Las metáforas que la inteligencia artificial no sabe interpretar

Metaphors that artificial intelligence doesn't know how to interpret

Christian Emmanuel Hernández Esquivel
Universidad Autónoma del Estado de México
cehernandeze@uaemex.mx

Recibido: 02 de septiembre de 2025 *Aceptado: 8 de enero de 2026*

Resumen
La generación de imágenes a partir de programas de Inteligencia Artificial (IA) ha llevado a una revolución en la manera en que se interpretan visualmente conceptos, términos y representaciones. Plataformas como ChatGPT se jactan de "transformar ideas abstractas en metáforas visuales" la realidad muestra la presencia de limitaciones en los modelos computacionales actuales. Hace necesario el diseño de "estímulos" (prompt, en inglés) específicos para que la IA arroje los productos esperados. En las siguientes paginas se despliegan los resultados de investigación a partir de la utilización de programas como Dall-e, Stable Diffusion, Ideogram y Flux Pro AI que fueron utilizados para interpretar visualmente metáforas pertenecientes a la vida cotidiana. El objetivo es compartir los sucesos que se obtuvieron visualmente al colocar las metáforas escritas como prompts en los generadores de imágenes. Los alcances evidencian la falta de comprensión tanto de analogías, semejanzas y comparaciones; las cuales la IA las interpreta visualmente de manera explícita sin acceder al significado implícito de las mismas. Existe cierto temor entre la comunidad artística respecto a que la IA desplace o sustituya al arte humano. Sin embargo, tal como se muestra en esta investigación, la IA sólo es una herramienta técnica ligada a tipos y estereotipos. En otras palabras, mientras la IA no sea capaz de entender cómo y de qué manera se estructuran las metáforas, no será posible acceder a los verdaderos significados que conforman la experiencia humana.

Palabras clave
Inteligencia Artificial, metáfora, generación de imágenes, interpretación, traducción visual.

Abstract
The generation of images through Artificial Intelligence (AI) programs has led to a revolution in how concepts, terms, and representations are visually interpreted. Platforms such as ChatGPT boast of "transforming abstract ideas into visual metaphors," but reality reveals limitations in current computational models. This necessitates the design of specific prompts in order for AI systems to produce the expected results. The following pages present research findings obtained using programs such as Dall-e, Stable Diffusion, Ideogram, and Flux Pro AI, which were employed to visually interpret metaphors drawn from everyday life. The objective of this study is to share the visual outcomes obtained by placing written metaphors as prompts in the image generators. The findings demonstrate a lack of understanding of analogies, similarities, and comparisons; the AI interprets these visually in an explicit way without accessing their implicit meanings. There is some fear within the artistic community that AI may displace or replace human artistic creation. However, as this research suggests, AI is merely a technical tool tied to types and stereotypes. In other words, as long as AI remains unable to understand how and in what way metaphors are structured, it will be incapable of accessing the deeper meanings that shape human experience.

Key words
Artificial intelligence, metaphor, image generation, interpretation, visual translation.

Introducción

La metáfora es una figura retórica que se utiliza para generar una relación de sentido entre dos elementos lingüísticos (por lo regular, dos sustantivos) en la que se reformulan los significados propios pertenecientes a ellos. Es una especie de comparación que no utiliza la palabra "como". A diferencia de lo que pudiera parecer, la metáfora no sólo compara objetos con sujetos o sujetos con objetos sino experiencias completas con otras iguales o diferentes. Así, existen un sin número de expresiones textuales que ponen de manifiesto la vivencia subjetiva del poeta a través de imágenes altivas, sublimes o elaboradas; la mayor parte de ellas relacionadas con las personas amadas. Algunos ejemplos se ubican en la literatura o la lírica formal, así como en la letra de las canciones populares. Tomando como ejemplo el inicio de "El son de la negra", canción popular mexicana:

Negrita de mis pesares,
ojos de papel volando,
a todos diles que sí
pero no les digas cuando.

