

Límites y desafíos de los grandes modelos de lenguaje para la escritura académica y científica: una revisión crítica en función del “uso experto”



Victoria Scotto

Consejo Nacional de Investigaciones Científicas y Técnicas / Instituto de Investigaciones en Humanidades y Ciencias Sociales, Universidad Nacional de La Plata, Argentina
vscotto@fahce.unlp.edu.ar

Trabajo recibido el 19 de febrero de 2025 y aceptado el 15 de abril de 2025.

Resumen

Desde 2022 el mundo científico ha acusado recibo del impacto que representó la masificación del ChatGPT. Su utilización en contextos académicos y científicos ha despertado el interés de investigaciones diversas que, hasta ahora, no han sido ordenadas en función de criterios de uso por parte de los individuos, y tampoco han hecho foco en las características comunes que presentan los textos que genera. El presente trabajo es una exposición ordenada de los problemas y límites del uso de grandes modelos de lenguaje en contextos académicos y científicos identificados en la bibliografía especializada, realizada en función de un criterio novedoso, el de uso experto. Por un lado, se referirán consecuencias del uso no experto, y por otro, se desarrollará qué problemas y limitaciones se observan en textos generados por estos modelos, problemas que pueden observarse (pero aún no prevenirse) con el uso experto: la falta de trazabilidad y los sesgos, las alucinaciones, el *parroting*, el plagio, el desarrollo argumentativo endeble y las exposiciones inconsistentes, incoherentes y contradictorias.

Palabras clave: grandes modelos de lenguaje, escritura científico-académica, alfabetización académica, alfabetización digital, ChatGPT.

Limits and challenges of large language models for academic and scientific writing: a critical review based on “expert use”

Abstract

Since 2022, the scientific world has acknowledged the impact of the massification of ChatGPT. Its use in academic and scientific contexts has sparked the interest of various research projects that, until now, have not yet organized their analysis according to criteria of individual expert use, nor have they focused on the common characteristics of the texts they generate. This paper is an organized exposition of the problems and limitations of the use of large language models in academic and scientific contexts observed in the specialized literature, based on a novel criterion: expert use. On the one hand, we will address the consequences of non-expert use, and on the other, we will name and explain the problems and limitations observed in texts generated by these models, problems that can be observed (but not yet prevented) with expert use: lack of traceability and biases, hallucinations, parroting, plagiarism, weak argumentative development, and inconsistent, incoherent, and contradictory expositions.

Keywords: large language models, scientific and academic writing, academic literacies, digital literacies, ChatGPT.

Limitações e desafios dos principais modelos de linguagem para escrita acadêmica e científica: uma revisão crítica baseada no “uso especializado”

Resumo

Desde 2022, o mundo científico sente o impacto do uso generalizado do ChatGPT. Seu uso em contextos acadêmicos e científicos tem despertado o interesse de diversos projetos de pesquisa que, até o momento, não foram organizados segundo critérios de uso por indivíduos, nem se concentraram nas características comuns dos textos que gera. Este artigo é uma exposição organizada dos problemas e limitações do uso de grandes modelos de linguagem em contextos acadêmicos e científicos identificados na literatura especializada, com base em um novo critério: o uso por especialistas. Por um lado, serão discutidas as consequências do uso não especializado e, por outro, serão desenvolvidos os problemas e limitações observados em textos gerados por esses modelos, problemas que podem ser observados (mas ainda não prevenidos) com o uso especializado: falta de rastreabilidade e vieses, alucinações, *parroting*, plágio, fraco desenvolvimento argumentativo e exposições inconsistentes, incoerentes e contraditórias.

Palavras-chave: grandes modelos de linguagem, escrita científico-acadêmica, alfabetização acadêmica, alfabetização digital, ChatGPT.

1. Introducción

1.1. Una introducción (amigable) a cómo funcionan los GML

Desde 2023, el mundo científico ha acusado recibo del impacto que representó la masificación del ChatGPT-3, la tercera versión integrada a un *chatbot* del modelo computacional capaz de generar texto en lenguas naturales humanas de forma automática (GPT, *Generative Pre-Trained Transformer*). Este modelo computacional, llamado en inglés *Large Language Model* (LLM) o gran modelo de lenguaje (GML) es un tipo de inteligencia artificial (IA) generativa. Opera en base a redes neuronales computacionales¹, específicamente de transformadores o de tipo transformacional², como lo indica el nombre de GPT: esto quiere decir que opera por medio de un modelo complejo en el que se ingresan datos que luego se asocian entre sí y por medio de esta asociación los datos sufren transformaciones no reversibles (es decir que los datos, una vez ingresados, se transforman y pasan a formar parte del modelo y no vuelven a su estado original). Si bien la mayor parte de los GML más usados, como Meta AI o GPT-3 en sus diferentes versiones, responden completamente a este modelo, otros GML lo combinan con diferentes mecanismos como los motores de búsqueda.³ La interfaz de funcionamiento de estos modelos, por lo general, es la de un *chatbot*, un “contestador automático” que en lugar de ofrecer opciones limitadas (como los *chatbots* que ofrecen ciertas empresas para atención del público) inserta el GML como reemplazo de las opciones precargadas. Por motivos de practicidad y de popularidad, en adelante nos referiremos de forma alternada a estos modelos —como GPT-3— como GML o *chatbot*, aunque no todos los *chatbots* cuenten con GML o viceversa.

En este punto, destacaremos dos características que interesan a los efectos del análisis de la generación de textos por parte de los GML. En primer lugar, la operación básica que habilita la generación de texto por parte de un modelo es la de predecir la siguiente palabra más probablemente adecuada

1 Las redes neuronales son un modelo computacional de aprendizaje automático (también conocido como *machine learning*). Toma su nombre e intenta emular la forma de conexión y asociación de las neuronas de los cerebros animales. La estructura de estas redes asocia nodos de información de un programa por medio de señales numéricas que luego se envían a otros nodos, incorporando la información del primer nodo. Esta asociación de nodos construye una capa de información; en un modelo a gran escala, como los GML, hay una gran cantidad de capas que procesan ingentes cantidades de información. Las diversas capas de esta red neuronal contienen una alta cantidad de datos organizados no solo en nodos, sino fundamentalmente en las asociaciones entre ellos: lo que permite este tipo de tecnología es elaborar modelos estadísticos a gran escala, puesto que permite extraer patrones y asociaciones entre elementos a partir de la recurrencia de asociaciones predominantes o no predominantes en cada nodo de información de las múltiples capas de una red neuronal.

2 Un transformador es un tipo de red neuronal computacional. Su procedimiento utiliza decodificadores y codificadores, y un mecanismo de autoatención para procesar entradas (Vaswani *et al.* 2017). Un transformador, en este caso de texto, transforma su objeto en una entrada: segmenta el texto, reemplaza sus términos en *tokens* (identificadores numéricos), para luego vectorizarlos y que estos entren en asociación con otros identificadores representados en esa entrada. Los vectores, una vez asociados, sufren transformaciones no reversibles que los insertan en un circuito mayor. El mecanismo de autoatención, que asigna diversos grados de importancia a diferentes elementos de la entrada, permite (re)construir computacionalmente la estructura de dependencias lógicas y sintácticas del texto original. A la vez, permite aportar datos que se almacenan como asociaciones más y menos predominantes en cada uno de los vectores. Este procedimiento, llevado adelante a gran escala, y habilitando la asociación entre los vectores y sus posiciones relativas en las diversas estructuras, permite realizar predicciones estadísticamente adecuadas sobre una palabra ubicada a continuación de un texto.

3 Este es el caso de Perplexity, un motor de búsqueda que utiliza un GML para producir respuestas por medio de la elaboración de resúmenes del contenido de sitios web que encuentra su motor.

en función del *prompt* introducido por el usuario del modelo. Así lo explica el artículo principal de OpenAI que describe GPT-2, una versión previa:

Nuestro modelo, llamado GPT-2 (sucesor de GPT), fue entrenado simplemente para predecir la siguiente pregunta con 40 gigabytes de texto proveniente de internet [...] GPT-2 es un gran modelo de lenguaje transformacional basado en mil quinientos millones de parámetros, entrenado en una base de datos de 8 millones de sitios web. GPT-2 fue entrenado con un objetivo simple: predecir la siguiente palabra, dadas todas las previas palabras en algún texto. La diversidad de la base de datos genera que este simple objetivo implique demostraciones espontáneas de una multiplicidad de tareas en diferentes áreas. GPT-2 es una versión magnificada de GPT, con diez veces más parámetros, entrenados en diez veces más cantidad de datos (OpenAI 2019)⁴.

De este texto también se puede desprender el segundo elemento clave para comprender cómo operan estos GML: su construcción se sostiene en un modelo estadístico de asociaciones alimentado con una ingente masa textual, que luego se pone al servicio del modelo predictivo. Esto quiere decir que cuando un usuario ingresa una pregunta o un texto con instrucciones (un *prompt*), los *softwares* de este tipo generan una respuesta en base a lo que los textos que se usaron para entrenarlo indican que es lo más (estadísticamente) probable que siga al *prompt*; realiza esa predicción a partir de los miles de millones de otros textos similares que el modelo ya ha descompuesto en datos asociados entre sí. Estos GML son, entonces, “generadores” de texto en el mismo sentido en el que la aplicación *Gmail* “genera” un saludo completo cuando predice que luego de “quedo a la espera” es probable que el usuario desee agregar “de su respuesta”, debido a los miles de millones de correos anteriores que han utilizado esa fórmula.

