrolladas para fungir como asistentes de escritura. A diferencia de las herramientas de generaciones anteriores, que se limitaban a identificar y corregir errores ortográficos y gramaticales, las nuevas herramientas de IA sugieren modificaciones al texto que mejoran su coherencia y claridad (Figura 2). De similar manera, también pueden sugerir cambios en la estructura y selección de palabras para que permitan una comunicación más efectiva de las ideas principales.

Figura 2. Creada con la herramienta Copilot (DALL-E 3) con IA usando el prompt: “Crea una imagen donde se aprecien las diferentes herramientas con IA empleadas en la academia”

# 4. PROS Y CONTRAS DE LA IA EN LA ACADEMIA

¡Parece magia! ¡está increíble!, son algunas de las expresiones que surgen cuando a ChatGPT se le pide que haga un cuestionario con preguntas en diferentes formatos sobre algún tema en particular y lo genera con bastante acierto y de manera instantánea. Como hemos visto, la IA en la academia puede ofrecer diferentes beneficios como la generación instantánea de información (ChatGPT, *Inciteful*, *Litmaps*, *Jenni*, etc.), la interacción personalizada (que se refiere a que las aplicaciones responden a solicitudes específicas de cada usuario), el aprendizaje adaptativo y una creciente accesibilidad (Samala et al., 2024). Por ejemplo, ChatGPT puede proporcionar respuestas rápidas y estructuradas, escritos con un aparente buen nivel de contenido (Cotton et al., 2023) o bien, simplificar temas complejos a los usuarios. Adicionalmente y por lo general, logra entender el tema que se está abordando en una conversación y dar respuestas coherentes dentro del contexto. En la historia del desarrollo de ChatGPT esas habilidades son un avance significativo respecto a las primeras versiones que se desarrollaron en 2018 que no ofrecían esas posibilidades (Samala et al., 2024). También se postula que la IA podría ayudar a mejorar los procesos de enseñanza y aprendizaje, auxiliar en el desarrollo de habilidades en los estudiantes y hacer más eficientes las actividades académicas (Pisica et al., 2023). Por ejemplo, por la aproximación personalizada y asincrónica que permitiría un acceso al conocimiento en periodos más flexibles, por el uso de nuevas herramientas para transmitir el conocimiento de manera más lúdica o extensa como asistentes de voz y porque puede proveer diferentes recursos como traductores, programas de estudio, experiencias interactivas y retroalimentación instantánea.

Sin embargo, también hay algunos aspectos que necesitan ser atendidos por su falta de claridad o desarrollo, como la existencia de errores o inexactitudes. Por ejemplo, se sabe que algunos modelos extensos de lenguaje (LLM por sus siglas en inglés) como ChatGPT pueden inventar literatura científica e incluso construir detalles sobre esa literatura (las llamadas “alucinaciones”) (Bahrini et al., 2023). Como

podemos imaginar, tales alucinaciones y su uso representan un problema potencial en la academia si las personas usan esa información inventada sin verificarla. De hecho, si se observa con atención, ChatGPT presenta en su pantalla de chat un aviso de “descargo de responsabilidad” que indica “ChatGPT puede cometer errores. Comprueba la información importante”. A lo que podríamos preguntarnos ¿qué información es “importante”? y si toda la información obtenida por estos medios debería ser verificada por los usuarios.

Otro aspecto crítico es la falta de inteligencia emocional en la IA, lo que genera interacciones mecánicas desprovistas de empatía genuina. Existen también preocupaciones sobre la privacidad de la información, ya que los usuarios no solo obtienen información de las IAs, sino que también proporcionan información personal (por ejemplo durante el registro) y contribuyen al entrenamiento de estos sistemas con sus solicitudes específicas (Bahrini et al., 2023). Todos hemos escuchado que cuando alguien realiza búsquedas en la web o las redes sociales sobre algún tema o necesidad en particular (por ejemplo, un viaje), después empiezan a recibir anuncios publicitarios relacionados con lo buscado previamente. Se piensa incluso que en el ámbito de la academia, la retroalimentación de estudiantes y académicos es potencialmente más especializada e incluso útil, pero aún no está claro cómo se usa dicha retroalimentación y a quién beneficia.