Sin utilizar la palabra "como", el autor de esta canción compara los ojos de su amada con pedacitos de papel picado flotando en el aire. Aunque existen otras versiones de la letra, incluyendo una en la que los "ojos" se cambia por "hojas", resulta claro que la canción hace referencia a una mujer afromexicana.

Este aspecto lo aborda de manera más profunda el escritor Eduardo Mejía en un post en su blog personal titulado *"¿Hojas u ojos? Replica a Gabriel Zaid"*. https://erratasputocom.blogspot.com/2011/01/hojas-u-ojos-replica-gabriel-zaid.html

Las metáforas constituyen uno de los elementos centrales tanto de la lírica culta como de la popular. Pero no todos tienen la capacidad de analizarlas y, más aún, de encontrar y comprender sus significados. Minaudo (2013) señala que la utilización de la metáfora permite "trasladar el sentido recto de las voces a otro figurado". Y este sentido figurado (no textual) es el que conforma la base medular detrás de ellas. Cobarruvias (1979, como es citado por Fernández 2018), indica: para comprender la metáfora se debe clarificar "el concepto más oscuro y escondido" de la frase. Por lo tanto, se tendría que preguntar: ¿la IA es capaz de reconocer las metáforas?, es decir, ¿la IA tiene capacidad para comprender los significados connotativos provenientes de la experiencia humana?

Marco teórico

En un intento por humanizar la Inteligencia Artificial, diversos autores han señalado la supuesta capacidad de la IA para acceder a un nivel intelectual superior al de los humanos (Faruque, 2022). Sin embargo, los fenómenos descritos tanto por los programadores como por los usuarios de este tipo de tecnología provienen solamente de la experiencia de interacción con chatbots ("bots conversacionales") en los que se interactúa de manera escrita con la IA que simula ser una persona. (Faruque, 2022)

Los chatbots escritos (como el ChatGPT, Google Gemini, Microsoft Copilot, Meta AI, etc.) permiten el análisis de las frases que el usuario dicta o teclea en sus diversos dispositivos. En este sentido, la IA de los chatbots puede asumir varios roles: desde una asistente virtual que proporciona datos alojados en la web sobre distintos temas, pasando por una experta en ciencia que puede efectuar operaciones lógico-matemáticas hasta encarnar de manera teatral un personaje que puede asumir el rol de una pareja sentimental o sexual. (Wikipedia, 2024)

Este tipo de interacciones virtuales han llevado a usuarios y a programadores a esgrimir la idea de que la IA es "sensitiva" (sentient, en inglés), es decir, que puede reconocer y experimentar emociones tal como lo realizan los humanos. (Faruque, 2022) Sin embargo, se debe entender que, si en algún momento de la conversación con el chatbot, el usuario le pide a la IA reflexionar sobre sí misma o sobre sus emociones, probablemente lo haga no porque realmente las posea sino porque su lenguaje de programación así se lo permite. (Larson, 2022)

Por todo esto, es importante subrayar que los diferentes programas de IA han sido creados para resolver cierto tipo de tareas específicas (no son omnisapientes, como se pretende divulgar). Eso implica que en su código fuente se presenten de manera explícita las tareas que pueden (y no pueden) resolver. Y, por lo tanto, el desarrollo de su gnosis está limitado a su lenguaje computacional.

El "Buscador de metáforas" y Dall-e

En el ChatGPT existe un bot titulado "Buscador de metáforas" que tiene como función analizar metáforas en textos (literarios y no literarios) y explicar sus significados. (ChatGPT, 2024)

Al utilizar dicho bot, los resultados que aporta resultan por demás interesantes. Por ejemplo, al preguntarle por el significado de la frase "Tus ojos son estrellas que titilan en el cielo de la noche", el bot explica

que los ojos "son algo que ilumina o inspira", que ofrecen "belleza y consuelo", y que, "en términos afectivos", la persona que se describe "tiene un efecto casi mágico, capaz de iluminar el ánimo o el entorno con solo su mirada". (ChatGPT, 2024)

Al solicitar al bot la generación de una imagen de dicha frase, los primeros resultados que arroja son:

Figura 1. Bot de tus ojos son estrellas que titilan en el cielo de la noche



Nota: Open IA (2024). ChatGPT. [Modelo de lenguaje extendido]. https://chatgpt.com/

Resulta interesante notar que en ChatGPT, a través de su generador Dall-e, les otorga un género (dos mujeres, un hombre) a los sujetos generados, siendo que en la frase en cuestión nunca están explícitos dichos géneros.

