



Cómo hacer cosas con palabras. Algunas reflexiones en torno a las inteligencias artificiales generativas de imágenes

How to do things with words. Some remarks on generative AI for images

Walter Koza

Universidad Nacional de General Sarmiento / CONICET / Asociación de Dibujantes de Argentina

Email

wkoza@campus.ungs.edu.ar

ORCID

0000-0002-5043-6561

RESUMEN. Actualmente, la inteligencia artificial (IA) se ha convertido en un tema de relevancia y que afecta a diversas disciplinas. Una IA puede pensarse como la capacidad de las máquinas para realizar tareas que normalmente requieren de la inteligencia humana. Dentro de las posibilidades que ofrece el campo, una de las más notables es la IA generativa. Esta se focaliza en la generación de contenido, desde cierta perspectiva, original, como ser texto, imagen, voz, video, entre otras. Para ello, se requiere que el usuario le proporcione a la IA una indicación del producto que desea, mediante un texto escrito, conocido como *prompt*. Ahora bien, un fenómeno observado es la dificultad de que el resultado generado se corresponda totalmente con lo que el usuario desea, siendo necesario explicitar al máximo la indicación y de la manera menos ambigua posible, lo que en ocasiones se dificulta debido a la naturaleza misma del lenguaje. Aquí presentamos un pantallazo general de lo que son las IA generativas de imágenes y el fenómeno de la ambigüedad lingüística, que incide en la redacción de *prompts* para la generación de imágenes.

Palabras Clave: inteligencia artificial generativa, generación de imágenes, *prompt*, ambigüedad lingüística, lenguaje natural.

ABSTRACT. Currently, artificial intelligence (AI) has become a relevant topic that affects various disciplines. AI can be thought of as the ability of machines to perform tasks that typically require human intelligence. Among the possibilities offered by the field, one of the most notable is generative AI. This focuses on generating, from original content, from a certain perspective, such as text, image, voice, video, among others. To do this, the user needs to provide the AI with an indication of the desired product, through a written text known as a *prompt*. However, an observed phenomenon is the difficulty of generating a result fully corresponding to what the user desires, necessitating the need to explicitly indicate as much as possible and in the least ambiguous way, which is sometimes challenging due to the nature of language itself. This article provides an overview of generative AI for images and the phenomenon of linguistic ambiguity. Subsequently, it reflects on how this affects the formulation of *prompts* for image generation.

Keywords: Artificial intelligence, images generation, *prompt*, linguistic ambiguity, natural language

*Un tipo encuentra una lámpara y, al frotarla, hace
aparecer un genio que le dice que le va a conceder un deseo.
El tipo dice que no quiere ver más pobreza en el mundo.
El genio lo deja ciego.*

1 | INTRODUCCIÓN

Desde hace tiempo, la inteligencia artificial (IA) viene siendo el tema del momento y sus usos pueden ser diversos: laborales o prácticos, como manejar autos; artísticos, como retocar fotografías de manera automática; científicos, como predecir evolución de tumores en el área de la medicina, o simplemente lúdicos. De manera sencilla, puede definirse una IA como un dispositivo o algoritmo diseñado para la resolución de problemas sociales para los que se requiere el uso de la inteligencia.

Entre los tipos de IA que actualmente se están desarrollando, las que mayor atención han recibido son las IA generativas de contenido. Estas se caracterizan, precisamente, por generar nuevo contenido a partir de datos procesados previamente en una etapa denominada *de entrenamiento*. Así, una IA como *ChatGPT*¹ (OpenAI, 2023) es capaz de crear textos de diferentes géneros: recetas de cocina, planes dietarios, ejercicios para estudiar, gacetillas de prensa, monografías, etcétera. Para lograr esto, es necesario que el usuario le indique al dispositivo lo que desea mediante un texto escrito, generalmente en inglés, denominado *prompt*. Este consiste en una cadena de caracteres (normalmente, organizados en grupos de palabras) que describe aquello que queremos que la IA realice. Por ejemplo, mediante la siguiente instrucción (1)² extraída de Lexic.art³, se obtiene como resultado la imagen de la Figura 1:

- (1) Coloring pages: a group of joyful little kids, adorned in angelic wings and halos, play on a heavenly playground filled with fluffy clouds, swings, and slides.

‘Páginas para colorear: un grupo de niños alegres, adornados con alas y halos angelicales, juegan en un parque celestial lleno de nubes esponjosas, columpios y toboganes’.

Al respecto, si bien la calidad de los resultados suele ser destacable e incluso sorprendente estos varían y no siempre son los que los usuarios desean. Con el ejemplo (2), extraído de la misma plataforma, se obtiene un resultado menos feliz (Figura 2):

- (2) Portrait of a blond mother and bald father holding a 8 age old young girl kid, digital art, realistic painting, character design

‘Retrato de una madre rubia y padre calvo sosteniendo a una niña de 8 años, arte digital, pintura realista, diseño de personajes’

¹<https://chat.openai.com/>

²Para mayor claridad, presentamos los prompts en su idioma original, cuando este sea en inglés, agregamos nuestra traducción.