La escala masiva de su entrenamiento permite la generación de textos en temas diversos y con muy buenos resultados, especialmente cuando se ingresan *prompts* adecuados: la precisión en las instrucciones y la aparición de una mayor cantidad de palabras, que el modelo transforma en datos y que condicionan los resultados, permite optimizar las condiciones de generación de textos. A esta escala del entrenamiento se incorpora, además, el reentrenamiento a nivel profundo del *software*: si bien es posible que un usuario ingrese datos que alteren las respuestas de su chat privado, a partir de su historial, este entrenamiento individual (llamado *fine tuning*) no necesariamente implicará un reentrenamiento del modelo, que, en general, solo se produce a partir de una intervención directa en su programación.

Este último tipo de intervenciones debe distinguirse de la intervención que hacen los informáticos que hacen el mantenimiento del modelo o que corrigen sectores de su programación. Este tipo de intervenciones no busca “personalizar” los textos generados por el modelo, sino que suelen querer,

4 En el original: “Our model, called GPT-2 (a successor to GPT), was trained simply to predict the next word in 40GB of Internet text [...] GPT-2 is a large transformer-based language model with 1.5 billion parameters, trained on a dataset of 8 million web pages. GPT-2 is trained with a simple objective: predict the next word, given all of the previous words within some text. The diversity of the dataset causes this simple goal to contain naturally occurring demonstrations of many tasks across diverse domains. GPT-2 is a direct scale-up of GPT, with more than 10X the parameters and trained on more than 10X the amount of data”. La traducción es de nuestra autoría.

por ejemplo, corregir errores de código (conocidos como *bugs*); comenzar procesos de desaprendizaje (o *unlearning*, según Liu et al., 2024) de datos perniciosos para el modelo o potencialmente peligrosos para los individuos vinculados a ellos; eliminar determinado tipo de respuestas que el modelo ofrece, o identificar *prompts* peligrosos que los programadores consideren que deben sufrir alguna forma de restricción. La mayoría de estas intervenciones, sin embargo, no suelen hacerse sobre el entrenamiento general del modelo, sino que se ejercen mediante diferentes mecanismos de control que analizan el texto crudo con diversos propósitos, como detener repeticiones excesivas o asegurar respuestas acordes a determinados parámetros éticos o de seguridad. Estos mecanismos de control, que los programadores se encargan de sofisticar y ajustar de manera permanente, representan una parte importante de los procesos por los que pasa el texto generado para ser admisible a los ojos de sus usuarios y programadores, y son a la vez una dimensión del programa que resulta más accesible en términos de intervención, en tanto no es necesario corregir aspectos del entrenamiento inicial del programa (masivo y en muchos casos irreversible) sino filtros posteriores. A medida que sus usuarios crecen y el *feedback* aumenta, cada GML se sofisticada y refina sus datos de salida, logrando producir textos que, progresivamente, se asemejan cada vez más a textos producidos por humanos.

La integración de estas diversas capas de procesamiento de información tiene resultados que permiten una amplia variedad de posibilidades para la escritura: los GML no solo sirven para predecir el cierre habitual de un *mail* de carácter formal, sino que pueden utilizarse como puntapié para diversos géneros textuales, para corregir la ortografía, para listar posibles soluciones a problemas, para ensayar traducciones iniciales, para emular estilos y automatizar tareas de escritura, entre otros usos. Bastaría, en principio, apenas un pedido inicial en forma de *prompt*, no demasiado específico, o una orientación o sucesivas correcciones en una ventana de *chat*, para recibir versiones crecientemente sofisticadas de un texto sin demasiado esfuerzo por parte de una persona. Pero, como veremos a continuación, esta práctica no carece de riesgos: especialmente, cuando se hace de ella un uso en el que no media una alfabetización académica y digital completa, y un conocimiento relativamente solvente sobre aquello que el *chatbot* devuelve como datos (aparentemente) certeros.

1.2. Objeto

En este trabajo exploraremos algunos aspectos problemáticos de la utilización de los GML tanto en contextos académicos como científicos; específicamente, en los modos en los que los mecanismos de generación de estos textos afectan directamente las tareas académicas y científicas. Tenemos en cuenta que los GML son un tipo de tecnología “progresivamente agentiva”⁵ cuyo uso y productos tienen características no proyectadas, deseadas o planificadas por sus programadores, y por lo tanto resulta de vital importancia atender rápidamente a la responsabilidad humana en la prevención

⁵ Traducción literal de “increasingly agentive” (Chan et al. 2023).

y anticipación de daños resultantes de este tipo de tecnología (Chan *et al.* 2023). En este caso, la tarea de monitoreo de sus consecuencias se hará en función de una diferenciación entre dos formas de uso que pueden hacerse de un GML en contextos académico-científicos.

El primer objetivo general es ofrecer un concepto novedoso, el de “uso experto”, para colaborar con el análisis de la incidencia de la inteligencia artificial generativa en actividades académicas y científicas. Por “uso experto” nos referiremos a una forma de utilización de un GML por parte de un usuario que posee una alfabetización académica completa (Lillis 2003, 2021) en conjunto con una alfabetización digital completa y actualizada (Barroso y Cabero 2011), y a su vez domina el campo de estudios que tematiza el texto producido. En este sentido, una persona alfabetizada académica y digitalmente puede hacer un uso experto de un GML siempre y cuando lo utilice para producir textos cuyas áreas domina de forma relativa; en caso de que carezca de formación en el asunto por el que consulta al modelo, este uso no será “experto”. A su vez, si quien lo utiliza no sabe qué es o cómo utilizar un GML, no tiene una alfabetización académica completa o no domina completamente la lengua con la que ingresa un *prompt*, no podrá advertir problemas textuales básicos que presentan los GML, y tampoco podrá hacer un uso experto. Así, la alfabetización digital y académica completa de un usuario es *conditio sine qua non* para el uso experto, pero no todo uso por parte de una persona alfabetizada digital y académicamente puede ser calificado de experto. Dentro de este objetivo general incluimos, entonces, el objetivo específico de exponer las dificultades intrínsecas que el uso no experto presenta en la bibliografía especializada: abordaremos el uso de esta tecnología para el cumplimiento de tareas académicas en el contexto de la universidad, restringiendo el análisis específicamente a su uso por parte de estudiantes del nivel superior que no han terminado su formación y, en este sentido, hacen uso inexperto de la herramienta.

El segundo objetivo general de este trabajo es ofrecer un panorama de las limitaciones que suponen los textos generados por los GML incluso cuando sus usuarios tienen alfabetizaciones académicas y digitales completas y conocen el campo con el que trabajan. Ofreceremos una sistematización de las limitaciones textuales que aparecen en la bibliografía especializada e incluiremos, además, un problema prácticamente no identificado más que lateralmente, que es la composición textual endeble posibilitada por el carácter no determinista de la generación de textos de los GML.

1.3. Metodología

El presente artículo presenta una sistematización realizada en función de un aspecto puntual del análisis: no se trata de un mero relevamiento, sino de un estudio de los aportes bibliográficos identificados en función de las limitaciones que presentan los textos producidos por los GML a la luz del concepto de “uso experto”. Aportaremos argumentos en torno a las dificultades que el uso de los GML supone para usuarios no expertos y también explicaciones a propósito de cómo su composición textual interfiere directamente con la construcción de artículos científicos de calidad.

En ambos casos, la sistematización realizada en este artículo se sostiene en la bibliografía especializada; específicamente, en dieciocho artículos publicados en revistas con referato, en inglés y en español, publicados entre mayo de 2023 (seis meses después del lanzamiento de ChatGPT-3) y mayo de 2024. Estos artículos, o bien ofrecen en parte o en todo el trabajo un análisis de problemas y beneficios del uso experimental de esta herramienta en contextos científicos o académicos, o analizan el producto de un GML puntual en base a datos propios o a una sistematización de la bibliografía existente⁶. Los artículos que cumplieron el criterio de inclusión fueron publicados en revistas indexadas de cuatro continentes (Europa, América, Asia y Oceanía), con referato, disponibles en la red *Google Scholar*. Se encontraron a través de una búsqueda por palabras clave en inglés (“LLM”, “AI”, “scientific use”, “university” y “higher education”). De un total de cuarenta artículos que se obtuvieron en el relevamiento se descartaron veintidós por tratarse de ensayos o elaboraciones teóricas, o publicaciones preliminares que comunican únicamente metodologías o encuestas a realizar. Tampoco nos detendremos en artículos dedicados a analizar los aspectos por los cuales el mero desarrollo y utilización de GML implican problemas éticos⁷.

Los artículos que reunieron los requisitos para ser tenidos en cuenta se mencionan a continuación: entre los que abordan usos no expertos, Darwin *et al.* (2023); Dempere *et al.* (2023); García Peñalvo, Llorens-Largo y Vidal (2023); García Sánchez (2023), Kostka y Toncelli (2023); Montenegro Rueda *et al.* (2023); Zhang y Tur (2023) y Niloy *et al.* (2024). Entre los que refieren problemas de los textos que pueden observarse incluso en el uso experto, se trabaja con Al Afnan *et al.* (2023); Alkaissi y McFarlane (2023); Imran y Almusharraf (2023); Lo (2023); Mai, Da y Van Hahn (2023); Martino, Ianelli y Truong (2023); Qi, Zhu y Wu (2023); Ray (2023); Sullivan, Belly y McLaughlan (2023); Vargas-Murillo, Pari-Bedoya y Guevara-Soto (2023) y nuevamente con Dempere *et al.* (2023) y Niloy *et al.* (2024). Las referencias por fuera de estos trabajos mencionados se utilizan únicamente como antecedentes. De los artículos seleccionados se extrajeron los aspectos en los que se elaboraban las limitaciones de los textos y se los reunió luego en dos grupos: los suscitados por el uso inexperto y aquellos que son independientes de un posible mal uso; así se presentan y se explican más adelante.