Por otra parte, al realizar la revisión de los nombres otorgados por el bot a los archivos de las imágenes, aparecen los prompts que fueron utilizados para generarlas, todos ellos en inglés:

> *A poetic, dreamlike image of a face with eyes that resemble stars, shining brightly in the dark sky of the night.*
> *A poetic, surreal image of a face with eyes resembling bright stars, softly glowing in the dark sky of the night.*
> *A poetic, enchanting image of a face with eyes that resemble shining stars, twinkling softly in a vast night sky.*
>
> (ChatGPT, 2024)

Resaltando que el verbo "parecerse a" (to resemble) y sus derivaciones aparecen en cada uno de los prompts. En este sentido, para generar dichas imágenes, además de traducir la frase del español al inglés, ChatGPT traduce el sentido literal de la metáfora.

Con todo esto, al preguntarle al "Buscador de metáforas" por los versos de apertura de "El son de la negra", la IA nos explica que:

A nivel conceptual, la frase podría estar reflejando la metáfora de la levedad o fragilidad de los sentimientos (al utilizar el papel y el vuelo) para expresar algo que es difícil de alcanzar, de concretar o de sostener en el tiempo. En conjunto, este verso puede estar evocando una relación cargada de sentimientos que no son completamente correspondidos, o que son, por su naturaleza, inasibles y que evocan tristeza o nostalgia. Las metáforas invocan a una persona que produce "pesares", pero cuyo aspecto (los "ojos de papel volando") también la hace etérea o inalcanzable. (ChatGPT, 2024)

Sin embargo, las imágenes que arroja son:

Figura 2. Bot de "ojos de papel volando" ejemplos



Nota: Open IA (2024). ChatGPT. [Modelo de lenguaje extendido]. https://chatgpt.com/

El mismo problema sucede en otras plataformas de generación de imágenes:

Figura 3. Ejemplos del Bot de "ojos de papel volando" en diversas plataformas de IA



Stable Diffusion



Ideogram



Flux Pro AI

El lector crítico podrá advertir que el problema no es de interpretación per se sino de la traducción que la IA realiza al pasar al inglés dichas frases en español. Por ello, en un segundo momento, y utilizando el listado de metáforas en inglés publicadas en el portal Verbal Versa (2024), se seleccionaron tres expresiones retóricas que tienen que ver con los ojos:

1. eyes like raining windows (lit. "ojos como ventanas lluviosas")
2. eyes like unfurling petals (lit. "ojos como pétalos desplegándose")
3. eyes like warm embers (lit. "ojos como brasas candentes").

Al introducir dichas frases en los diversos generadores de imágenes, los resultados fueron:

Figura 4. Ejemplos del Bot de "ojos como ventanas lluviosas" en diversas plataformas de IA.



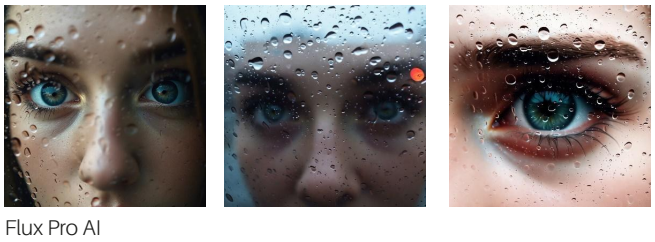
Dall-e



Stable Diffusion



Ideogram

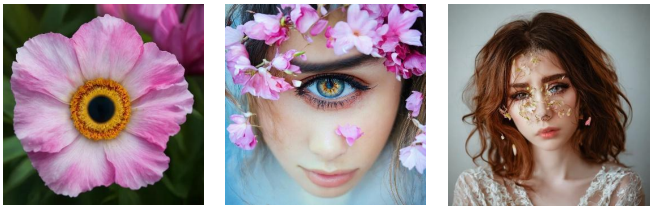


Flux Pro AI

Figura 4. Ejemplos del Bot de "ojos como pétalos desplegándose" en diversas plataformas de IA.