³<https://lexica.art/>



FIGURA 1 Imagen generada por Lexic.art a partir del *prompt* de (1) (Fecha de consulta 12/6/2023)



FIGURA 2 Imagen generada en Lexic.art a partir del *prompt* (2). (Fecha de consulta 11/6/2023)

Esto puede explicarse tanto por limitaciones propias del algoritmo como por ciertas complejidades derivadas del sistema de la lengua. En este caso, la instrucción ‘retrato de una madre rubia y padre pelado’ ha sido entendida por el algoritmo como una combinación de ambas personas. Notemos que en la instrucción faltó colocar *un* antes de *padre* para que la instrucción fuera ‘Retrato de una madre rubia y **un** padre calvo...’ y así la IA entendiera que se trataba de dos entidades y no de una que sea madre, padre, rubia y calvo a la vez. Cosas que pasan...

Ante estas particularidades, vamos a presentar un pantallazo general de lo que son las IA generativas de imágenes y cómo influye la ambigüedad lingüística en el resultado. Para esto, reflexionaremos sobre qué significa *inteligencia artificial* (sección 2) y cómo se construyen (sección 3). Luego, nos focalizamos en las IA generativas de imágenes y la manera en que la ambigüedad lingüística incide en la redacción de *prompts*. Finalmente, les dejamos un par de reflexiones que surgieron a partir de nuestro interés por el tema. Pasen y lean.

2 | ¿DE QUÉ HABLAMOS CUANDO HABLAMOS DE IA?

2.1 | Inteligencia, inteligencia artificial, polisemia

En primer lugar, empecemos definiendo qué es una IA. Partamos de dos obviedades, en primer lugar, una IA es ‘algo’ a lo que le podemos pedir ‘algo’. Esto equivale a decir que es un dispositivo creado para realizar actividades que llevan a la consecución de un objetivo que nos interesa. En segundo lugar, una IA es una inteligencia que no es natural. Lejos de ser irónico, una inteligencia no natural va a permitir reflexionar sobre la noción misma de inteligencia y si esta es exclusiva o no de los seres humanos. Respecto del concepto de *inteligencia* hay abundantes estudios que tratan de explicarla, clasificarla y medirla (al final de este ensayo les dejo algo de lo que leí sobre el tema). De manera general, Rubén Ardilla, un investigador de la Universidad Nacional de Colombia, la define como un conjunto de características referidas a la capacidad de solucionar problemas, de razonar y de adaptarse al ambiente.

Para entrar en tema ya que andamos en estos asuntos, le pedimos a la IA ChatGPT que nos facilitara una lista de definiciones de inteligencia (Figura 3).

La síntesis obtenida nos permite observar que hay diversas acepciones de inteligencia, que se focalizan en distintos aspectos. Sin embargo, todos ellos parecen estar relacionados. Es decir, en mayor o menor medida, podemos etiquetar a la inteligencia como una *capacidad*. Ahora bien, una pregunta que sería lícita realizar es ¿cómo es posible que en una misma palabra se incluyan significados tales como *la habilidad para adaptarse* o *el potencial para la creatividad*? La respuesta a esta pregunta está dada por un concepto fundamental de la lingüística que es el de *polisemia*. La polisemia estudia los posibles significados que tiene un mismo término, según el contexto en el que se hallen (si hay entusiastas por este tema, al final mencionamos algunas lecturas). Esta palabra, desde una perspectiva etimológica, proviene del griego *poli* ‘muchos’ y *sema* ‘significado’.