6 El corte temporal se debe a que este trabajo se basa en una investigación realizada para las *Primeras Jornadas de Políticas Lingüísticas en Publicaciones Científicas* organizada por el Instituto de Lingüística de la Universidad de Buenos Aires en agosto de 2024.

7 Aquí solo se mencionarán algunos: los peligros del determinismo tecnológico, la influencia monopólica en la sociedad de la información, la falta de soberanía de datos de sus usuarios, la monetización de obras no autorizadas para ello y el extractivismo de conocimiento, la pérdida de recursos naturales que supone cada hora de utilización de un GML y la pérdida de soberanía territorial que implica la reorientación de esos recursos (Floridi 2023, Pasquinelli y Joler 2023, Zhuk 2023, Christou 2024 y Archuby 2025, entre otros). Otros problemas éticos no observados en la bibliografía, y que fueron discutidos en torno a este asunto en las *Primeras Jornadas de Políticas Lingüísticas en Publicaciones Científicas*, podrían incluir dilemas de carácter filológico, como la pérdida de la humanidad en la escritura, el relegamiento del registro textual de la humanidad a algo no humano, la desaparición de la responsabilidad sobre la autoría de los textos y la demanda de que personas humanas lean textos que nadie ha tenido la deferencia de escribir.

2. El uso en ámbito académico

2.1. Definición del problema: el uso inexperto

La primera hipótesis que presenta este trabajo es que no es posible hacer un uso experto de una herramienta como lo es cualquier GML si no se cuenta con una alfabetización digital y académica completas. Para desarrollar esta hipótesis se propondrán definiciones de estas dos formas de alfabetización y las implicancias de su ausencia *a priori*, para luego, en apartados subsiguientes, exponer qué consecuencias identifica la bibliografía en esos casos.

La principal característica del uso de GML en contexto universitario en particular y en el de educación superior en general es que presenta, de acuerdo con la bibliografía disponible, un porcentaje de uso experto mínimo. Según Sullivan, Belly y McLaughlan (2023) la mayoría de los estudios sobre uso de IA indican que la mayor parte de las personas posee una alfabetización digital mínima con respecto a cómo funcionan estas herramientas, y en general su conocimiento está muy mediado por la prensa, que frecuentemente se presenta como sensacionalista o ignora problemas éticos y políticos de su uso. Este panorama presentado por los autores puede observarse en la tendencia hallada en diferentes artículos científicos que registran que el mayor problema de la utilización de esta aplicación en contextos de educación superior radica en la falta de conocimiento sobre cuáles son los límites de lo que estos *chatbots* pueden hacer, cómo lo hacen y en qué medida requieren corrección o verificación cuando ofrecen información. Esto está relacionado, naturalmente, con una diferenciación entre dos dimensiones de la alfabetización que los trabajos anteriores no mencionan: la digital y la académica.

La noción de alfabetización digital implica la capacitación imprescindible para utilizar dispositivos tecnológicos, pero también para sobrevivir y actuar críticamente en una sociedad centrada en la información provista por ellos (Barroso y Cabero 2011). La evolución hacia la inmediatez y el desarrollo de interfaces cada vez más alejadas del lenguaje de máquina y cada vez más orientadas hacia un uso intuitivo (Macaranas, Antle y Riecke 2015) han colaborado los últimos diez años de forma progresiva hacia la ampliación del uso de software sin entrenamiento ni reflexiones conscientes sobre los procesos que ocurren detrás de la interfaz. El fenómeno de expansión del uso de internet en un mundo profundamente desigual ha dado lugar a una marcada disparidad en la alfabetización digital (Tinmaz, Fanea-Ivanovici y Baber 2022), y la sofisticación de su programación, a su vez, colaboró con un fenómeno muy referido en los últimos años: la opacidad de los algoritmos (Bonneau, Grondin-Robillard, Ménard y Mondoux 2022). Este fenómeno produce que la sofisticación de los algoritmos que regulan interfaces hoy hayan terminado por resultar inaccesibles para la gran mayoría de las personas que las utilizan, colaborando con que las personas sin una alfabetización digital completa confíen en la información que diversos programas les ofrecen sin conocer (quizás en lo absoluto) cómo se produce esta información.

La alfabetización académica (Lillis 2021), por otra parte, refiere a un proceso propio de la educación superior, que requiere la enseñanza y el aprendizaje de un tipo particular de escritura y lectura asociadas a la academia por oposición a otras esferas de uso, y que es parte de la responsabilidad de las instituciones. Esta perspectiva, que pone el foco en el aprendizaje de la escritura académica, según Lillis, incluye además un aspecto nodal en su concepción de la lengua: la entiende como un conjunto de prácticas discursivas socialmente situadas e inscriptas ideológicamente; por ello Lillis entiende que no debe suponerse que los estudiantes reconocen estos géneros de forma implícita sin ningún tipo de asistencia, sino que se requiere un proceso consciente de señalamiento de sus estructuras (Lillis 2003, 194). La alfabetización académica promueve la familiarización con los géneros académicos, su lectura y escritura, y se caracteriza por su autoconciencia lingüística, pero también por proponer un paradigma de coexistencia de géneros discursivos, voces y perspectivas diversas que no pretenden anularse entre sí; busca enseñar advirtiendo las reglas específicas de cada género como marcadas, no siempre evidentes, siempre políticamente situadas. El reconocimiento de estas reglas permite un acercamiento reflexivo a la escritura y la lectura de textos académicos tanto como permite identificar aquellos elementos que los diferencian de textos producidos en otras esferas. En este punto, una alfabetización académica completa permite que los sujetos adviertan rupturas a las convenciones académico-científicas, como la falta de sustento de afirmaciones, la no consecución entre hipótesis, desarrollo y conclusiones, y el plagio, entre otras.

Entendemos entonces que, en la era digital, y en la era de la información, no es únicamente necesario conocer los mecanismos básicos de utilización de computadoras o de celulares para ejercer de forma crítica la ciudadanía: es necesario, por un lado, poseer una alfabetización digital completa; esto es, contar con conocimientos básicos sobre el funcionamiento de la producción y la circulación de la información, y utilizar estratégicamente el software a disposición. A su vez, tiene una importancia crítica poseer una alfabetización académica completa, que le aporta a cualquier sujeto la perspectiva suficiente para interpretar los textos que lee, analizar sus limitaciones y producir textos que respondan a sus necesidades educativas, pragmáticas o de otro tipo. A partir de aquí se abre un interrogante a propósito de cuál será el rol, en el futuro, de los GML en la alfabetización académica y digital, y cuáles son las consecuencias a mediano y largo plazo de su utilización masificada: quedará para futuras investigaciones avanzar en las respuestas a estas preguntas.

2.2. *El uso inexperto evidenciado en la bibliografía*

El desarrollo de una alfabetización digital (y, agregamos, académica) pareciera ser fundamental para lograr un uso experto de los GML. Darwin *et al.* (2023) manifiestan el peligro, expresado por los mismos estudiantes, de que un uso inadecuado o no selectivo de inteligencia artificial puede generar “endoso cognitivo”, una tendencia a la delegación irreflexiva del proceso de lectura e interpretación de las tareas para que las realice, sin supervisión,

el GML.⁸ En concordancia, Zhang y Tur (2023) señalan el peligro de que los estudiantes se vuelvan excesivamente dependientes de este tipo de tecnologías como una de las preocupaciones más destacadas de docentes y académicos en la bibliografía sobre educación; y casi todos los trabajos mencionados hasta ahora advierten de alguna forma que no parecería ser conveniente realizar un uso acrítico de *chatbots* como ChatGPT. Algunos estudios, como el de Kostka y Toncelli (2023), presentan evidencia de elaboración propia extraída de encuestas realizadas a estudiantes de inglés como segunda lengua y concluyen que, si bien la importancia de la alfabetización digital precede al lanzamiento de ChatGPT-3, la popularización del uso de los GML resalta su urgencia, fundamentalmente debido a que los jóvenes necesitan poner una mirada crítica sobre la información que estos modelos presentan. En esta misma línea, la sistematización bibliográfica realizada por Dempere *et al.* (2023) sostiene que la alfabetización debería incluir también a los docentes de educación media y superior, posicionamiento compartido por otros autores (García Peñalvo, Llorens-Largo y Vidal 2023; García Sánchez 2023; Zhang y Tur 2023).

La recurrencia de aparición de estas preocupaciones parece ir en concordancia con la creciente evidencia del uso inexperto en ámbitos educativos, sostenida en las conclusiones finales de al menos tres trabajos sistemáticos de análisis de datos a propósito del uso de *chatbots* en educación superior (Niloy, *et al.* 2024; Montenegro Rueda *et al.* 2023; García Sánchez 2023). El estudio de García Sánchez (2023) se dedica a medir el impacto de esta tecnología en la educación superior mediante cuestionarios realizados a estudiantes de nivel universitario. En este estudio, el 79% de los estudiantes que utilizaban ChatGPT regularmente se encuentra en desacuerdo o muy en desacuerdo con la afirmación de que esta herramienta haya mejorado su capacidad de realizar investigaciones académicas, y el 76% expresó insatisfacción con respecto a las respuestas que el *chatbot* le proporcionó para sus tareas académicas, dos números marcadamente diferentes a lo esperable teniendo en cuenta las tasas de éxito de ChatGPT-3 en exámenes de prueba realizados en ensayos como los de Niloy *et al.* (2024). La principal hipótesis de García Sánchez frente a los resultados integrales de su encuesta es que existe una falta de conocimiento sobre cómo hacer un uso adecuado de la herramienta.

