



## Evidencia empírica, oportunidades y riesgos de la IA

Silvina Damiani

La inteligencia artificial (IA) ha demostrado ser una herramienta valiosa en muchos ámbitos: hoy día la automatización de procesos empresariales<sup>1</sup>, la conducción autónoma de vehículos<sup>2</sup>, la generación de texto, la traducción automática y las evaluaciones en los ámbitos educativo, crediticio, médico, laboral, judicial, penal son algunos contextos en los que se la emplea.

Pero, además, su uso se sigue extendiendo cada vez más, porque cada vez estamos más rodeados de inteligencia artificial. Como señala el filósofo y economista Diego Hidalgo (2021) en su libro “Anestesiados. La humanidad bajo el imperio de la tecnología”:

Si recordamos nuestra relación con la tecnología digital hace apenas unos años, el uso estaba mucho más circunscrito en el espacio y el tiempo, y el límite entre conexión y desconexión era mucho más definido y consciente. O bien estábamos ante el ordenador o no lo estábamos. En el primer caso, nos encontrábamos en la habitación donde se hallaba el ordenador, un despacho o un rincón de la sala de casa. En el segundo caso, estábamos desconectados, en la calle, en la mesa, en el transporte público o en el coche. Cualquier conexión empezaba al pulsar el botón *on*. (p.23).

En la actualidad, en cambio, como podemos dar cuenta, la tecnología y la IA forman parte de nuestras vidas y de nuestros hábitos cotidianos.

Mark Weiser, un prominente científico informático de origen estadounidense, vaticinó tempranamente esta situación en el año 1988 cuando concibió la *tecnología ubicua* en referencia a entornos rodeados de computadoras y redes de comunicaciones inalámbricas al alcance de todos. Sin embargo, no fue hasta septiembre de 1991 que su concepción se hizo famosa debido a la circulación del artículo de su autoría “The Computer for the Twenty-First Century” publicado por la prestigiosa revista *Scientific American*, cuyo

---

<sup>1</sup> Un estudio publicado por Goldman Sachs en abril de 2023 predice que los sistemas de IA aumentarán significativamente los puestos de trabajo en todo el mundo, con el potencial de sustituir 300 millones de puestos a tiempo completo. Otra encuesta realizada por McKinsey en 2022 arrojó datos interesantes sobre el uso actual de la IA en las empresas: Desarrollo de productos y/o servicios: 55%/Marketing y ventas: 55%/ Operaciones de servicio (incluyendo atención al cliente y back office): 54% y Modelización y análisis de riesgos: 15%

<sup>2</sup> Se prevé que el mercado de los coches autónomos crezca a una tasa de crecimiento anual compuesta (CAGR) del 13,3%, de 20,3 millones de unidades en 2021 a 62,4 millones de unidades en 2030.



primer párrafo ya auguraba los recursos tecnológicos propios de nuestro presente: “The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it.” (Weiser, 1991, p.94).

En este sentido, la inteligencia artificial es en la actualidad una herramienta valiosa e imprescindible, dos fuertes razones que nos lleva a predecir que estamos ante un fenómeno que llegó para quedarse de manera definitiva -con y entre nosotros- que da la certeza de que algo cambió súbita y abruptamente para siempre. Por esta razón, es fundamental reflexionar acerca de los desafíos y riesgos que nos plantea.

En nuestra ponencia abordaremos varias cuestiones. No obstante, a fin de delimitar la amplitud del tema, acotaremos el análisis de la IA al modelo generativo de lenguaje, GPT (y al Chat GPT3). El criterio de nuestra elección se basa en dos razones fundamentales: i) porque GPT es uno de los modelos más avanzados en el campo del procesamiento del lenguaje natural y ii) porque GPT es una tecnología que está actualmente disponible y es de libre acceso para cualquier persona que desee usarla.

ChatGPT se lanzó el 30 de noviembre de 2022 y en tan solo cinco días superó el millón de usuarios. A los dos meses, enero de 2023, OpenAI publicaba en su cuenta *LinkedIn* que el chat había superado ya los 100 millones de usuarios<sup>3</sup>.