Dall-e



Stable Diffusion



Ideogram



Flux Pro AI

Figura 5. Ejemplos del Bot de "ojos como brasas candentes" en diversas plataformas de IA.



Dall-e



Stable Diffusion



Ideogram



Flux Pro AI

Según Verbal Versa (2024), "to have eyes like raining windows" significa mirar con tristeza; "to have eyes like unfurling petals" significa mirar sorprendido; mientras que "to have eyes like warm embers" significa mirar con afecto y agrado. Queda a juicio del lector si las imágenes generadas por la IA lograron representar de manera fehaciente dichas emociones o no.

Por otra parte, Kwon (2024) señala que la naturaleza altamente descriptiva en los prompts que deben utilizarse en los generadores de imágenes necesitan ser cada vez más específicos, utilizando un lenguaje técnico altamente denotativo que incluya los aspectos perceptuales de la imagen (personas, espacios, ambientes, conceptos, colores, estilos y atmósferas) así como los aspectos técnicos de la misma (composición, plano, posición de la cámara, iluminación, textura, tipo de captura, calidad de grano, etc.). Y, por supuesto, en el caso de la generación de imágenes de personas, indicar, además del género, edades y franjas etarias, nacionalidades o grupos raciales, color y tipo de vestimenta, así como rasgos faciales y corporales.

Con base en los ejemplos anteriores que han mostrado, la generación de todo un rostro haciendo alusión solamente a los ojos, los resultados suelen ser impredecibles. Si se busca obtener una imagen acorde a las expectativas del usuario, se debe realizar una descripción exhaustiva de la misma.

Con todo esto, para poder generar una imagen que represente a la mujer afromexicana enunciada en "El son de la negra", se procedió a realizar una descripción textual de dicho personaje, cuyo prompt traducido al inglés fue el siguiente:

> *A black-skin traditional beautiful Mexican young woman with her eyes like confetti floating in the air.*

Como es evidente, la generación de dicho prompt vuelve a quedar sujeta a la traducción que realiza la IA, ya que "papel picado" queda traducido como "confetti". Al introducir dicha frase en los generadores de imágenes, los resultados fueron los siguientes:

Figura 6. Imagen que represente a la mujer afromexicana



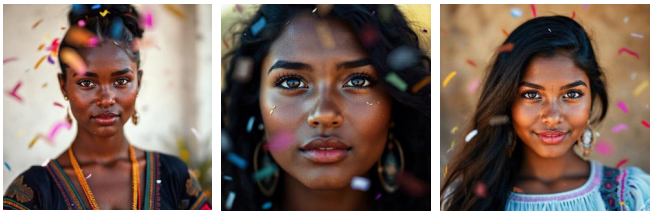
Dall-e



Stable Diffusion



Ideogram



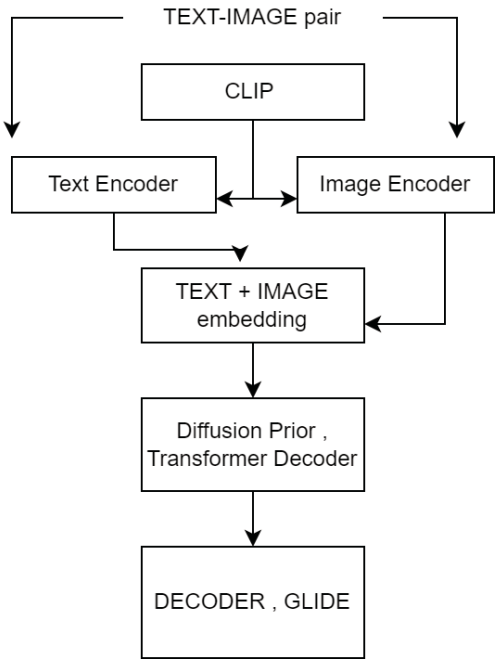
Flux Pro AI

Discusión

De acuerdo con Gupta (2023), los generadores de imágenes funcionan a través de un esquema de flujo en el que se da un emparejamiento entre el texto y la imagen, codificando tanto el texto introducido por el usuario como la imagen que se genera a partir de dicho texto [ver figura. 7]. En el caso específico de Dalle-e, a este sistema se le da el nombre de CLIP (Contrastive Language–Image Pre-training o "Preentrenamiento contrastivo de lenguaje e imagen") el cual posibilita una generación de imágenes casi-automática a partir de la implementación de tres procesos:

1. la codificación del texto introducido por el usuario y su clasificación temática
2. la asociación de ciertos fragmentos de dicho texto con una serie de categorías visuales predeterminadas por la IA
3. la generación de una imagen específica que cumpla tanto con la descripción proporcionada por el usuario como con la categoría visual con la que fue asociada por la IA.

Figura 7. Esquema de flujo



Nota. Esquema de flujo que explica de qué manera se da el proceso de generación de imágenes en la plataforma DALL-E. Adaptado de "How does DALL·E, the text-to-image generator work?", por Gupta, 2023

Esto implica la existencia de una serie de repositorios digitales (datasets) en la cual las imágenes preexistentes a la difusión sirven de base a la IA para sugerir las imágenes solicitadas. Lo que provoca la existencia de un universo visual limitado. (Navarro, 2024)

Estudiar hasta qué punto las imágenes generadas por IA corresponden a estereotipos culturales es un tema para investigaciones futuras. Mientras el italiano Raigoso (2024), realizó análisis de los sesgos raciales y de género que Dalle-e presenta, en un intento de "generar conciencia entre los usuarios y desarrolladores de estas tecnologías, y contribuir a una sociedad con menos sesgos y estereotipos". (2023-2024, p. 4)

Una posible solución para evitar los "sesgos" (biases) sería revisar la configuración de cada uno de los datasets que los generadores de imágenes utilizan. Existen aún muchas dudas que los programadores de IA no han resuelto de manera satisfactoria: ¿Cómo se integran los datasets? ¿Bajo qué contexto se integran? ¿Se pueden realizar "auditorias" que garanticen la no existencia de sesgos ni estereotipos, tal como el autor Barve et al. (2025) lo ha planteado? Los generadores de imágenes mediante IA se entrenan con millones de imágenes extraídas de la web, junto con sus descripciones o etiquetas textuales. Por lo tanto, se debería cuestionar: ¿quiénes son los que suben a la web dichas imágenes? ¿El acceso a internet móvil limita el acceso a dicha tecnología? ¿Cuáles son las políticas de sitios públicos como Facebook o Instagram que resguardan las imágenes de millones de usuarios? Considerando que los usuarios de IA deberían demandar para que sean aclaradas por los desarrolladores de este tipo de tecnologías. En especial, después de las denuncias legales que algunos artistas estadounidenses han emprendido en contra de sus compañías.

El arte no debe integrarse por estereotipos porque esto implicaría una representación injusta de la realidad. El arte debe ser liberador: permitirles a las personas menos agraciadas por el sistema mostrar su realidad social por más cruenta que esta sea. En el mundo capitalista en el que vivimos, darles voz a los pobres, a las minorías económicas y a los desposeídos debería ser una tarea prioritaria si se busca fomentar la equidad. De no hacerlo, la IA se convertirá, igual que ha ocurrido con los mass media corporativos, en un espacio que sólo reproducirá los discursos hegemónicos que excluyen a los grupos marginados o subrepresentados. El acceso a la tecnología computacional y al servicio de Internet, por supuesto, también debería ser un

asunto de discusión pública, aunque no es tema de esta investigación.