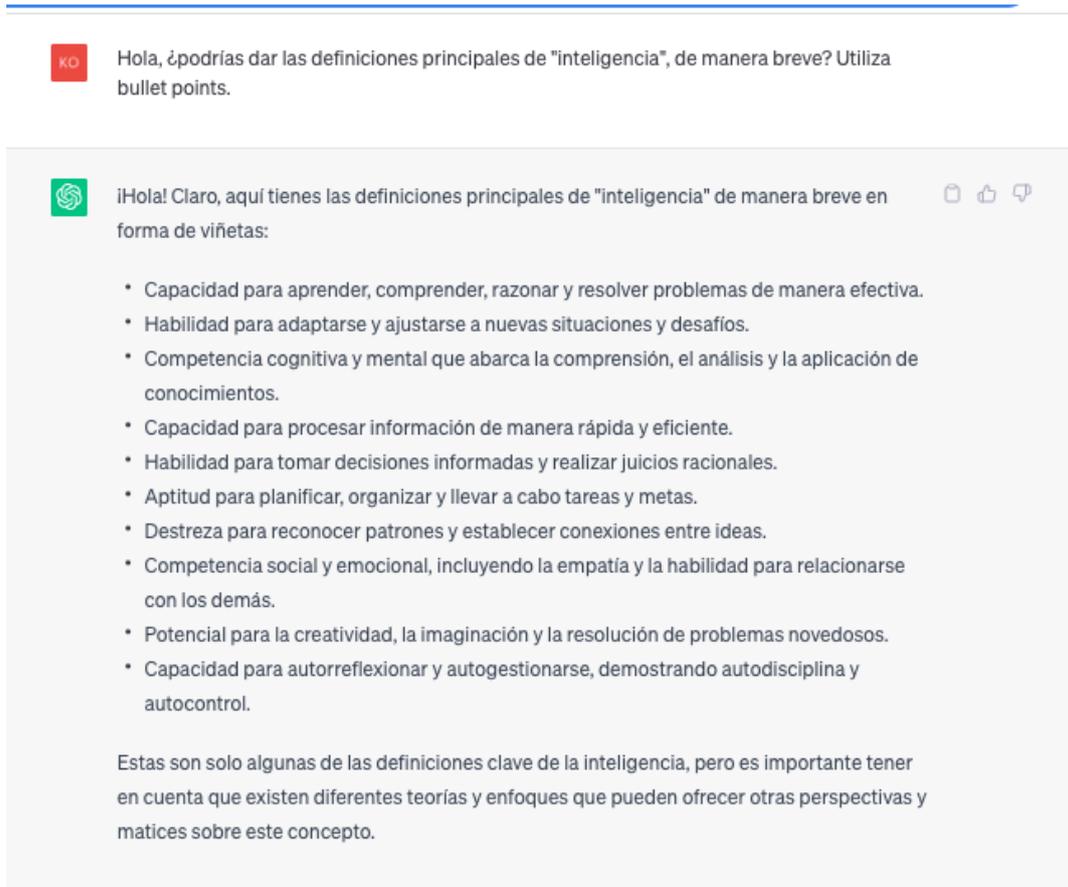


FIGURA 3 Respuesta de ChatGPT sobre las definiciones de inteligencia (consultado el 1/6/2023)

Pensemos en ejemplos sencillos como *carne* o *radiografía*. Ahora, fijémonos en las oraciones (3) y (4):

- (3) a. Juan no come carne de vaca.
b. Juan no come carne vegana.
- (4) a. El radiólogo le hizo una radiografía de tórax al paciente.
b. El paciente sacó la radiografía del sobre y se la pasó al médico para que la revisara.

Entonces, ¿la carne es carne cuando proviene de un animal (3a) o también cuando es fabricada con alimentos vegetales (3b)?, ¿una radiografía es un procedimiento que hace un radiólogo (4a) o es una placa negra y celeste –resultante del proceso anterior– que muestra cómo tenemos los huesos (4b)? Y la respuesta es que todas las opciones son correctas. Esto lo podemos explicar señalando que las palabras cobran significado solo cuando las usamos, y, como se ve, a algunas de ellas, los hablantes las podemos usar de diversas maneras, precisamente, porque son polisémicas. No obstante, los distintos usos deben estar relacionados entre sí: la *carne vegana* va a tener la forma y la apariencia de la ‘carne animal’, mientras que el producto *radiografía* va a ser el resultado de ‘sacarse una radiografía’, por ejemplo. Si se ponen a pensar, van a salir muchos ejemplos más: la *nube* que está en el cielo y la *nube* donde guardamos los archivos de la computadora; la *ventana* de nuestra casa y la *ventana* de Windows, etcétera.

Entonces será tan válido referirse a *inteligencia artificial*, como a *revólver de juguete* o *carne vegana*. De

esta manera, si bien es cierto lo expresado por la investigadora Karina Pedace⁴, respecto de que las IA no son “tan inteligentes en el sentido de que no son capaces de generar pensamientos por sí solas”, sí lo serían en su capacidad de resolver algunos problemas para los que se requiere inteligencia.

En definitiva, la IA es una rama de las ciencias computacionales que se encarga del estudio de modelos de cómputos capaces de realizar actividades propias de los seres humanos a partir del razonamiento y la conducta. En la sección siguiente, les contamos brevemente cómo se logra esto.

3 | ¿CÓMO SE CONSTRUYE UNA IA?

En la creación de las IA están involucradas diversas metodologías, siendo el aprendizaje automático (en inglés, *machine learning*, ML) la que más aceptación ha tenido. Hablar de ML implica referirse a una noción especial de lo que significa ‘aprendizaje’, que difiere de la noción que pueda aplicarse a los seres vivos. Decimos que una computadora aprende cuando es capaz, a partir de una indicación, de imitar una conducta humana inteligente. Una manera sencilla de definir al ML es concebirlo como una técnica que busca automatizar el aprendizaje, a partir de la exploración estadística de un conjunto de datos que permiten construir un modelo, es decir, una abstracción que funciona como molde para nuevas soluciones. Esto permite automatizar la construcción y programación de modelos, a partir del descubrimiento sistemático de patrones significativos.

Pensemos un ejemplo sencillo con una serie numérica elaborado a partir de la propuesta de Dot CSV⁵, supongamos que tenemos los siguientes datos estructurados con un número una operación matemática que no conocemos ‘(...)’ y un resultado:

- 1 (...) = 10
- 2 (...) = 20
- 3 (...) = 30
- 4 (...) = 40

¿Cuál sería el resultado de ‘5 (...)’, ‘8 (...)’ y ‘12 (...)’? La respuesta es sencilla: 50, 80 y 120 respectivamente. Aún más, vamos a poder decir sin mucho esfuerzo que ‘(...)’ significa multiplicar por 10 el número de la izquierda. De esta forma, gracias a los datos previos, hemos creado el siguiente molde ‘X x 10 = Y’, que nos va a servir para dar la respuesta correcta a cualquier valor que tome ‘X’. Pues bien, básicamente, es esto lo que se pretende lograr en un dispositivo computacional: que, sin necesidad de explicitación, pueda descubrir el patrón subyacente en un conjunto de elementos.

Mediante la utilización de datos de muestra o entrenamiento, los algoritmos de ML logran crear automáticamente un modelo matemático para tomar decisiones sin que se los haya programado específicamente para ello. A partir de las inferencias aprendidas, estos algoritmos clasifican nuevos elementos de la misma naturaleza o predicen cualidades y tendencias. Por otra parte, los algoritmos involucrados en las IA generativas permiten crear nuevos datos, a partir de la consideración de las características estadísticamente relevantes de los datos previos con las que fueron entrenadas.