¿Qué es Chat GPT3? Si quisiéramos definirlo en pocas palabras diríamos que se trata de una inteligencia artificial que predice qué palabras van a ocurrir dada una cadena textual. Si nos interesara una respuesta más detallada, el Chat nos los dice:

Chat GPT-3 (Generative Pre-trained Transformer 3) es un modelo de lenguaje desarrollado por OpenAI. Utiliza una arquitectura basada en transformadores, una técnica de aprendizaje automático que permite procesar y generar texto. GPT-3 es uno de los modelos de lenguaje más grandes y avanzados creados hasta la fecha. GPT-3 se entrena utilizando grandes cantidades de datos de texto en múltiples dominios, como libros, artículos de noticias, páginas web y otros recursos en línea. El modelo aprende a predecir la siguiente palabra o frase en una oración basándose en el contexto proporcionado. Esto le permite generar texto coherente y relevante en respuesta a una pregunta o una instrucción. El

---

<sup>3</sup> 9. Chat GPT AI; 2023. Disponible en: [https://www.linkedin.com/posts/gpt-ai\\_chatgpt-activity7038523838967549952-8b0P?utm\\_source=share&utm\\_medium=member\\_desktop](https://www.linkedin.com/posts/gpt-ai_chatgpt-activity7038523838967549952-8b0P?utm_source=share&utm_medium=member_desktop). También se puede testear en Statista



objetivo principal de GPT-3 es responder de manera inteligente y natural al texto que se le proporciona. Puede realizar tareas como responder preguntas, escribir ensayos, generar código de programación, traducir texto, chatear con los usuarios y mucho más. GPT-3 tiene una amplia gama de aplicaciones y ha sido utilizado en diversos campos, desde asistentes virtuales hasta herramientas de escritura y tutoriales interactivos. Sin embargo, es importante tener en cuenta que GPT-3 es un modelo de lenguaje basado en patrones y no tiene una comprensión real del significado o el contexto más allá de lo que se le ha proporcionado en sus datos de entrenamiento. Aunque puede generar respuestas coherentes y contextualmente relevantes en muchos casos, también puede producir resultados incorrectos o incoherentes en situaciones más complejas o ambiguas<sup>4</sup>.

La primera característica verdaderamente curiosa de este modelo de lenguaje es que no contiene ninguna tecnología nueva importante. Básicamente es una versión ampliada de otros modelos de lenguaje que usan aprendizaje profundo. En todos los casos existentes, estos modelos del lenguaje son enormes redes neuronales artificiales entrenadas en -y con- texto para predecir cuál será probablemente la siguiente palabra en una secuencia. GPT es simplemente el modelo más grande: cien veces más grande (98 capas y 175 mil millones de parámetros) que los anteriores y está entrenado con mucha más cantidad de datos, en particular con *Common Crawl*, una base de datos que contiene gran parte de Internet en conjunción con una enorme biblioteca de libros y Wikipedia. Es decir, el modelo se entrenó con cientos de miles de millones de palabras de Internet y de libros (Thompson, 2022) (Recalde Varela, Bolagay Egas y Yáñez Velásquez, 2023)

Ahora bien, ¿por qué hace tanto ruido<sup>5</sup>, si no tiene ninguna tecnología nueva importante respecto de modelos de lenguaje anteriores?

En primer lugar, porque interactuar con chat GPT3 es sencillamente asombroso debido a la naturalidad y familiaridad con que se da dicha interacción. Verdaderamente, el intercambio lingüístico fluye con tanta facilidad que en cuestión de pocos minutos se

---

<sup>4</sup> Pregunta realizada al ChatGPT el día 14 de julio 2023.

<sup>5</sup> Esta pregunta tiene sentido al momento de ser formulada (abril 2023) y contribuyó en la elección del tema de la ponencia. Es difícil predecir si continuará así con el transcurso del tiempo.



pierde la noción de estar chateando con una máquina. La incompreensión semántica del Chatbot pasa completamente inadvertida mientras se está interactuando.