La dependencia excesiva o gradual que se puede llegar a tenerse de la IA (Samala et al., 2024) es otro aspecto por cuidar. Por un lado, si se llegase a tener una dependencia excesiva de las herramientas con IA (como ChatGPT, *Inciteful*, *Litmaps*, *Jenni*, etc.), se corre el riesgo de fomentar una falta de autonomía del usuario pudiendo conllevar a una disminución del pensamiento crítico y de la motivación (Samala et al., 2024). Adicionalmente, si la mayoría de las personas empiezan a usar las herramientas con IA para realizar tareas y en consecuencia, las realizan más rápidamente que una persona que no las usa, se podría empezar a generar la presión de usar dichas herramientas con IA para mantener la competitividad, creando así la llamada “dependencia por competitividad” (Bahrini et al., 2023). Aunque no hay que perder de vista que hacerlo rápido no significa hacerlo bien y justo ahí es donde toma importancia la responsabilidad que cada persona puede dar al uso de las herramientas con IA.

Otro tema preocupante son los potenciales sesgos en las respuestas generadas por algunas herramientas con IA. En teoría, las respuestas generadas con IA deberían estar desprovistas de sesgos, discriminación y ser objetivas. Sin embargo, algunos expertos advierten que las respuestas podrían contener sesgos de origen debido al proceso de recopilación de datos y cómo se clasifica la información que se usa para generar las respuestas. Como dato curioso, se ha indicado que para el entrenamiento inicial de ChatGPT se utilizaron los artículos publicados por el New York Times, quien después demandó a ChatGPT (Bahrini et al., 2023), pero fuera de eso, podemos deducir que si basaron su entrenamiento (al menos inicial) en los artículos del New York Times, las respuestas basadas en esa información podrían contener la visión de ese diario estadounidense. Adicionalmente, el funcionamiento automático y predictivo de la IA puede llevar a generalizaciones incorrectas y su funcionamiento poco claro no se alinea con principios democráticos (Cossette-Lefebvre & Maclure, 2023) o de transparencia. Por lo anterior, resulta importante cuidar y garantizar que esos aspectos en la compilación y generación de datos y funcionamiento de la IA se enmarquen en regulaciones éticas, de justicia, equidad, objetividad y no discriminación.

De ahí que otros autores han identificado problemas adicionales a atender como el de las bases éticas (Bahrini et al., 2023) y legales (Regulation (EU) 2024/1689) del uso de la IA. Por un lado, se postula que son necesarias bases éticas como la aplicación de principios y prácticas que aseguren que el desarrollo, la implementación y el uso de las IAs de manera transparente, equitativa, sin discriminación, responsable y respetuosa de los derechos y el bienestar de todas las personas. Por otro lado, parece necesario disponer de leyes que regulen tanto a los desarrolladores como a las empresas que utilizan IA y que provean de lineamientos de uso claros a los usuarios. La Comunidad Europea (CE) ha tomado la delantera en ese sentido al publicar la primera legislación sobre IA, conocida como la Ley de IA, de la CE (Regulation (EU) 2024/1689) (Figura 3).