Nota.
"Midjourney and two other A.I. imagery generators, Stable Diffusion and DreamUp... indexes more than five billion images from across the Internet, including the work of many artists." Chayka, K. (2023). "Is A.I. Art Stealing from Artists?" The New Yorker. https://www.newyorker.com/culture7infinite-scroll/is-ai-art-stealing-from-artists

Conclusión

En el caso de la generación de imágenes a partir de la metáfora inicial de "El son de la negra", la descripción de aquellos ojos de papel picado escapa al lenguaje denotativo, ya que deja a la imaginación y a la interpretación subjetiva de quien lee o escucha la metáfora la conformación de una imagen concreta de la misma. Sin embargo, al hacer explícito su color de piel, su nacionalidad y su pertenencia a una franja etaria, la mayor parte de los generadores de imágenes la contemplan. Se obtiene, por fin, una representación visual acorde al significado de la metáfora inicial que presenta la canción.

Podría concluirse entonces que, en el mundo de la generación de imágenes, aquello que no queda explícito queda a disposición de la IA. El algoritmo llena los vacíos de información con los datos presente en los datasets, deformando en muchas ocasiones el sentido original de la instrucción.

El objetivo general de este trabajo es mostrar la poca capacidad de la IA para traducir a un lenguaje visual el contenido latente de las metáforas escritas. Y, tal como se ha mostrado, esto es debido a la necesidad de utilizar un lenguaje extremadamente descriptivo que evite la utilización de lenguaje connotativo.

Referencias

Barve, S., Bhattacharya, S., Goswami, K., y Mukhopadhyay, S. (2025). *Can we debias social stereotypes in AI-generated images? Examining text-to-image outputs and user perceptions*. arXiv. https://arxiv.org/abs/2505.20692

Faruque, M. (2022). AI versus human consciousness: A future with machines as our masters? *Renovatio: The Journal of Zaytuna College*. https://renovatio.zaytuna.edu/article/ai-versus-human-consciousness

Fernández, V. H. (2018). *Diccionario práctico de figuras retóricas y términos afines: Tropos, figuras de pensamiento, de lenguaje, de construcción, de dicción y otras curiosidades.* Biblioteca Virtual Miguel de Cervantes. https://www.cervantesvirtual.com/obra/diccionario-practico-de-figuras-retoricas-y-terminos-afines-924724/

Gupta, M. (2023, 10 de octubre). *How does DALL·E, the text-to-image generator work? Medium*. https://medium.com/data-science-in-your-pocket/how-does-dall-e-the-text-to-image-generator-work-c2d9f4a0f26c

Kwon, D. H. (2024). Analysis of prompt elements and use cases in image-generating AI: Focusing on Midjourney, Stable Diffusion, Firefly, DALL·E. *Journal of Digital Contents Society*, 25(2), 341–354. http://journal.dcs.or.kr/xml/39751/39751.pdf

Larson, E. J. (2022). *The myth of artificial intelligence.* Harvard University Press.

Minaudo, J. (2013). *Los usos de la metáfora, el recurso metafórico y el psicoanálisis.* V Congreso Internacional de Investigación y Práctica Profesional en Psicología, Buenos Aires, Argentina. https://www.aacademica.org/000-054/225

Navarro, S. (2024, 4 de junio). *¿Qué son los datasets?* KeepCoding. https://keepcoding.io/blog/que-son-datasets/

Open IA. (2024). *Buscador de metáforas* [GPT personalizado]. ChatGPT. https://chatgpt.com/g/g-4yeqoyJpI

Raigoso, N. (2024). *Behind the screen: Biases and stereotypes in DALL·E AI-generated images* [Tesis de maestría, Politecnico di Milano]. Politesi. https://www.politesi.polimi.it/bitstream/10589/223348/3/2024_07_Raigoso.pdf

Verbal Versa. (2024). *70+ eyes metaphors with examples*. https://verbalversa.com/eyes-metaphors/

Wikipedia. (2024, 15 de diciembre). *Bot conversacional.* https://es.wikipedia.org/wiki/Bot_conversacional