⁴<https://www.pagina12.com.ar/528835-la-inteligencia-artificial-es-un-certificado-de-poder> (consultado el 6/6/2023).

⁵<https://www.youtube.com/watch?v=oT3arRRB2Cw&t=112s> (consultado el 6/6/2023).

Se pueden mencionar tres tipos de ML: aprendizaje supervisado, aprendizaje no supervisado y aprendizaje por refuerzo. Veamos de qué se trata cada uno de ellos.

3.1 | Aprendizaje supervisado

En este tipo de aprendizaje, los algoritmos usan datos que fueron etiquetados u organizados previamente, de forma manual. De este modo, se indica cómo deberá ser categorizada la nueva información. Por ejemplo, si se quiere obtener un algoritmo que identifique gatos, habría que entrenarlo con imágenes de estos animales, junto con la etiqueta pertinente, como se muestra en la Figura 4:

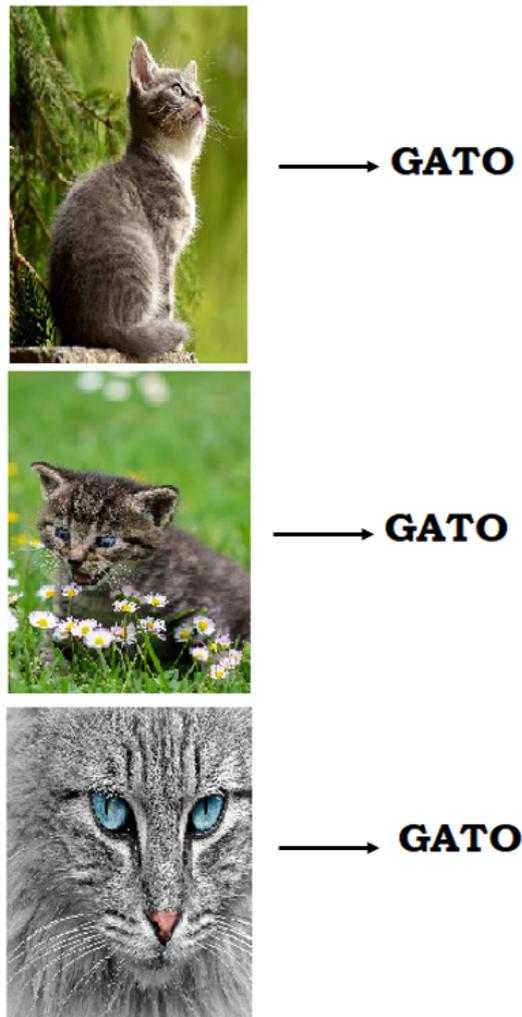


FIGURA 4 Estructura de datos para el entrenamiento de ML supervisado⁶

El propósito del ML supervisado es crear una función que pueda predecir el valor de salida para cualquier entrada dada, luego de haber aprendido con los datos de ejemplo. En este caso, el objetivo es que pueda asignar la etiqueta 'GATO' a imágenes de gatos que no estén en el corpus de entrenamiento y solo a esas (es decir, que no asignen esa etiqueta a la foto de un perro, por ejemplo).

Como dato de color, podemos cerrar esta subsección diciendo que todos los usuarios de internet hemos

⁶Imágenes extraídas de <https://pixabay.com/es/images/search/gato/> (consultado el 5/5/2023)

colaborado, voluntaria o involuntariamente, en el etiquetado de imágenes cuando quisimos a entrar a algún sitio que nos exigía, previamente, resolver un *captcha*, como el que se ve en la Figura 5.

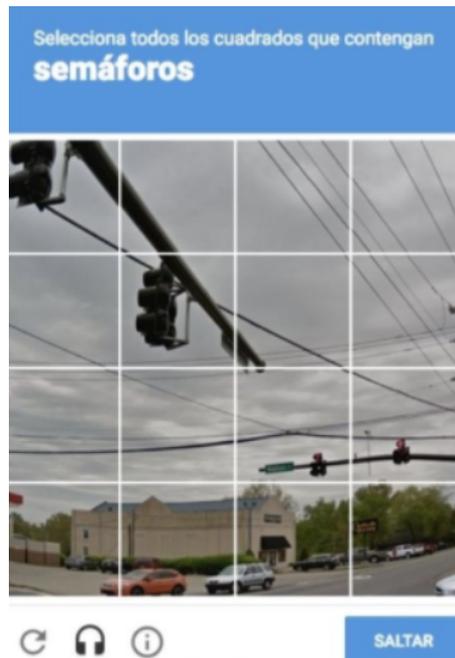


FIGURA 5 Ejemplo de captcha⁷

Cuando clickeamos los fragmentos de la fotografía que contienen semáforos, al mismo tiempo “ayudamos” a los algoritmos a contar con datos más fidedignos.

3.2 | Aprendizaje no supervisado

A diferencia del anterior, para el entrenamiento, no se usan datos etiquetados ni organizados previamente, sino que es el dispositivo el que debe encontrar patrones de clasificación y no se requiere de intervención humana. Así, la figura 6 presenta un set de imágenes sin etiquetar⁸:



FIGURA 6 Conjunto de imágenes de perros y gatos

⁷Imagen extraída de <https://talentocorporativo.com/wp-content/uploads/2021/01/captcha-semaforo.png> (consultado el 12/9/2023).