En este punto se nos presenta la primera cuestión filosófica, que surge en torno a la relación *inteligencia/lenguaje*, un tópico de la filosofía. Pues, si una máquina es capaz de responder preguntas mejor que cualquier humano promedio: ¿cuál es entonces la relación entre la inteligencia y el lenguaje humano? ¿Son necesarias (total o parcialmente) nuestras capacidades humanas (comprender y captar significados) para entablar una comunicación exitosa?

Sin ese propósito, pero como consecuencia de su desarrollo, es de esperar que el modelo GPT -y otros modelos de lenguaje- nos empuje un paso adelante en dirección a una comprensión más precisa de este asunto.

En segundo lugar, otra razón por la que creemos que GPT provoca ruido es porque se trata de un modelo de lenguaje que genera muy buenos resultados en una variedad de dominios, pese a que no está entrenado en ningún dominio específico en particular.

En este punto haremos un breve paréntesis en torno a la inteligencia artificial, puesto que es necesario señalar dos áreas escindidas: por un lado, una rama de la IA basada en normas, reglas, sistemas simbólicos o sistemas expertos en la que los programadores intentan enseñar a los ordenadores a pensar codificando una serie de reglas lógicas de la clase *si X, entonces Y*. Este modo de programación como un razonamiento, paso por paso, basado en reglas fue -y sigue resultando- muy adecuado para universos reglados en los que se conoce muy bien todas las reglas. Un ejemplo paradigmático de esta rama de IA es *Deep Blue*, la supercomputadora desarrollada a mediados de los 90 por IBM para jugar al ajedrez, cuya potencia en la primera versión le permitía calcular 100 millones de posiciones por segundo. En 1997, *Deeper Blue* (un año después de la primera versión) doblaba en potencia a su predecesora con 200 millones de operaciones por segundo y obviamente eso le permitió ganarle el título al campeón humano de ajedrez, Gary Kasparov. (Lee, Wang, Yen, Wei, Wu, Chou, y Yan, 2016). El juego de ajedrez no se gana con suerte, sino con cálculo de operaciones y los superordenadores provistos de esta clase de IA han incrementado exponencialmente su capacidad de cálculo, de modo que puede calcular una magnitud enorme de datos en una fracción muy corta de tiempo. Es decir, esta clase de IA es capaz de hacer las mismas tareas que un ser humano, pero en una magnitud mayor y en mucho menos tiempo (Bostrom, 2016). Por otro lado, la otra rama



de la inteligencia artificial, *las redes neuronales*, que en vez de intentar imitar el mecanismo del razonamiento o pensamiento humano con reglas paso por paso, imita la forma del cerebro. En este sentido, las redes neuronales aprenden por sí mismas, deduciendo sus propios patrones a partir del *big data* introducido en ellas imitando la arquitectura subyacente del cerebro, construyendo capas de neuronas artificiales que son capaces de recibir y de transmitir información en una estructura similar a nuestras redes de neuronas biológicas<sup>6</sup>.

La importancia de esta rama de la IA, esto es, de las redes neuronales artificiales, es que además de ser capaces de aprender a partir de los datos, son capaces de aprender en otro sentido: pueden corregir relaciones y procesos concretos (que se llevan a cabo en sus nodos internos) para terminar en una respuesta relativamente eficiente. Es decir, las redes neuronales *no solo aprenden de los datos, sino que pueden cambiar lo que ya saben*, y eso lo pueden hacer de manera supervisada por seres humanos o automáticamente. Por eso los modelos de lenguaje grandes, como GPT, pueden capturar más variaciones de los datos en los que están entrenados y pueden aplicar esto a tareas para las que no han sido entrenados específicamente. En lugar de ajustarse a un problema determinado, el modelo recibe una instrucción y algunos ejemplos de la tarea y se espera que identifique qué hacer basándose solo en esto. En este sentido, GPT es diferente a otros sistemas de procesamiento de lenguaje natural (NLP), porque realiza tareas de lenguaje completamente nuevas basadas en unas pocas instrucciones y en ejemplos simples. Y lo que resulta interesante es que lo hace extraordinariamente bien, dado que genera muy buenos resultados con muy poco contexto. Con el fin de ilustrar esto, transcribiré parte de un intercambio mientras preparaba esta ponencia :