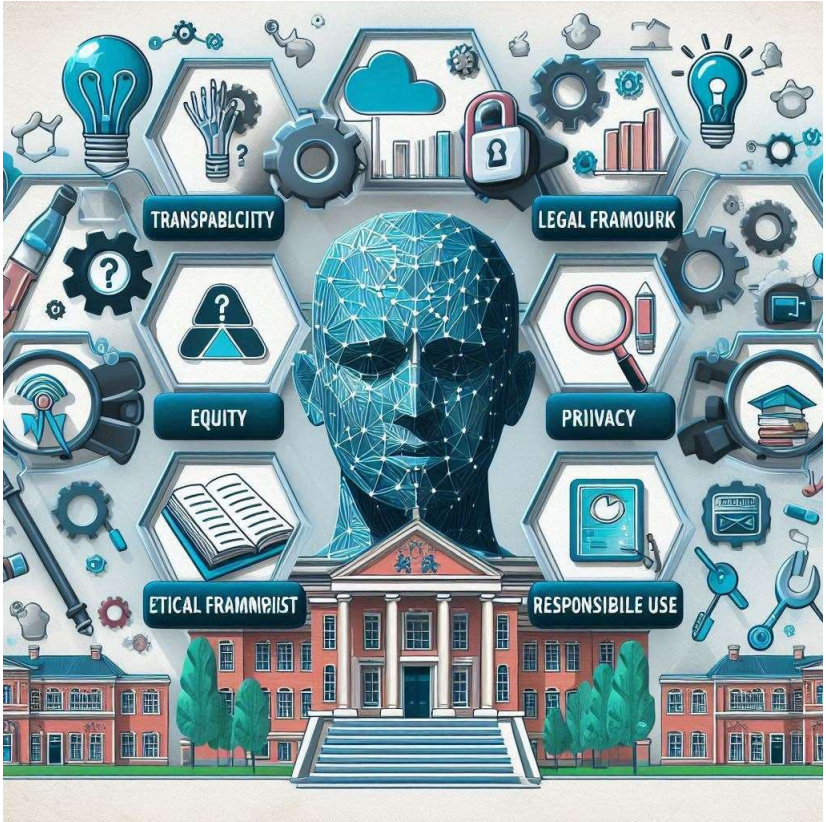


Figura 3. Creada con la herramienta Copilot (DALL-E 3) con IA usando el prompt: “Crea una imagen donde se aprecien los desafíos del uso de las herramientas con IA en la academia como la transparencia, la equidad, el marco legal, la privacidad, el uso ético y responsable”

# 5. PERSPECTIVAS

Como se describió previamente, el uso de la IA en la academia ofrece diferentes beneficios, como la generación rápida de información, la interacción personalizada, el aprendizaje adaptativo y una creciente accesibilidad.  Existen muchas herramientas que pueden ayudar en diferentes etapas del quehacer académico. Sin embargo, antes de centrarse en mejorar estas herramientas y extender los beneficios, es fundamental atender los aspectos éticos y legislativos que regulen la creación y el uso de herramientas basadas en IA. La regulación puede provenir tanto de los gobiernos, como en el caso de la Comunidad Europea, como de las instituciones académicas. En efecto, algunos expertos proponen que las instituciones podrían desarrollar sus propias infraestructuras de IA, estableciendo lineamientos éticos y claros sobre su creación. Esto no solo fomentaría la transparencia, la independencia y el uso responsable (Bahrini et al., 2023), sino que también permitiría mejorar las herramientas de IA a través de la retroalimentación, permitiendo aplicaciones más versátiles, robustas, equitativas y útiles. Implementar modelos piloto y probarlos gradualmente ayudaría a fortalecer el uso responsable de estas tecnologías (Yawson, 2024). Adicionalmente, la integridad académica se robustecería frente al uso de la IA si se implementaran estrategias como el desarrollo de políticas de uso, la capacitación, el aprendizaje y la mejora de los métodos de detección de plagio (Cotton et al., 2023; Miao & Holmes, 2023). Con el establecimiento de bases éticas y jurídicas, el desarrollo de futuras mejoras o innovaciones relacionadas con la IA se harían en un marco de respeto, ética y responsabilidad.

En cuanto a la mejora de las herramientas con IA, se ha identificado que los sesgos presentes en los datos de entrenamiento y las limitaciones en el razonamiento de algunas herramientas son áreas potenciales por optimizar (Yawson, 2024). Los datos utilizados en el entrenamiento de la IA generativa es un componente clave que debería ser revisado y evaluado por expertos externos para evitar sesgos y

huecos de información. Además, aunque herramientas como ChatGPT pueden generar respuestas rápidas, aún tienen capacidad limitada para realizar análisis profundos y razonamientos complejos, lo que posiblemente lleva a las llamadas “alucinaciones” o inventos de literatura científica. Estas limitaciones deberían ser abordadas antes de que la IA generativa se adopte de manera generalizada en la academia.