⁸Imágenes extraídas de <https://pixabay.com/es/images/search/gato/> y de <https://pixabay.com/es/images/search/perro/> (consultado el 5/5/2023).

Este tipo de algoritmos, a partir de la coincidencia de rasgos comunes, podría agrupar cada fotografía como se muestra en la Figura 7.

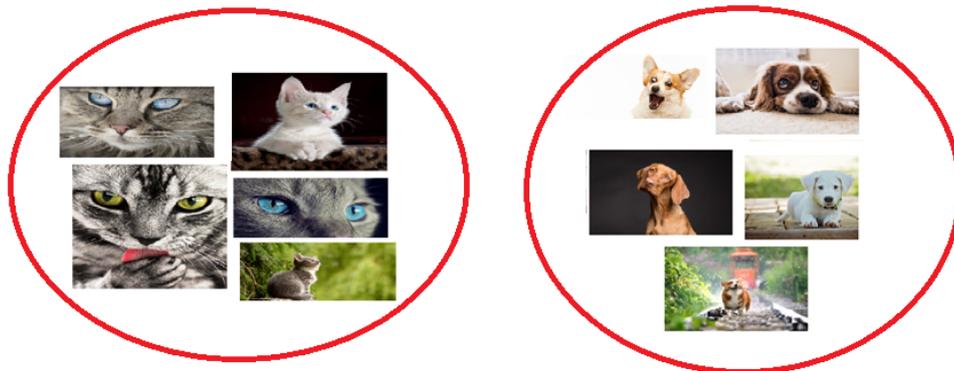


FIGURA 7 Ejemplo de clasificación automática de perros y gatos

Aquí, el sistema decidirá la agrupación que considere más pertinente de acuerdo con las características de cada dato ingresado al sistema de entrenamiento.

3.3 | Aprendizaje por refuerzo

El aprendizaje por refuerzo se diferencia de los modelos anteriores en que, mientras que los primeros no realizan modificaciones en el modelo, este se reajusta y mejora constantemente a partir de una recompensa, entendida esta como una señal que indica la conveniencia o no de una serie de acciones llevadas a cabo. Por ejemplo, en una IA generativa de imágenes, los dispositivos suelen dar un conjunto de resultados para que el usuario elija el más adecuado. En la Figura 8, se presentan 4 imágenes generadas en Playground AI⁹ a partir del siguiente *prompt*:

- (5) A dog playing with a cat in the garden, in the morning, children book style.
'Un perro jugando con un gato en el jardín, en la mañana, estilo de libro para niños.'



FIGURA 8 Generación automática de un perro jugando con un gato¹⁰

⁹<https://playgroundai.com/> (consultado el 6/6/2023).

¹⁰Generado en Lexic.art (12/6/2023).

Si observamos con atención, de las 4 imágenes, la tercera sería la que más se acerca a lo declarado en el *prompt* (aunque la ilustración parece presentar a los dos animales posando más que jugando). Por ende, seleccionarla implicará recompensar ese resultado por sobre los demás.

En resumen, el aprendizaje por refuerzo genera un valor interno para los estados o acciones intermedias con el propósito de acercarse al objetivo propuesto y así conseguir la recompensa. Una vez aprendido el mecanismo, el algoritmo tomará los caminos en los que ha recibido recompensa, descartando otras posibilidades.

4 | INTELIGENCIAS ARTIFICIALES GENERATIVAS DE IMÁGENES Y COMPLEJIDAD LINGÜÍSTICA

4.1 | Inteligencias artificiales generativas de imágenes

Como se mencionó más arriba, existen diversas aplicaciones para las IA. Entre ellas, se puede mencionar a las IA generativas de contenido o IA generativas. Estas tienen como fin la creación de nuevos datos, similares a los de su entrenamiento. Por ejemplo, en el caso de ChatGPT, a partir de los textos involucrados en su entrenamiento, es posible la obtención de nuevos textos originales. Este tipo de inteligencias se logra mediante una técnica denominada redes neuronales: una estructura que se organiza de manera similar a la neurona humana y es capaz de producir distintos productos a partir del procesamiento de grandes masas de datos¹¹.

Aquí nos centraremos en la creación de imágenes. Como se mencionó más arriba, para obtener una imagen es necesario escribir un *prompt*, es decir, un input lingüístico. La Figura 9 ejemplifica el resultado:

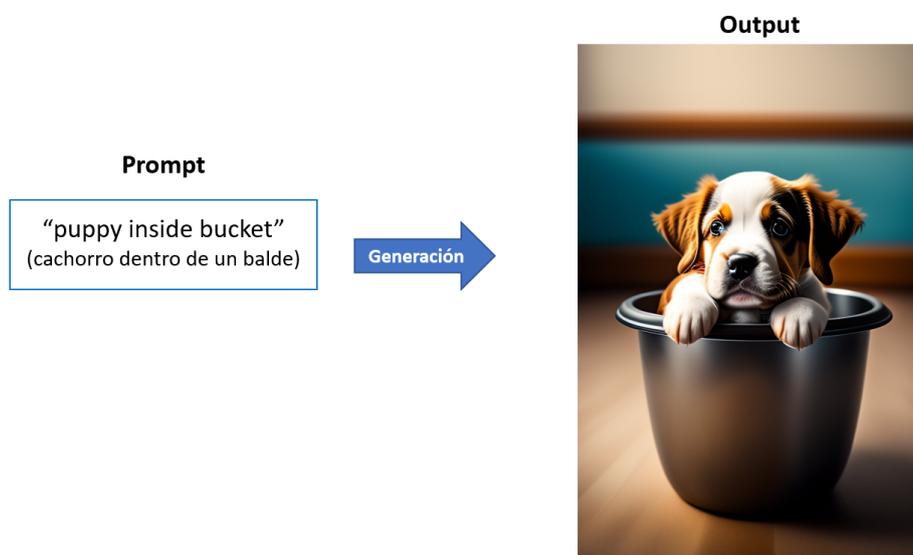


FIGURA 9 Ejemplo de proceso de creación de imágenes. Fuente: Lexic.art¹²

¹¹ Al respecto, es necesario señalar que, desde hace un tiempo, se vienen cuestionando las maneras en que se accedieron a los datos para los entrenamientos. Por ejemplo, en el caso de las IA generativas de imágenes, estas fueron entrenadas con millones de imágenes extraídas de Internet, sin consentimiento de sus autores. Esto ha dado lugar a una serie de causas judiciales que se vienen desarrollando, a la vez que diversas comunidades de artistas se han manifestado en contra de estas prácticas, en relación al origen y el uso de este tipo de inteligencias. En el caso de Argentina, la Asociación de Dibujantes de Argentina (ADA) viene llevando a cabo diversas actividades en pro de regular el uso de estos dispositivos a fin de garantizar un uso ético y que no atente contra la creatividad artística ni contra el campo laboral de la ilustración. Si les interesa este tema, pueden encontrar más información en la página web de la asociación: <https://ada.dibujantes.ar/>.

¹² <https://lexica.art/?prompt=9755dba7-beaf-4288-bfe4-e7327ad49a06> (consultado el 7/6/2023).

Para utilizar estas IA, se puede ingresar a sus respectivas páginas web y realizar una serie de pasos administrativos como registrarse, elegir una modalidad (gratuita, de pago, etcétera), aceptar las condiciones legales, entre otras. A partir de allí, el procedimiento para lograr estas imágenes es sencillo: el usuario debe realizar las instrucciones por escrito, es decir, el mentado *prompt* y la IA actuará. Al respecto, si bien estas pueden lograr una calidad elevada, los resultados muchas veces presentan diversas imperfecciones. Por un lado, se encuentran las limitaciones propias de los algoritmos. Un error frecuente en este sentido supo ser la generación de seres humanos con más de cinco dedos en una mano. Por otro lado, se observan limitaciones en los datos que derivan en sesgos (por ejemplo, responder a una idea de belleza, pobreza, etcétera hegemónicas). Finalmente, existen ciertas complejidades lingüísticas derivadas de la ambigüedad. Este último punto es lo que desarrollamos en la siguiente subsección.

4.2 | El chiste del genio de la lámpara (o las complejidades del *prompt*)

Recapitulando, el *prompt* constituye un input lingüístico que nos permite obtener distintos tipos de respuesta. Por lo general, se requiere que esté escrito en inglés, no obstante, se cuenta también con IA que permiten la escritura en otras lenguas. En este caso, vamos a ejemplificar con aquellas que permite la escritura de *prompts* en español.

Un problema que se observa al respecto es que, muchas veces, lo que el usuario *tiene en mente* no suele corresponderse con la imagen generada. Esto puede deberse a diversas razones: limitaciones propias del modelo (no contar con un corpus de datos de entrenamiento suficiente, por ejemplo), falta de especificidad en la escritura del *prompt*, sesgos producto de los datos con los que fue entrenado el algoritmo o, simplemente, interpretación particular de la IA. En relación con estas particularidades, uno de los casos más mediáticos fue una serie de imágenes de “un salmón nadando río abajo” (“a salmon swimming down a river”), como se muestra en la Figura 10:



FIGURA 10 Interpretación del *prompt* “a salmon swimming down a river”. Fuente Reddit¹³

Al respecto, surgen diversas preguntas: (i) ¿qué ocurrió aquí?; (ii) ¿se equivocó la IA interpretando el *prompt* ‘un salmón nadando río abajo’ (‘a salmon swimming down a river’)?; (iii) ¿El autor no escribió el *prompt* correctamente? La respuesta a la primera pregunta es sencilla, lo que ocurrió fue que un usuario dio una orden y la IA generó un output, en este caso, un (filé/filete de) salmón yendo por el río. Las respuestas a las otras dos son un poco más complejas, dado que ambas pueden ser contestadas tanto afirmativa como negativamente. Vamos a analizarlo con un poco de detalle.

En primer lugar, observemos las palabras que componen el input lingüístico, a los fines prácticos se presenta la traducción al español:

(6) Un salmón nadando río abajo.

En esta frase pueden observarse palabras con contenido léxico y con contenido gramatical. Para aclarar esto, a grandes rasgos, señalamos que existe un criterio para clasificar las palabras según estas sean léxicas o funcionales (o gramaticales). Las primeras se caracterizan por ser portadoras de un contenido que nos permite referir a cosas, eventos, estados o cualidades. Las funcionales, por su parte, son aquellas que no poseen contenido léxico, sino de tipo gramatical (género, número, etcétera) y tienen la función de combinar las palabras de contenido léxico para formar oraciones. Las palabras léxicas, precisamente, son las que brindan los elementos que van a estar representados (o no¹⁴) en la imagen que deseemos generar. Por lo tanto, en el *prompt* ‘un salmón nadando río abajo’, podemos discriminar ‘salmón’ y ‘nadando’, como palabras léxicas, y ‘un’ y ‘abajo’, palabras gramaticales.