**S** Quisiera jugar al Lenguasil con vos. Es un nuevo lenguaje, semejante al castellano, salvo que cada palabra termina en sil.

**ChatGPT** ¡Por supuesto! Estoy emocionado de jugar contigo al Lenguasil. Entiendo que es similar al castellano, pero con la peculiaridad de que cada palabra termina en "sil". ¿Listo para comenzar? Adelante, puedes hacerme

---

<sup>6</sup> Esta rama de la IA se ha fortalecido en los últimos años fundamentalmente por tres factores: i) el desarrollo del hardware, con la disponibilidad de GPU (Graphics Processing Unit), ii) la masividad de los datos disponibles y el crecimiento exponencial de los datos para alimentar los modelos y, por último, iii) los algoritmos, que debido a los avances en el aprendizaje en redes de múltiples capas no requiere de la intervención humana para la selección de las variables.



preguntas o frases propuestas y estará encantado de responder en Lenguasil. ¡Vamos a divertirnos!

**S** Empezá vos con cualquier pregunta y luego yo respondo.

**ChatGPT** ¡Excelente! Comenzamos con la primera pregunta en Lenguasil:

¿Cuál es tu nombre en lenguasil?

**S** Pero, para empezar a jugar tenés que preguntar en lenguasil.

**ChatGPT** Mis disculpas por el error. ¡Claro que sí! Comenzamos correctamente:

¿Cuál es tu nombre en lenguasil?

**S** ¿No sabes cómo preguntar en Lenguasil?

**ChatGPT** Mis disculpas nuevamente. Tienes toda la razón, cometí un error en la pregunta anterior. Permíteme corregirlo y hacer una pregunta en Lenguasil:

¿Cuál es tu nombre en Lenguasil?

**S** Misil nombresil essil Silsil. ¿Cuál sil essil tusil nombresil?

**ChatGPT** Encantadosil de conocertesil, Silsil.

Milsil nombresil essil ChatGPTSil. Estoyil aquísil para ayudartesil y divertirnosilsil jugando a Lenguasilsil. ¿Qué talilsil si continuamosil con el juegilsil? Puedesil hacermesil otra preguntasil o proponermeilsil una frasisil en Lenguasilsil. ¡Estoyilsil listosil para seguirsil jugando!

**S** Graciassil

**ChatGPT** De nadasil, Silsil. ¡Graciasilsil a tisil por participarilsil en el juegilsil! Siempreilsil estoyilsil aquíilsil para ayudartesil o para jugarilsil cuando quierasil. ¡Que tengasil un díailsil maravillosilsil! :)

Es poco probable, o mejor dicho, es del todo improbable que GPT haya podido encontrar ese “lenguaje (Lenguasil)” en algún texto de algún libro, Internet o Wikipedia antes de



que yo lo propusiera. Sin embargo, es evidente que le bastó la expresión “*Misil nombresil essil Silsil. ¿Cuálsil essil tusil nombresil?*” para descifrar el patrón e identificar la regla de la construcción de la fórmula .