El trabajo sistemático de Montenegro Rueda *et al.* (2023) aporta un panorama de las ventajas y desventajas que diversos autores recogen a propósito del uso del ChatGPT y tecnologías similares. En orden de frecuencia, estas fueron: representa una novedad y eso le aporta impacto; genera satisfacción en los estudiantes; representa un soporte de trabajo; se trata de una herramienta utilizada tanto por docentes como estudiantes; ahorra u optimiza tiempo; es eficiente; sirve para el desarrollo de ideas creativas; puede desarrollar el pensamiento crítico. Por otro lado, entre las desventajas referidas por estudiantes se encuentran la falta de un uso apropiado o de entrenamiento con respecto a la herramienta; que los resultados ofrecidos por la herramienta no son satisfactorios; que los estudiantes desconfían de

8 En el original: "cognitive offloading" (Darwin *et al.* 2023, 12).

la herramienta; que su uso puede replicar desigualdades vinculadas con la brecha tecnológica y les permite a los estudiantes entregar trabajos sin haber aprendido nada. Resulta notorio que mientras que las ventajas tienen menos frecuencia de aparición en la bibliografía (ninguna de las nombradas llega a aparecer más de cinco veces en los doce artículos científicos), el desconocimiento sobre cómo funciona o cómo utilizar correctamente ChatGPT apareció en diez de los doce trabajos examinados por los autores en Montenegro Rueda *et al.* (2023).

El trabajo de Niloy *et al.* (2024) analiza, por medio de encuestas a estudiantes de tres países de Asia, los motivos de utilización de ChatGPT en educación. Y demuestra con los datos que ofrece que no parece haber una correlación directa entre el grado de conocimiento sobre cómo funciona el modelo GPT-3 y la intención de uso, mientras que sí se observa una relación casi directa entre la intención de uso y el uso real del *software*. Esto supone que a pesar de que la mayor parte de sus usuarios en educación superior no conoce los mecanismos que operan detrás de la interfaz, esto no los disuade de utilizarlo.⁹ A su vez, este trabajo cuantifica las motivaciones detrás del uso que hacen estudiantes del ChatGPT: las más importantes se resumen en: ahorro de tiempo y gestión de tareas e inseparabilidad del contenido¹⁰ en los primeros puestos relacionados con intención de uso del programa, y facilidad de acceso, aprendizaje asistido, y mezquindad cognitiva del usuario¹¹ en un segundo nivel de motivaciones vinculadas con su uso posterior.

La falta de uso experto es el mayor problema de la utilización de la herramienta, y sin embargo es inescapable en ámbitos de educación superior si se tiene en cuenta que sus estudiantes no poseen una alfabetización digital y académica completas. A partir de la detección de esta falta de uso experto, Barman, Wood y Pawlowsky (2024) en su artículo sobre el tema proponen un abordaje novedoso y concebido especialmente para estos usos, atendiendo a la falta de lectura crítica que presentan la mayoría de los usuarios de este tipo de modelos. Barman, Wood y Pawlowsky proponen que los GML presenten guías de uso, indicando qué pueden hacer y qué no, y qué tareas requieren intervención y refinamiento por parte del usuario; por último, también consideran que los usuarios deberían ser entrenados en cómo componer *prompts* adecuados y en estrategias y verificación de los *outputs* que produce el GML.

Este enfoque parece presentarse más en línea con el de las alfabetizaciones académicas de Lillis: pretender que el conocimiento de las reglas que estructuran los intercambios lingüísticos es innato o se desarrolla sin

9 Vale la pena aclarar que la totalidad de los datos válidos obtenidos en ese estudio fueron aportados por estudiantes. De ellos, casi el 75% cursaba al momento de la entrevista un Bachelor in Arts (BA), es decir que se trata de estudiantes en sus primeros años de la carrera de grado, mientras que apenas el 20% correspondían a estudiantes de maestría o superior: de esto puede desprenderse que en ningún caso el uso que estudiantes pueden hacer de un *chatbot* tendrá estándares altos de supervisión, ya que los propios estudiantes se encuentran en período de formación.

10 *Inseparabilidad* refiere a un concepto de marketing: la producción de un bien ofrecido y su consumo resultan indistinguibles para el consumidor.

11 Refiere a un término proveniente de la psicología (*cognitive miser*) que la American Psychological Association define como una disposición a buscar respuestas rápidas y adecuadas frente a otras lentas y cuidadosas; se trata de una tendencia general en las personas (APA 2018).

reflexión metadiscursiva implica perder la oportunidad de intervenir para garantizar las alfabetizaciones digitales y académicas que los sujetos necesitan para ejercer su ciudadanía de forma autónoma y crítica. En este punto, solo en el trabajo de Barman, Wood y Pawlowsky encontramos un enfoque que exceda la mera advertencia general y subraye la importancia de la formación de docentes y estudiantes sobre “posibles peligros” y contra el exceso de optimismo en lo relativo al uso de GML, como puede observarse en las conclusiones del trabajo conducido por Dempere *et al.*, por ejemplo. Poner el foco en la necesaria alfabetización académica y digital de los usuarios de GML por medio de guías de utilización puede ser un punto de partida, aunque, como se argumentará a continuación, ni siquiera el uso experto está exento de problemas.

3. Uso en el ámbito científico

Si bien el uso inexperto supone peligros específicos que no deben ser pasados por alto, la forma en la que generan texto los GML supone riesgos para sus usuarios expertos que deben ser tenidos en cuenta. Los referiremos a continuación centrándonos en el uso de los GML en contextos científicos: siguiendo la definición de uso seguro, inicialmente resulta lógico suponer que un científico formado tanto en su tema de trabajo como en cuestiones básicas de alfabetización académica, debería poder utilizar críticamente un GML. Y sin embargo, resulta indispensable no homologar la alfabetización académica con la alfabetización digital: en este punto, vale la pena hacer una salvedad advirtiendo que un científico que conoce vastamente su campo pero no conoce funcionamiento básicos de una IA como un GML podrá caer en los mismos lugares comunes de sobrevaloración de sus productos como cualquier otro usuario. Entendemos en este punto que una alfabetización académica completa no es suficiente si no se considera la importancia de un manejo crítico de la tecnología: suponemos entonces que, nuevamente, solo podrán hacer un uso seguro aquellos científicos que, además, posean una alfabetización digital completa y actualizada.

3.1. Resistencias

Un mes después de su lanzamiento, el grupo *Nature* publicó una advertencia sobre el uso académico y científico de ChatGPT, especialmente dirigida a profesores universitarios (Stokel-Walker 2022); dos meses más tarde, la revista *Science* publicó una nota de su editor afirmando que “ChatGPT es divertido, pero no es un autor” (Thorp 2023) y afirmó que todo manuscrito que el grupo *Science* detecte que ha sido producido con un GML sería automáticamente descartado. Sin embargo, en los primeros meses de 2023 una serie de autores optaron por incluir a ChatGPT entre los autores de sus artículos (cf. Transformer y Zhavoronkov 2022) con el propósito de dar cuenta de su utilización y, al mismo tiempo, evitar acusaciones de plagio. Esta no fue una práctica que se haya mantenido, y sin embargo muchos científicos continuaron utilizándolo sin admitirlo abiertamente (Conroy 2023). En el próximo apartado nos dedicaremos a exponer cinco aspectos, referidos en la bibliografía, en los que los textos producidos por diversos GML son

deficientes para el trabajo científico y el académico y que, esperamos, colabore con disuadir a la comunidad científica del uso no supervisado de estas herramientas en la escritura de trabajos de diversa índole. Se espera que estos cinco aspectos (sesgos y falta de trazabilidad, alucinaciones, plagio, *parrotting* y composición textual endeble) demuestren una importancia tal en la tarea de escritura de trabajos científicos y académicos que no puedan ser obviados cuando se trabaja con *chatbots*. En este sentido, aportaremos argumentos en torno a la composición textual referidos en la bibliografía que, entendemos, interfieren directamente con la construcción de artículos científicos de calidad.

3.2. Límites del modelo

3.2.1. Sesgos y falta de trazabilidad

Uno de los problemas que manifiesta la bibliografía especializada es el de la calidad de la información que presentan los textos que devuelven los *chatbots* de los GML a sus usuarios, especialmente con respecto a sus sesgos (Sullivan, Belly y McLaughlan 2023; Lo *et al.* 2023) pero también debido a la forma en la que su propia construcción atenta contra la trazabilidad de las fuentes que informaron las respuestas del chat (Ray 2023)¹². Como se ha explicado, los GML son programas de predicción del resultado más probable, obtenido a partir de un entrenamiento previo del modelo a base de datos ya ingresados y transformados, y no son capaces de rastrear los textos que permitieron las asociaciones estadísticas que devuelven en forma de predicciones. Si bien las empresas como OpenAI conocen las bases de datos que entrenaron sus GML y podrían reponer las fuentes textuales, estas han asociado palabras con *tokens* numéricos, y las transformaciones que el modelo predictivo ejerció sobre cada uno de los nodos de información no pueden deshacerse. Esto quiere decir que, incluso si la lista de textos estuviese disponible, no resultaría en absoluto claro, ni siquiera para un programador, qué asociaciones produjeron los resultados que arroja un GML. En este punto, trabajos como los de Ray (2023) comienzan a mostrar la importancia de que las empresas garanticen la transparencia y la “explicabilidad”¹³ de los modelos, debido a que la ausencia de trazabilidad de fuentes de los textos generados por el modelo representa un problema a la hora de asignar responsabilidades por lo que aparece allí.