La palabra ‘salmón’, es definida en su acepción de nombre por la RAE de la siguiente manera:

1.m. **Pez** teleósteo de hasta metro y medio de longitud, de cuerpo rollizo, cabeza apuntada y una aleta adiposa dorsal junto a la cola, de lomo azulado, vientre plateado, con reflejos irisados en los costados, que remonta los ríos para desovar, y **cuya carne, rojiza y sabrosa**, es muy apreciada¹⁵ (RAE, online, consulta 8/6/2023, el resaltado es mío).

Podemos ver dos significados de este término en las oraciones (7a) y (7b):

- (7) a. El salmón procrea en agua dulce.
b. Juan preparó una ensalada con salmón ahumado.

Aquí estamos nuevamente ante el fenómeno de la polisemia, pues vemos que la palabra *salmón* puede ser tanto animal como alimento. Otros ejemplos que sirven para ilustrar este caso son *pollo*, *conejo*, *cerdo*. A esto se lo llama polisemia regular y refiere al fenómeno que se produce cuando, en una determinada lengua, dos sentidos se diferencian regularmente de la misma manera. En este caso, la polisemia se da entre el animal y la carne del animal cuando esta se utiliza como alimento. También vemos polisemia regular en la referencia a un árbol y su madera (*algarrobo*, *pino*, *abeto*) o a nombres que remiten a ‘territorio’ y a ‘organización administrativa o de gobierno’, como en el caso de *califato* y *colonia*.

Veamos ahora el caso de *nadando*, el verbo *nadar* indica desplazarse por el agua de manera voluntaria (8a) o no y que puede afectar tanto a seres animados como inanimados (8b):

¹³https://www.reddit.com/r/Damnthatsinteresting/comments/yijd3c/they_asked_an_ai_engine_to_recreate_a_salmon/?utm_source=embedv2&utm_medium=post_embed&utm_content=post_body (consultado el 6/6/2023)

¹⁴Algunas aplicaciones tienen la opción de *prompts* negativos, en los que el usuario declara aquello que no desea que aparezca en la imagen, por ejemplo, puede pedir un bosque en el que no haya animales.

¹⁵<https://dle.rae.es/salm%C3%B3n>.

- (8) a. Tarzán nadó hasta la orilla.
b. El barquito de papel nadó por la canaleta hasta la alcantarilla.

La expresión ‘un salmón nadando río abajo’ puede significar al pez vivo desplazándose por el agua (para buscar alimento) o, y aquí apelando a la imaginación del lector, un trozo de carne de dicho animal yéndose por el río (imaginemos que justo pasaba un cocinero por el río, con la compra de la pescadería y la bolsa se le abre y se le caen los filetes de salmón al agua). Por lo tanto, podemos esquematizar las combinaciones de significado como se ve en la Figura 11.

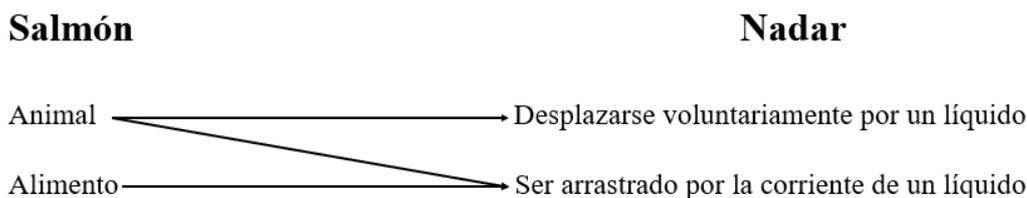


FIGURA 11 Combinaciones de los significados de *salmón* y *nadar*

Más allá de que para el humano resulte forzada la interpretación del algoritmo, recordemos que una IA es un dispositivo que se limita a hacer cálculos matemáticos, por lo que no va a tener en cuenta factores pragmáticos, ni de conocimiento de mundo. Desde ese punto de vista, ambas opciones son igualmente posibles.

Tanto el usuario como la IA recrearon el típico chiste del genio de la lámpara. Pues el usuario realizó un pedido y el algoritmo cumplió. Esto evidencia un problema clásico en el procesamiento del lenguaje natural que es la ambigüedad, fenómeno que se da cuando una expresión puede tener más de una interpretación posible. El lingüista Danilo Hidalgo menciona en un artículo de 1978 tres tipos principales de ambigüedad: (i) léxica; (ii) estructural patente, y (iii) estructural subyacente.

La ambigüedad léxica alude, como hemos estado viendo, a palabras con más de un sentido. Hidalgo recupera, entre otros ejemplos, la oración ‘Encontró la pluma’, en donde *pluma* puede aludir a lo que recubre la piel de las aves o al instrumento para escribir, entre otros significados posibles.

La ambigüedad estructural patente remite a aquella en las que las palabras no necesariamente son ambiguas, pero sí la construcción. Veamos el caso de (9), en el que podemos ver que *enojado* puede referirse al profesor o al estudiante:

- (9) El profesor encontró al estudiante enojado.

Por último, la ambigüedad estructural subyacente, que no está dada ni por el léxico ni por las posibilidades de agrupar, sino por la oración entendida como un todo. Un ejemplo clásico es el de (10):

- (10) El pollo está listo para comer.