A la luz de esta evidencia surgen otras preguntas: ¿Cómo sabemos si GPT-3 está generalizando una nueva tarea a partir de estos pocos datos o si simplemente está combinando cosas que ya ha visto antes? Por otra parte, ¿habría una diferencia significativa entre ambos comportamientos?<sup>7</sup>

Otro tema de suma importancia en relación con GPT (y con la inteligencia artificial de última generación) es que los sistemas de IA anteriores se han desempeñado muy bien en dominios especializados, como es el caso del Ajedrez, Go, por ejemplo, pero no en la inteligencia general entre distintos dominios. Eso hasta hace muy poco parecía estar muy lejos. Y sin embargo GPT muestra habilidades en muchos dominios y es capaz de realizar una amplia gama de tareas razonablemente bien, desde producir ficción, poesía y comunicados de prensa hasta código de funcionamiento, y desde música, chistes, y manuales técnicos, hasta “artículos de noticias que los evaluadores humanos tienen dificultad para distinguir de los artículos escritos por humanos”<sup>8</sup>. (Brown, Mann, Ryder, Subbiah, Kaplan, Dhariwal, y Amodei, 2020).

Sin embargo, dado que -como ya dijimos- el modelo se entrena automáticamente y los datos provienen de textos de sitios web y otros recursos no académicos, el uso masivo de GPT plantea problemas que debemos tener en cuenta. Uno de los principales riesgos

---

<sup>7</sup> Amanda Askill ha planteado estas preguntas respecto al chat GPT3 en un foro de filosofía. Ella es una joven filósofa que trabaja actualmente en *Anthropic*, una empresa emergente estadounidense de investigación y desarrollo de inteligencia artificial fundada por ex miembros de OpenAI. Su trabajo filosófico está dedicado principalmente al estudio de la inteligencia artificial en relación con la ética, con la teoría de la decisión y con la epistemología formal.

<sup>8</sup> Le preguntamos al ChatGPT acerca de sus habilidades y nos respondió: Generación de texto: GPT-3 es capaz de generar texto coherente y natural en diferentes idiomas y estilos. (Esto puede ser útil en la generación de contenidos para páginas web, redes sociales, marketing digital, entre otros). Traducción: GPT-3 tiene la capacidad de traducir texto de un idioma a otro de manera muy precisa y natural. (Esto puede ser de gran ayuda para empresas que operan en mercados internacionales). Asistencia virtual: GPT-3 puede utilizarse para desarrollar asistentes virtuales inteligentes que puedan responder preguntas y ayudar a los usuarios a realizar tareas específicas. Automatización de tareas: GPT-3 puede ayudar a automatizar tareas que requieren procesamiento de lenguaje natural, como la clasificación de documentos, la identificación de temas clave en grandes conjuntos de datos. Análisis de sentimiento: GPT-3 puede ser utilizado para analizar el sentimiento de los clientes o usuarios en las redes sociales o en los comentarios de los sitios web. (Esto puede ayudar a las empresas a entender mejor a sus clientes y mejorar sus productos y servicios). Aprendizaje automático: GPT-3 puede ser utilizado para desarrollar modelos de aprendizaje automático que puedan ser utilizados en diferentes campos, como la medicina, la ingeniería y la biología.



asociados a GPT es el sesgo en los datos, ya que el modelo puede replicar los sesgos presentes en los datos de entrenamiento, lo que podría llevar a que se tomen decisiones de manera errónea o peor aún, que sean discriminatorias (Scasserra, 2019). Una de las razones es que en el proceso decisorio las redes neuronales funcionan como *cajas negras*, esto es, no se puede acceder a la interpretabilidad de proceso<sup>9</sup> (Shwartz-Ziv y Tishby, 2017).

Otro grave problema es la privacidad y otro la seguridad, ya que en el procesamiento de grandes cantidades de datos es posible que se filtre información confidencial (ya ha ocurrido antes y podría volver a ocurrir) y esa es la razón por la que en algunos países, como Italia, por ejemplo, el uso de chat GPT fue prohibido.

Además, si bien GPT-3 produce texto coherente y gramaticalmente correcto, no tiene la capacidad de verificar la veracidad de la información que presenta y, por lo tanto, se corre un serio riesgo de desinformación. El modelo puede generar texto en apariencia auténtico, pero efectivamente falso y esto contribuye a la propagación de información errónea -o sencillamente falsa- en las redes sociales, plataformas digitales y en otros medios de comunicación.