A su vez, el estudio sistemático de Ray (2023) organiza los desafíos que supone el uso de veintiséis modelos diferentes, y dedica un apartado entero a los sesgos observables en los textos que produce este tipo de *software*: entre ellos se encuentran los sesgos presentes en los textos que alimentaron el modelo, los sesgos que supone la selección del material con el que se lo entrenó y la falta de diversidad de perspectivas y orígenes de los textos, y,

12 Para abordar la cuestión de los sesgos seguimos a Ray (2023) que propone una visión amplia de su alcance. Reconstruimos la definición *ad hoc*: un sesgo es, en este caso, un defecto en el diseño de la herramienta o una sobrerrepresentación en los datos que alimentan la herramienta de determinadas perspectivas, temas o grupos sociales, que conduce a producciones que no son precisas o no dan cuenta de la diversidad o complejidad de los problemas que abordan.

13 Traducción directa del inglés *explainability*, referido a la característica de una forma de tecnología de poder ser explicada por humanos en sus procesos internos (Voosen 2017).

vinculado con estos últimos, los sesgos supuestos por las diferenciaciones entre la información contenida en textos en diferentes lenguas. Por otro lado, Ray también señala que no solo debido a su alimentación, sino a su forma de procesar los textos, un modelo de lenguaje ejercerá necesariamente generalizaciones excesivas en casi cualquier asunto en torno al cual genere un texto: debido a su entrenamiento en cantidades masivas de material textual, es menos probable que pueda generar respuestas que den cuenta de divergencias, o que acuse tendencias hasta el momento no observadas en nuevos conjuntos de datos, algo central para cualquier trabajo científico.

Una gran excepción para este funcionamiento es la implementación unificada de los modelos de lenguaje y los motores de búsqueda: los dos mejores ejemplos de este tipo de modelo son Perplexity AI y Gemini. Estos dos casos, impulsados respectivamente por la empresa Perplexity AI y por Google, utilizan los GML para transformar en conversacionales las interacciones con los motores de búsqueda, y para resumir el contenido de los sitios web y ofrecer respuestas rápidas. La ventaja de este método por sobre otros es que permite ofrecer hipervínculos a las fuentes de las que se obtuvieron los datos resumidos; la desventaja, al menos para la credibilidad del *chatbot* y del modelo, es que esta misma exposición de la fuente puede evidenciar que se trata de textos que no deberían ser tomados como fuentes factuales, como blogs personales o sitios web de compra y venta, por solo nombrar algunos casos. Es importante destacar que, de cualquier manera, los problemas de sesgos y de falta de trazabilidad siguen presentes en las respuestas producidas por Perplexity y Gemini, porque si bien son capaces de vincular la fuente, el entrenamiento del modelo de lenguaje en sí es tan poco explicable como los producidos por otros GML.

3.2.2. Alucinaciones

Las alucinaciones de GMLs están definidas en la bibliografía como fabulaciones que producen los GML frente a la falta de información en su entrenamiento (Alkaissi y McFarlane 2023); frecuentemente, su comportamiento es errático y no manifiesta razonamiento (Qi, Zhu y Wu 2023; Vargas-Murillo, Pari-Bedoya y Guevara-Soto 2023). En trabajos anteriores (Scotto 2024) hemos expuesto mediante una conversación extraída de ChatGPT-3 que es capaz de asignar falsamente un texto a un autor e, inmediatamente después, asignar el mismo texto a otro autor; es capaz de fabricar argumentos para esos textos e incluso críticas. Además, modelos como GPT-3, al ser indagados por sus fuentes (que ya hemos analizado que no es posible reponer), responden que no son capaces de indicarlas, pero que sí pueden recomendar textos que abordan el tema: casi sin excepciones, las citas que devuelve ChatGPT-3 de artículos son falsas, sus DOIs no existen, los nombres de los autores corresponden a científicos de otras áreas e incluso las revistas referidas son ficticias. Sobre esto se expresan muchos autores:

Este estudio descubre que el texto generado por el chatbot de esta IA y el material de referencia, incluyendo las citas y otros aspectos de las publicaciones, como los nombres de los autores, y los DOIs, contienen errores. Nadie puede fiarse sin mayores dudas en

los resultados que produce o en las sugerencias que hace. Incluso ChatGPT reconoció estos errores, manifestó lamentarlos y produjo otros resultados (Roumeliotis & Tselikas, 2023) (Imran y Almusharraf 2023)¹⁴.

Las alucinaciones son un problema en el que diferentes empresas de IA han demostrado interés y existe un notable y permanente esfuerzo por su erradicación. Una estrategia posible para su mitigación consiste, según artículos como los de Martino, Ianelli y Truong (2023), en ofrecer como *prompt* al GML la directiva de que solo extraiga información de un texto en particular, que se carga en el mismo *chat*. Si bien la reducción de alucinaciones es muy notoria en estos casos, no desaparecen del todo: incluso cuando se lo condiciona para que solo responda con el contenido de un texto que se le ofrece, GPT-3 y otros GML terminan por utilizar entrenamiento previo para responder preguntas que no encuentran respuesta en el propio texto ingresado. Este fenómeno está detallado en el estudio de Dempere *et al.* 2023:

Muchos autores han estudiado las alucinaciones que genera la IA. De hecho, Cao *et al.* (2017) demostraron que el 30% de los textos generados ante la solicitud de que un chatbot produzca resúmenes de trabajos del estado de la cuestión por medio de redes neuronales presentaban problemas de alucinaciones. De forma similar, Falke *et al.* (2019) estudian los sistemas más recientes y tecnológicamente avanzados de resumen de textos, y encontraron que presentan cerca de un 25% de errores de alucinaciones. Maynez *et al.* (2020) observaron que más del 70% de resúmenes de una sola línea producidos por modelos basados en IA (recursivos, convolucionales y transformacionales) muestran alucinaciones intrínsecas y extrínsecas (Dempere *et al.* 2023, 7)¹⁵.

Resulta especialmente importante señalar que, según Dempere *et al.*, la detección de este tipo de alucinaciones requiere tanto un profundo conocimiento del tema como un ejercicio de lectura lo suficientemente incisivo como para descubrir afirmaciones no sustentadas. En este sentido, no solo esto puede resultar problemático para personas que no pueden hacer uso experto, sino que puede atentar directamente contra la calidad de la labor de cualquier científico que lo utilice, ya sea para evitar leer otros artículos o para generar partes de un artículo.

3.2.3. Plagio

El siguiente aspecto que aparece con más frecuencia en la bibliografía especializada, particularmente la que refiere al uso de modelos de lenguaje en contextos académicos y científicos, es el problema del plagio;

¹⁴ En el original: "this study found that text generated by this AI chatbot and proposed searched material, including references and other publication feathers like authors' names and DOIs, contain errors. No one can mindlessly rely on the results produced and suggestions made. Even ChatGPT accepted those errors, expressed regret for them, and tried to produce similar results (Roumeliotis & Tselikas, 2023)". La traducción nos pertenece.

¹⁵ En el original: "Several authors have studied AI-generated hallucinations. Indeed, Cao *et al.* (2017) find that 30% of the outputs generated by state-of-the-art neural summarization applications suffer from hallucination problems. Similarly, Falke *et al.* (2019) study the most recent technologically advanced summarization systems and find that they produce about 25% of hallucination errors in their summaries. Likewise, Maynez *et al.* (2020) see 70%+of single-sentence summaries show intrinsic/extrinsic hallucinations in AI-based systems (Recurrent, Convolutional, and Transformers)". La traducción es de nuestra autoría.

específicamente, la posibilidad de detectar uso de IA, práctica en sí misma entendida en diversos contextos académicos como plagio (Sullivan, Kelly y McLaughlan 2023; Lo 2023; Niloy *et al.* 2024; Imran y Almusharraf 2023; Dempere *et al.* 2023; entre otros).

Según exponen Mai, Da y Van Hahn (2023) en la serie de pruebas que analizan, ChatGPT-3 generó texto que pasó las pruebas de integridad académica del sistema *Turnitin*, una herramienta de detección de plagio que diferentes autores refieren como la más utilizada a nivel mundial (Chaudhry *et al.* 2023). El artículo de Al Afnan *et al.* demuestra que GPT-3 y otros modelos son capaces de superar detectores de plagio en ciertas tareas, exactamente tres de las seis que presentaron en su investigación. Estas tareas exitosamente realizadas sin detección de plagio eran: producir un reporte en base a un estudio de caso ya ingresado en forma de *prompt* en torno al fenómeno de la sobreinformación, responder una pregunta teórica sobre economía, y responder dos preguntas de interpretación de un texto (de qué se trata y cuál es tu opinión). Los casos en los que *Turnitin* logró detectar plagio fueron las siguientes tareas: producir una respuesta teórica sobre ciencias de la información, componer un mail para resolver un caso de mal servicio de una empresa, y generar un ensayo sobre aspectos favorables de los autos eléctricos. En el último caso, los autores llegan a afirmar que

el índice de coincidencias para todos los ensayos es alto, y no hay referencias a los originales. Los ensayos fueron calificados sin analizar el grado de coincidencias o similitud general. Si la similitud hubiese sido tenida en cuenta, todos los ensayos, excepto el n°29, hubiesen recibido un cero o como mínimo un informe de mala conducta académica, debido a que existían coincidencias en todos los ensayos con respecto a varias fuentes (Al Afnan *et al.* 2023, 64)¹⁶.