Algunos expertos sugieren que la IA debe ser adoptada de manera gradual como herramientas complementarias (el principio de la equivalencia sustancial), sin reemplazar la labor de los usuarios y siempre respetando principios éticos y de responsabilidad. Sin embargo, esta estrategia podría seguir generando dependencia hacia las compañías que desarrollan estas tecnologías, sin garantizar la transparencia ni la ausencia de sesgos en los datos de entrenamiento (Yawson, 2024). Por otro lado, hay quienes proponen que se aplique el principio de precaución antes de integrar la IA en la academia. Esto implica evaluar primero los riesgos potenciales de la IA, no solo sus beneficios potenciales, antes de adoptarla de manera más generalizada. Lo cierto es que la IA ya se está volviendo común en el ámbito académico, con muchos estudiantes utilizando herramientas como ChatGPT. Por ello, es urgente establecer las bases éticas y legislativas que regulen su uso, así como capacitar a los integrantes de la comunidad académica para manejar estas tecnologías de manera ética y responsable. Esto permitirá ir a la par del desarrollo tecnológico y mantener los estándares académicos asegurando un uso consciente y regulado de estas herramientas en la educación superior.

# 6. REFERENCIAS


Bahrini, A., Khamoshifar, M., Abbasimehr, H., Riggs, R. J., Esmaeili, M., Majdabadkohne, R. M., & Pasehvar, M. (2023). ChatGPT: Applications, Opportunities, and Threats. *2023 Systems and Information Engineering Design Symposium, SIEDS 2023*, 274–279. https://doi.org/10.1109/SIEDS58326.2023.10137850

Cossette-Lefebvre, H., & Maclure, J. (2023). AI’s fairness problem: understanding wrongful discrimination in the context of automated decision-making. *AI and Ethics*, *3*(4), 1255–1269. https://doi.org/10.1007/s43681-022-00233-w

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1–12. https://doi.org/10.1080/14703297.2023.2190148

Jackson, P. (1998). *Introduction to Expert Systems* (3rd ed.). Addison-Wesley Longman Publishing Co., Inc.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

McCarthy, J. (1974). Artificial intelligence: a paper symposium: Professor Sir James Lighthill, FRS. Artificial Intelligence: A General Survey. In: Science Research Council, 1973. *Artificial Intelligence*, *5*(3), 317–322. https://doi.org/10.1016/0004-3702(74)90016-2

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://doi.org/10.54675/EWZM9535

Moor, J. (2006). The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years. *AI Magazine*, *27*(4), 87–91.

Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: an introduction. *Journal of the American Medical Informatics Association*, *18*(5), 544–551. https://doi.org/10.1136/amiajnl-2011-000464

Pisica, A. I., Edu, T., Zaharia, R. M., & Zaharia, R. (2023). Implementing Artificial Intelligence in Higher Education: Pros and Cons from the Perspectives of Academics. *Societies*, *13*(5), 118. https://doi.org/10.3390/soc13050118

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving Language Understanding by Generative Pre-Training. In *OpenAI blog*.

Radford, A., Wu, J., Child, R., Luan, D., Amodel, D., & Sutskever, I. (2019). Language Models are Unsupervised Multitask Learners. In *OpenAI blog* (Vol. 1).

Samala, A. D., Rawas, S., Criollo-C, S., Bondarenko, O., Gentarefori Samala, A., & Novaliendry, D. (2024). Harmony in Education: An In-Depth Exploration of Indonesian Academic Landscape, Challenges, and Prospects Towards the Golden Generation 2045 Vision. *TEM Journal*, 2436–2456. https://doi.org/10.18421/TEM133-71

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), 484–489. https://doi.org/10.1038/nature16961

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, *LIX*(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. https://doi.org/10.5555/3295222.3295349

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45. https://doi.org/10.1145/365153.365168

Yawson, R. M. (2024). Perspectives on the promise and perils of generative AI in academia. *Human Resource Development International*, 1–12. https://doi.org/10.1080/13678868.2024.2334983

Esta obra está bajo una licencia de Creative Commons
Atribución-NoComercial-Compartirlgual 4.0 Internacional.