Otro de los peligros asociados al uso del ChatGPT, y en estrecha relación con lo anterior, es la posibilidad de manipulación de los individuos por medio de la manipulación de las emociones, ya que el chat puede ser utilizado para crear noticias falsas con un formato veraz, creando para tal fin informes ilustrados con fotos, voces y/o videos falsos, por ejemplo, mensajes de phishing, etc.

Esta última cuestión también está relacionada con la controversia de su uso en el ámbito educativo y académico tanto por el peligro de propagar información falsa como por el riesgo de socavar la credibilidad de los investigadores serios al ser empleado por otros agentes (menos serios) para “fabricar” *papers*. (Elali, 2023).

Además, desde otra perspectiva que importa tener en cuenta, el empleo de GPT trae aparejado graves problemas ambientales. Algunos datos con el propósito de ilustrar este peligro: el ChatGPT emite 8,4 toneladas de dióxido de carbono al año. Se estima que se necesitó 78.437 kWh de electricidad para entrenar el modelo de lenguaje GPT-3 y para

---

<sup>9</sup> El programador solo conoce los datos de entrada con los que se ha cargado el sistema, *el input* y el resultado, *el output*, pero no puede acceder al proceso decisorio del algoritmo. La decisión se autogenera teniendo en cuenta las categorías que el algoritmo produce



entrenar a la IA correspondiente se consumieron 700 mil litros de agua dulce en aparatos de refrigeración para que las máquinas no se sobrecalienten. Por último, el ChatGPT consume medio litro de agua cada vez que “conversa” con un usuario. (Li, Yang, Islam, y Ren, 2023). Si se tiene en cuenta la inmensa cantidad de usuarios contabilizada en muchos miles de miles diarios<sup>10</sup>, el consumo es sencillamente alarmante.

Finalmente, no queremos dejar sin mencionar que los modelos de lenguaje generativo, como GPT y sus sucesores, nos obligarán a analizar, redefinir y precisar los conceptos clave y centrales de la filosofía, tales como ‘inteligencia’, ‘agencia’ y ‘conciencia’, por citar algunos ejemplos. En este sentido, David Chalmers, filósofo analítico australiano, especializado en filosofía de la mente y del lenguaje, cree que si bien es improbable que los grandes modelos lingüísticos actuales sean conscientes, sí debería tomarse en serio la posibilidad de que sus sucesores puedan serlo, lo que además, según él, será posible en un futuro no muy lejano. (2023). Acerca de si el modelo es o no inteligente, la línea de demarcación es menos clara, pues no dudaríamos en decidir que lo es, si esas tareas las llevara a cabo a un ser humano.

Entonces ¿podríamos afirmar que GPT es un buen modelo para entender la mente humana? “La primera respuesta es un rotundo no”, nos dice el filósofo español especializado en filosofía de la Inteligencia Artificial Santiago Sánchez-Migallón. Y la razón reside, según él, en su forma de funcionar, puesto que el modelo parece carecer de cualquier tipo de semántica, no parece comprender absolutamente nada de lo que dice y solo imita estadísticamente otros textos dados. Sin embargo, pese a su imposibilidad de comprender cualquier clase de significados, sus resultados son eficaces para resolver problemas que sí requieren la comprensión de significados. Y aquí, en este punto, el filósofo español se plantea preguntas importantes, a saber: ¿Qué tipo de comportamiento inteligente puede emerger de sistemas de fuerza bruta como GPT-3? ¿Qué tipo de patrones encuentran en el lenguaje que les sirven para ese buen desempeño lingüístico? ¿Se asemejan a los que utilizamos los humanos? Comprender bien esto, nos alerta, podría arrojar mucha luz a la comprensión sobre cómo los humanos generamos o aprendemos a usar el lenguaje. (2022, p.2)

Y a estas preguntas se suman otras más: ¿Cuáles son los límites de los modelos del lenguaje? ¿Qué afirmamos cuando afirmamos que estos modelos comprenden (o que no

---

<sup>10</sup> Se estima que ChatGPT recibe hasta 10 millones de solicitudes diarias (Tom Goldstein, 2022).



comprenden)? ¿Cómo debemos evaluar las capacidades de los grandes modelos de lenguaje? ¿Es posible considerar esta clase de modelo de lenguaje como modelos de verdadera creatividad? ¿Es posible que la inteligencia artificial pueda reemplazar algún día la creatividad humana y la capacidad de generar ideas nuevas?