En esta línea, trabajos como los de Xames y Shefa (2023) advierten a los científicos que la tendencia de ChatGPT-3 de reproducir texto sin la citación adecuada puede significar un desafío importante para investigadores que usen esta herramienta:

Es esencial que los desarrolladores y los investigadores aborden este asunto para asegurarse de que ChatGPT produce “outputs” adecuados y éticos, que estén a la altura de estándares académicos. (...) En última instancia, es responsabilidad de los investigadores que usan ChatGPT asegurarse de que citaron adecuadamente todas las fuentes y evitaron cualquier plagio no intencional (Xames y Shefa 2023, 3)¹⁷.

La promesa de un *chatbot* que produce respuestas diferentes de usuario a usuario es interesante como estrategia para evitar cualquier forma de

¹⁶ En el original: “the similarity index for all essays is high without any reference to the original. The essays in the grading exercise were graded without deducting any grade on similarity. If the similarity is considered in these essays, all essays, except test 29, would have been given zeros or at least reported for academic misconduct as the similarity is spread all over the essays from different resources”. La traducción es de nuestra autoría.

¹⁷ En el original: “It is essential for the developers and researchers to address this issue to ensure that ChatGPT produces accurate and ethical outputs that meet scholarly standards. This could involve incorporating mechanisms for identifying and citing sources, or training the model to recognize and properly attribute previous work. Ultimately, it is the responsibility of the researchers using ChatGPT to ensure that they properly cite all sources and avoid any unintentional plagiarism”. La traducción es de nuestra autoría.

plagio. Sin embargo, en principio, no pareciera que las reformulaciones que opera GPT-3 impliquen transformaciones lo suficientemente importantes como para que detectores de plagio no identifiquen copias, según lo que muestran Al Afnan *et al.*; especialmente si se lo utiliza con un *fine tuning* de artículos científicos determinados. A la vez, la advertencia de Xames y Shefa representa un llamado de atención con respecto a la posibilidad de que el entrenamiento del modelo realizado con determinados textos pueda reproducir conceptos o nociones ideadas por otros autores, a pesar de que no se utilicen los mismos términos que en el texto original; y si bien en este punto no se trataría de plagio en sentido estricto, es responsabilidad de los investigadores verificar que su trabajo no utiliza, sin citar, ideas de otro. Teniendo en cuenta que rastrear todas las fuentes que utiliza un GML no es enteramente posible a menos que se conozca de antemano el texto original, persiste, a pesar del uso cuidadoso, algún riesgo de que un GML reproduzca en un texto nuevo algo ya publicado por otro autor.

3.2.4. Parroting

En marzo de 2021, un año y ocho meses antes del lanzamiento al público de ChatGPT-3 y de que se produjera la verdadera masificación de los grandes modelos de lenguaje, Bender, Gebru, McMillan-Major y Shmitchel publicaron un trabajo titulado *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* (2021). En este trabajo los autores problematizan el alcance de sostener conversaciones que contravienen los principios básicos de la pragmática, que indican que toda interacción en un lenguaje natural supone que ambas partes intentan expresar significado: explican, así, que un modelo de lenguaje no construye significado, sino que devuelve asociaciones estadísticas que *parecen* tener sentido (pero, aclaran también, que la definición de que algo tiene sentido radica en la mirada de quien lo define). Así, los modelos de lenguaje son meros “loros estocásticos”, que parlotean sin intención, sin sentido, sin lógica algo que solo parece humano. En base a esta descripción apareció luego en otros artículos el término *parroting*: cháchara insustancial que repite palabras asociadas al contexto general que se crea a partir del *input*, pero que no están organizadas para producir lo que llamaríamos, en términos críticos, un texto, ya que no contiene una estructura coherente.

Este es el fenómeno más referido en el trabajo de Al Afnan *et al.* (2023), que analiza los resultados de múltiples pruebas remitidas a ChatGPT-3, las últimas de las cuales eran ensayos. En el análisis cuantitativo y cualitativo del artículo se informa lo siguiente sobre la composición de los ensayos:

Los ensayos generados por ChatGPT sufren de severos problemas de estructura y organización. A pesar de que logró exponer respuestas integrales a las preguntas de algunas áreas de las afirmaciones iniciales, no logró construir afirmaciones atinadas en todos los ensayos generados. La exposición de ideas era o bien meras enunciaciones o bien ideas no adecuadas, ya que no incluían los puntos principales abordados en el ensayo. Además, casi todos los ensayos carecían de sustento en términos de argumentos subsidiarios en párrafos del texto; los párrafos principales tampoco incluían suficientes fuentes para sostener al escritor/generador (ChatGPT). También advertimos que los

párrafos carecían de 'unidad', ya que no todas las oraciones en un párrafo colaboraban con el tópico (Al Afnan et al. 2023, 64)¹⁸.

Las conclusiones del artículo se aproximan a la observación de que un modelo de lenguaje no es, al menos todavía, capaz de producir un texto científico o ensayístico *totalmente* coherente; muchas de sus producciones ofrecen argumentos que no son retomados luego, afirmaciones que resultan insustanciales cuando no incorrectas, y allí donde hay afirmaciones adecuadas, se reducen en general a meras enunciaciones sin desarrollo. Reflexionar sobre el discurso científico de calidad y el posible rol de un modelo de lenguaje basado en inteligencia artificial implicaría necesariamente un nivel de supervisión extremadamente cuidadoso de la progresión argumental presente en el texto científico. Cabe hacerse la pregunta entonces, naturalmente, de cuánto trabajo extra supone para un científico la verdadera supervisión de este tipo de problemas de redacción, al menos en relación con el esfuerzo que supone, directamente, componer el texto.

3.2.5. Composición textual endeble: no determinismo y pérdida de rigor

Por último, existe un aspecto que no ha sido atendido extensamente por la bibliografía dedicada al uso científico y académico de los GML, aunque aparece mencionado muy lateralmente en los trabajos de Al Afnan *et al.* (2023) y en Ray (2023). Se trata de una característica que resulta muy atractiva del uso de modelos de lenguaje, especialmente para estudiantes, y que representa quizás el desafío más importante para docentes que persiguen su detección: un *chatbot* no ofrece dos veces la misma respuesta, o al menos no utiliza exactamente *todas* las mismas palabras. El motivo detrás de esta variación no es únicamente el hecho de que la respuesta depende del *prompt* ingresado por el usuario y por su historial, sino el hecho de que la mayoría de estos GML tienen un carácter no determinista: esto significa que, al momento de predecir la siguiente palabra, el modelo no siempre utiliza el primer resultado, el más probable, sino que frecuentemente alterna con otros resultados posibles de forma aleatoria (Ouyang, Zhang, Harman y Wang 2025). Algunos trabajos (especialmente Song, Wang, Li y Lin 2024) comienzan a abordar el problema de que esta característica de los modelos no se esté teniendo en cuenta al medir su impacto en áreas determinadas de implementación, cuando su evaluación está sostenida en resultados extraídos de un *prompt* ingresado una sola vez. Esto apunta al problema de que en general puede sobreestimarse el rendimiento de un modelo si únicamente se le pregunta una vez por algún tema, algo que no ocurre en la realidad.

¹⁸ En el original: "essays generated by ChatGPT also suffer from relatively major structuring and organization issues. Even though ChatGPT managed to provide comprehensive answers to the question on parts of thesis statements, it failed to construct accurate thesis statements in all generated essays. The thesis statements in the generated essays were either announcements or inadequate as they did not include all major points discussed in the essay. In addition, almost all essays lacked support in terms of minor points in the body paragraphs of the essays. The body paragraphs either did not include enough support to back the stand of the writer/generator (ChatGPT). It was also noticed that the paragraphs did not have 'unity' as not all sentences in a paragraph supported the topic sentence." La traducción es de nuestra autoría.

Para explicar más detalladamente la gravedad que supone para el uso académico y científico el carácter no determinista de los GML utilizaremos un ejemplo de elaboración propia, conducido en ChatGPT-4, la versión más avanzada de las disponibles de forma gratuita. Dos usuarios diferentes, en dos dispositivos electrónicos diferentes, separados por muy poco tiempo, ingresaron el mismo *prompt*: “Producí una respuesta admisible en nivel universitario para la siguiente pregunta: ¿qué son las modalidades según Maingueneau?”. La primera respuesta presentó serios problemas de adecuación, pero la clasificación contenía algunos aciertos: mencionaba “modalidades epistémicas, deónticas y apreciativas”, que no se corresponden con la categorización del autor mencionado, pero algunos de los ejemplos eran acertados, aunque estuvieran mal ubicados. En el segundo intento, el chat respondió con definiciones de términos completamente alejados del campo de la lingüística y las redujo a dos: “modalidades epistemológicas y deontológicas”. Es posible advertir en este ejemplo el problema con el grado de no determinismo de este GML en el contexto científico: la no repetición de términos puede conducir a crasos errores conceptuales debido a similitudes (por ejemplo, entre “deóntico” y “deontológico”) solo aparentes. Un GML no es capaz de identificar qué términos pueden reemplazarse por otros y cuáles no, habilidad básica cuya ausencia resultaría extraordinariamente problemática en un trabajo científico. Resulta importante aclarar que otros modelos como Perplexity, que utilizan los GML para procesar y resumir texto que obtienen con un motor de búsqueda, poseen una marcada reducción del grado de no determinismo de su modelo en relación con otros como los de Meta o OpenAI. Esto mejora estos problemas de desempeño de forma drástica: en el mismo ensayo, Perplexity ambas veces respondió acertadamente las tres categorías de modalidades, las definió correctamente y citó como referencia un apunte de la Facultad de Periodismo de la Universidad Nacional de La Plata en los dos casos.

4. El futuro llegó hace rato: algunas propuestas

Frente a esta avanzada tecnológica, algunos autores no consideran conveniente ni efectivo abordar estrategias prohibicionistas, dado que en general no parecen funcionar a corto, mediano o largo plazo (García-Peñalvo, Llorens-Largo y Vidal 2023; Grassini 2023). Sin embargo, por lo que hemos expuesto anteriormente resulta notorio que los GML no se tratan de herramientas que no presenten problemas tanto en su uso como en términos de su resultado; y, en el caso del uso académico, podrían resultar dañinas para los estudiantes en tanto pueden entorpecer el proceso de desarrollo de su alfabetización académica.

En principio, en términos del uso académico, especialmente en docencia, entender las generalidades del funcionamiento de los GML es un buen primer paso para docentes que deseen responder de alguna forma a la nueva presencia de las IA en las aulas. Un segundo paso puede radicar en mostrar, como parte de las tareas docentes, a qué punto estas tecnologías no son capaces de realizar las sofisticadas tareas que se exigen a estudiantes de educación superior, y su uso puede resultar riesgoso en tanto ellos

no podrán confiar en que les resulte útil lo que produce un GML, al menos hasta que no estén alfabetizados académicamente y dominen relativamente el campo sobre el que van a hacer una consulta. Debido a que los estudiantes no están en condiciones de realizar un uso experto, al menos en los primeros años, una aproximación para el asunto podría implicar demostrar a estudiantes los riesgos del uso inexperto y, sobre todo, no supervisado, en el cual no medien operaciones básicas como la verificación, el contraste con fuentes confiables, la corrección y la edición.

Con respecto al uso experto, existen muchos motivos prácticos para utilizar estos programas, pero una reflexión profunda sobre su funcionamiento y una actitud verdaderamente seria que analice los riesgos de utilizarlo pueden llegar a disuadir a colegas; sobre todo, porque el trabajo de persecución de posibles fuentes, verificación de datos, análisis crítico del desarrollo argumental y corrección general del texto producido por el chat puede terminar excediendo el esfuerzo de escribir un artículo. Hemos analizado que los GML *no pueden* escribir *por sí solos* artículos científicos sin presentar los problemas ya mencionados: no parecería posible ofrecer a un GML datos propios de una investigación y solicitar, por ejemplo, que produzca una introducción a un artículo. Los resultados pueden o bien no estar de acuerdo con los datos ingresados como *input*, o bien estar sesgados por las asociaciones establecidas por el entrenamiento del GML de formas que resultan desconocidas, o bien carecer de una estructura argumentativa adecuada. Quizás sea posible hacer uso de ellos como una asistencia ante el *horror vacui* de la hoja en blanco o una corrección formal menor, pero no parecieran poder reemplazar la producción de textos, o al menos no frente a una expectativa de estándares altos como los exigidos en la comunidad científica.

Bibliografía

- » AlAfnan, Mohammad A., Samira Dishari, Marina Jovic, y Koba Lomidze. 2023. “ChatGPT as an Educational Tool: Opportunities, Challenges, and Recommendations for Communication, Business Writing, and Composition Courses”. *Journal of Artificial Intelligence and Technology* 3: 60-68. Fecha de consulta: 15 de febrero de 2025. ChatGPT as an Educational Tool: Opportunities, Challenges, and Recommendations for Communication, Business Writing, and Composition Courses | Journal of Artificial Intelligence and Technology.
- » Alkaisi Hussam y Samy I. McFarlane. 2023. “Artificial Hallucinations in ChatGPT: Implications in Scientific Writing”. *Cureus*. 15: e35179. Fecha de consulta: 15 de febrero de 2025. <http://doi.org/10.7759/cureus.35179>.
- » American Psychological Association (APA). 2018. *APA Dictionary of Psychology*. Washington: American Psychological Association.
- » Archuby, Gustavo. 2025. “Repositorios institucionales y plataformas de inteligencia artificial ¿acceso abierto para todo?”. *Acceso Abierto*. Última modificación: 11 de febrero de 2025. Repositorios institucionales y plataformas de inteligencia artificial ¿acceso abierto para todo?.
- » Barman, Kristian, Nathan Wood, y Pawel Pawlowski. 2024. “Beyond Transparency and Explainability : On the Need for Adequate and Contextualized User Guidelines for LLM Use.” *ETHICS AND INFORMATION TECHNOLOGY* 26 (3). <https://doi.org/10.1007/s10676-024-09778-2>.
- » Barroso, Julio y Julio Cabero. 2011. *La investigación educativa en TIC*. Madrid: Síntesis.
- » Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?”. Trabajo presentado en *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAcT '21)*. Association for Computing Machinery, Nueva York, Nueva York, Estados Unidos, 1 de marzo de 2021. On the Dangers of Stochastic Parrots | Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency.
- » Bonneau, Jonathan, Laurence Grondin-Robillard, Marc Ménard, André Mondoux. 2022. “Fighting the ‘System’: A Pilot Project on the Opacity of Algorithms in Political Communication”. En *New Perspectives in Critical Data Studies. Transforming Communications – Studies in Cross-Media Research*, editado por Andreas Hepp, Julianne Jarke y Leif Kramp, 97-120. Cham: Palgrave Macmillan. http://doi.org/10.1007/978-3-030-96180-0_5.
- » Chan, Alan, Rebecca Salganik, Alva Markelius, Chris Pang, Nitarshan Rajkumar, Dmitrii Krashennnikov, Lauro Langosco, Zhonghao He, Yawen Duan, Micah Carroll, Michelle Lin, Alex Mayhew, Katherine Collins, Maryam Molamohammadi, John Burden, Wanru Zhao, Shalaleh Rismani, Konstantinos Voudouris, Umang Bhatt, Adrian Weller, David Krueger, y Tegan Maharaj. 2023. Harms from Increasingly Agentic Algorithmic Systems. En: *2023 ACM Conference on Fairness, Accountability, and Transparency (FAcT '23)*, June 12--15, 2023, Chicago, Illinois, Estados Unidos. Fecha de consulta: 28 de abril de 2025. Harms from Increasingly Agentic Algorithmic Systems | Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency.
- » Chaudhry, Ifat S., Sayed A. M. Sarwary, Ghaleb A., el Refae, Habeb Chabchoub. 2023. “Time to revisit existing Student’s performance evaluation approach in higher education sector in a new era of ChatGPT — a case study”. *Cogent Education* 10 (1): 1-30. Time

to Revisit Existing Student's Performance Evaluation Approach in Higher Education Sector in a New Era of ChatGPT — A Case Study.

- » Christou, Prokopis A. 2024. "A critical inquiry into the personal and societal perils of Artificial Intelligence". *AI Ethics* 2024. Fecha de consulta: 15 de febrero de 2025. <https://doi.org/10.1007/s43681-024-00556-w>.
- » Conroy, Gemma. 2023. "Scientific sleuths spot dishonest ChatGPT use in papers". *Nature. News*. Fecha de consulta: 15 de febrero de 2025. <http://doi.org/10.1038/d41586-023-03144-w>.
- » Darwin, Diyenti R., Nur Mukminatien, Nunung Suryati, Ekaning D. Laksmi, and Marzuki. 2023. "Critical Thinking in the AI Era: An Exploration of EFL Students' Perceptions, Benefits, and Limitations." *Cogent Education* 11: 1-18. Full article: Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations.
- » Dempere Juan, Kenedy Modugu, Hesham Allam y Lakshmana K. Ramasamy. 2023. "The impact of ChatGPT on higher education". *Frontiers in Education* 8: 1-13. Fecha de consulta: 15 de febrero de 2025. The impact of ChatGPT on higher education.
- » Floridi, Luciano. 2023. *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*. Oxford: Oxford University Press.
- » García Sánchez, Omar V. 2023. "Uso y percepción de ChatGPT en la educación superior". *Revista de Investigación en Tecnologías de la Información* 11: 98-107. Fecha de consulta: 15 de febrero de 2025. Uso y percepción de ChatGPT en la educación superior | Revista de Investigación en Tecnologías de la Información.
- » García-Peñalvo, Francisco J., Faraón Llorens-Largo, y Javier Vidal. (2024). "La nueva realidad de la educación ante los avances de la inteligencia artificial generativa". *RIED-Revista Iberoamericana de Educación a Distancia* 27: 9-39. Fecha de consulta: 15 de febrero de 2025. La nueva realidad de la educación ante los avances de la inteligencia artificial generativa | RIED-Revista Iberoamericana de Educación a Distancia.
- » Grassini, Simone. 2023. "Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings". *Education Science* 13: 692-711. Fecha de consulta: 15 de febrero de 2025. Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings.
- » Imran, Muhammad y Norah Almusharraf. 2023. "Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature". *Contemporary Educational Technology* 15 (4): 1-14. Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature.
- » Kostka, Ilka y Rachel Toncelli. 2023. "Exploring Applications of ChatGPT to English Language Teaching: Opportunities, Challenges, and Recommendations". *The Electronic Journal for English as a Second Language* 27: 1-19. Fecha de consulta: 15 de febrero de 2025. <https://doi.org/10.55593/ej.27107int>.
- » Lillis, Theresa. 2003. "Student Writing as 'Academic Literacies': Drawing on Bakhtin to Move from Critique to Design. Language and Education". *Language and Education* 17: 192-207. Fecha de consulta: 15 de febrero de 2025.
- » Lillis, Theresa. 2021. "Academic Literacies: Intereses Locales, Preocupaciones Globales? Academic Literacies: Local Interests, Global Concerns". En *Multilingual Contributions to Writing Research Toward an Equal Academic Exchange*, editado por Natalia Ávila Reyes, 35-59. Colorado: University Press of Colorado.
- » Liu, Sijia, Yuanshun Yao, Jinghan Jia, Stephen Casper, Nathalie Baracaldo, Peter Hase, Yuguang Yao, Chris Yuhao Liu, Xiaojun Xu, Hang Li, Kush R. Varshney, Mohit Bansal, Sanmi Koyejo, Yang Liu. 2024. Rethinking Machine Unlearning for Large Language Mo-