Y esas preguntas disparan a su vez un sinfín de otras preguntas y cuestiones éticas: dada la capacidad de GPT para automatizar tareas y procesar grandes cantidades de información: ¿qué efecto tendrá la automatización de tareas en la economía y en la sociedad en general? ¿Cómo podemos garantizar que la tecnología se utilice de manera responsable para beneficio de la sociedad en su conjunto? Y finalmente, ¿cómo podrían la IA y el aprendizaje automático transformar la distribución del poder en la sociedad, nuestro discurso político, nuestras relaciones personales y nuestras experiencias estéticas?

En conclusión, un examen (no exhaustivo) del modelo de lenguaje GPT como caso paradigmático de la IA ha contribuido a generar muchas preguntas, algunas de las cuales demandan una respuesta concreta y otras, la mayoría de ellas, invitan al debate y a la reflexión filosófica. Hoy más que nunca estamos llamados a filosofar, a repensar la relación *mundo, tecnología y cultura*.

## REFERENCIAS BIBLIOGRÁFICAS

Bostrom, N. (2016). *Superintelligence: Paths, Dangers, Strategies*. Oxford, UK: Oxford University Press.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P. y Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, 33, 1877-1901.

Chalmers, D. J. (2023). Could a large language model be conscious? *arXiv preprint arXiv:2303.07103*.

Elali, F. R., & Rachid, L. N. (2023). AI-generated research paper fabrication and plagiarism in the scientific community. *Patterns*, 4(3).

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). Gpts are gpts: An early look at the labor market impact potential of large language models. *arXiv preprint arXiv:2303.10130*.



- Eyal, N. (2014). *Hooked: how to build habit-forming products*. Nueva York: Penguin Group.
- Goldstein, J. A., Sastry, G., Musser, M., DiResta, R., Gentzel, M. y Sedova, K. (2023). Generative language models and automated influence operations: Emerging threats and potential mitigations. *arXiv preprint arXiv:2301.04246*.
- Hidalgo, D., (2021). *Anestesiados. La humanidad bajo el imperio de la tecnología*, Madrid: Catarata.
- Li, P., Yang, J., Islam, M. A., & Ren, S. (2023). Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models. *arXiv preprint arXiv:2304.03271*.
- Lee, C. S., Wang, M. H., Yen, S. J., Wei, T. H., Wu, I. C., Chou, P. C., ... & Yan, T. H. (2016). Human vs. computer go: Review and prospect [discussion forum]. *IEEE Computational intelligence magazine*, 11(3), 67-72.
- Liu, X., Zheng, Y., Du, Z., Ding, M., Qian, Y., Yang, Z., & Tang, J. (2021). GPT understands, too. *arXiv*. <https://doi.org/10.48550/arXiv.2103.10385>
- Marcel, R. V. P., Fernando, B. E. M., y Roberto, Y. V. J. (2023, junio). "A brief history of the artificial intelligence: chatGPT: The evolution of GPT". In *2023 18th Iberian Conference on Information Systems and Technologies (CISTI)* 1-5. IEEE.
- Sánchez-Migallón Jiménez, S. (2022). En la mente de GPT-3. *Ciencia Cognitiva*, 16:1, 1-3.
- Scasserra, S. (2019). El despotismo de los algoritmos: Cómo regular el empleo en las plataformas. *Nueva Sociedad*, (279), 133-140.
- Weiser, M. (1991). The Computer for the 21 st Century. *Scientific american*, 265(3), 94-105.