dels. *Nature Machine Intelligence*. En prensa. <https://doi.org/10.48550/arXiv.2402.08787>.

- » Lo, Chung K. 2023. "What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature". *Education Sciences* 13 (410): 1-15. Fecha de consulta: 15 de febrero de 2025. What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature.
- » Macaranas, Anna, Alissa N. Antle, Bernhard E. Riecke, 2015. "What is Intuitive Interaction? Balancing Users' Performance and Satisfaction with Natural User Interfaces". *Interacting with Computers* 2: 357-370. Fecha de consulta: 15 de febrero de 2025. What is Intuitive Interaction? Balancing Users' Performance and Satisfaction with Natural User Interfaces | Interacting with Computers | Oxford Academic.
- » Mai Duong T. T., Can Van Da y Nguyen Van Hanh. 2024. "The use of ChatGPT in teaching and learning: a systematic review through SWOT analysis approach". *Frontiers in Education* 9: 1-17. Fecha de consulta: 15 de febrero de 2025. The use of ChatGPT in teaching and learning: a systematic review through SWOT analysis approach.
- » Martino, Ariana, Michael Iannelli, Coleen Truong. 2023. "Knowledge Injection to Counter Large Language Model (LLM) Hallucination". En *The Semantic Web: ESWC 2023 Satellite Events. Lecture Notes in Computer Science*, editado por Catia Pesquita, Hala Skaf-Molli, Vasilis Efthymiou, Sabrina Kirrane, Axel Ngonga, Diego Collarana, Renato Cerqueira, Mehwish Alam, Cassia Trojahn, Sven Hertling, vol 13998, 182-185. Cham: Springer. Knowledge Injection to Counter Large Language Model (LLM) Hallucination | Springer-Link.
- » Montenegro-Rueda, Marta, José Fernández-Cerero, José María Fernández-Batanero, Eloy López-Meneses. 2023. "Impact of the Implementation of ChatGPT in Education: A Systematic Review". *Computers* 12: 1-13. Fecha de consulta: 15 de febrero de 2025. Impact of the Implementation of ChatGPT in Education: A Systematic Review.
- » Niloy, Ahnaf Chowdhury, Muhammad Ashraful Bari, Jakia Sultana, Rup Chowdhury, Fareha Meem Raisa, Afsana Islam, Saadman Mahmud, et al. 2024. "Why Do Students Use ChatGPT? Answering through a Triangulation Approach". *Computers and Education: Artificial Intelligence* 6: 200-208. Why do students use ChatGPT? Answering through a triangulation approach - ScienceDirect.
- » OpenAI. 2019. Better language models and their implications. Última modificación: 14 de febrero de 2019. Better language models and their implications | OpenAI.
- » Ouyang, Shuyin, Jie M. Zhang, Mark Harman, and Meng Wang. 2025. "An Empirical Study of the Non-Determinism of ChatGPT in Code Generation". *ACM Transactions on Software Engineering and Methodology* 34: 1-28. Fecha de consulta: 15 de febrero de 2025. An Empirical Study of the Non-Determinism of ChatGPT in Code Generation | ACM Transactions on Software Engineering and Methodology.
- » Pasquinelli, Matteo y Vladan Joler. 2021. The Noosphere manifested: AI as instrument of knowledge extractivism. *AI & Society* 36, 1263-1280. Fecha de consulta: 28 de abril de 2025. The Noosphere manifested: AI as instrument of knowledge extractivism | AI & SOCIETY.
- » Qi, Xiang, Zheng Zhu, Bei Wu, X. 2023. "The promise and peril of ChatGPT in geriatric nursing education: What we know and do not know". *Aging and Health Research*, 3: 100-136. Fecha de consulta: 15 de febrero de 2025. <https://doi.org/10.1016/j.ahr.2023.100136>.
- » Ray, Partha P. 2023. "ChatGPT: A Comprehensive Review on Background, Applications, Key Challenges, Bias, Ethics, Limitations and Future Scope". *Internet of Things and Cyber-Physical Systems* 3: 121-154. Fecha de consulta: 15 de febrero de 2025. ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope - ScienceDirect.

- » Scotto, Victoria. 2024. El ChatGPT y el fin de la historia: una elaboración teórica sobre sus peligros y promesas leídos desde la filología. *Recial*, 15 (25), 118-146. Fecha de consulta: 15 de febrero de 2025. <https://doi.org/10.53971/2718.658x.v15.n25.45626>.
- » Song, Yifan, Guoyin Wang, Li y Lin. 2024. "The Good, The Bad, and The Greedy: Evaluation of LLMs Should Not Ignore Non-Determinism". *ArXiv*. Última modificación: 15 de julio de 2024. [arXiv:2407.10457v1](https://arxiv.org/abs/2407.10457v1) [cs.CL] 15 Jul 2024.
- » Stokel-Walker, Chris. "AI bot ChatGPT writes smart essays - should professors worry?". *Nature*. Última modificación: 9 de diciembre de 2022. AI bot ChatGPT writes smart essays — should professors worry?.
- » Sullivan, Miriam, Andrew Kelly, y Paul McLaughlan. 2023. "ChatGPT in Higher Education: Considerations for Academic Integrity and Student Learning". *Journal of Applied Learning & Teaching* 6: 1-10. Fecha de consulta: 15 de febrero de 2025. ChatGPT in higher education: Considerations for academic integrity and student learning | Journal of Applied Learning and Teaching.
- » Thorp, Holden H. 2023. "ChatGPT Is Fun, but Not an Author". *Science* 379: 6630 (27 de enero de 2023): 313-313. Fecha de consulta: 15 de febrero de 2025. <https://doi.org/10.1126/science.adg7879>.
- » Tinmaz, Hasan, Mina Fanea-Ivanovici, Hasnan Baber. 2022. "A snapshot of digital literacy". *Library Hi Tech News* 40.1: 20-23. <https://doi.org/10.1108/LHTN-12-2021-0095>.
- » Transformer, Chat G. P. T. y Alex Zhavoronkov. 2022. "Rapamycin in the context of Pascal's Wager: Generative pre-trained transformer perspective". *Oncoscience* 9 (82): 82-84. Fecha de consulta: 15 de febrero de 2025. Rapamycin in the context of Pascal's Wager: generative pre-trained transformer perspective.
- » Vargas-Murillo, Alfonso R., Ilda N. M. A. Pari-Bedoya, y Francisco De Jesús Guevara-Soto. 2023. "Challenges and Opportunities of AI-Assisted Learning: A Systematic Literature Review on the Impact of ChatGPT Usage in Higher Education". *International Journal of Learning, Teaching and Educational Research* 22: 122-135. Fecha de consulta: 15 de febrero de 2025. Challenges and Opportunities of AI-Assisted Learning: A Systematic Literature Review on the Impact of ChatGPT Usage in Higher Education | Vargas-Murillo.
- » Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. "Attention Is All You Need". *Advances in Neural Information Processing Systems*. Fecha de consulta: 15 de febrero de 2025. [1706.03762] Attention Is All You Need.
- » Voosen, Paul. 2017. "How AI detectives are cracking open the black box of deep learning". *Science. News*. 6 de julio de 2017. Fecha de consulta: 15 de febrero de 2025. How AI detectives are cracking open the black box of deep learning | Science.
- » Xames, Doulotuzzaman y Jannatul Shefa. 2023. "ChatGPT for research and publication: Opportunities and challenges". *Journal of Applied Learning & Teaching* 6: 1-6. Fecha de consulta: 15 de febrero de 2025. <https://doi.org/10.37074/jalt.2023.6.1.20>.
- » Zhang, Peng y Gemma Tur. 2024). "A systematic review of ChatGPT use in K-12 education". *European Journal of Education* 59: e12599. Fecha de consulta: 15 de febrero de 2025. <https://doi.org/10.1111/ejed.12599>.
- » Zhuk, Alesia. 2023. "Artificial Intelligence Impact on the Environment: Hidden Ecological Costs and Ethical-Legal Issues". *Journal of Digital Technologies and Law* 4: 932-954. Fecha de consulta: 15 de febrero de 2025. Artificial Intelligence Impact on the Environment: Hidden Ecological Costs and Ethical-Legal Issues | Zhuk.