Aquí vemos también dos significados posibles: (i) el pollo (comida, alimento) ya fue cocinado y está listo para ser ingerido o (ii) el pollo animal (vivo) tiene su alimento listo y puede comer.

Pues bien, para resolver estos tipos de ambigüedades, los hablantes de una lengua recurrimos a diversas inferencias, gracias al contexto lingüístico (palabras y oraciones que anteceden o proceden a la expresión ambigua) o extralingüístico (propio de la situación comunicativa). Para el caso de la escritura de *prompts* a fin de comunicarnos con una computadora, estas posibilidades están suspendidas, lo que provoca que, en varias ocasiones, sea necesario desambiguar agregando más información o pidiendo generar varias veces hasta lograr lo que se desea. A continuación, presento una serie de imágenes generadas a partir de *prompts* ambiguos por las IA Dall.E 2¹⁶ y Bluewillow AI¹⁷. Se eligieron entre las muchas que hay porque permiten la escritura de *prompts* en español.

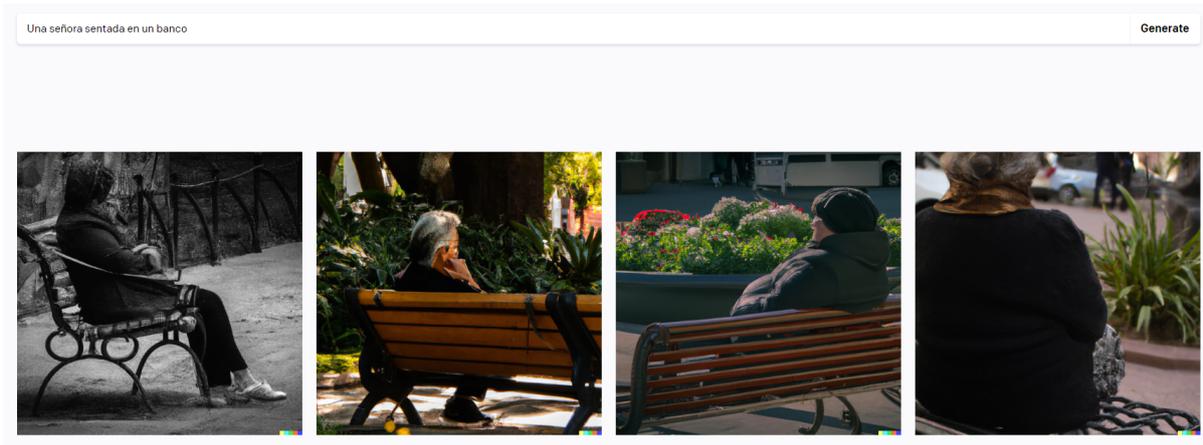


FIGURA 12 Output del *prompt*: 'Una señora sentada en un banco' (Dall-E-2) (consultado el 6/6/2023)

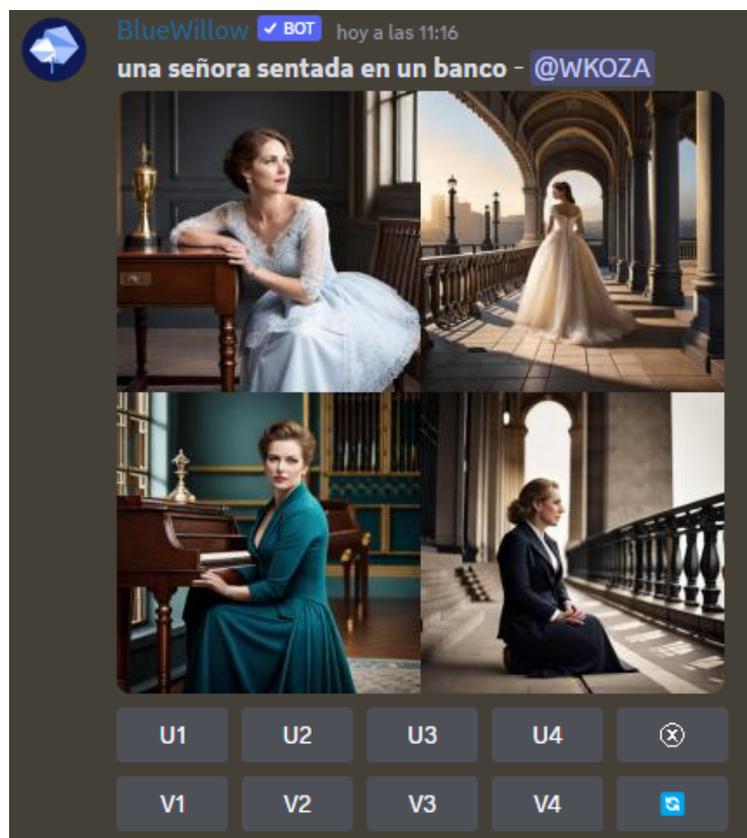


FIGURA 13 Output para 'una señora sentada en el banco' (Bluewillow) (consultado el 6/6/2023)

¹⁶<https://openai.com/blog/dall-e-2-extending-creativity>

¹⁷<https://www.bluewillow.ai/>



FIGURA 14 Primer grupo de imágenes generadas a partir del *prompt* ‘el pollo está listo para comer’ en Dall.E 2. (consultado el 6/6/2023)



FIGURA 15 Segundo grupo de imágenes generadas a partir del *prompt* ‘el pollo está listo para comer’ en Dall.E 2. (consultado el 6/6/2023)

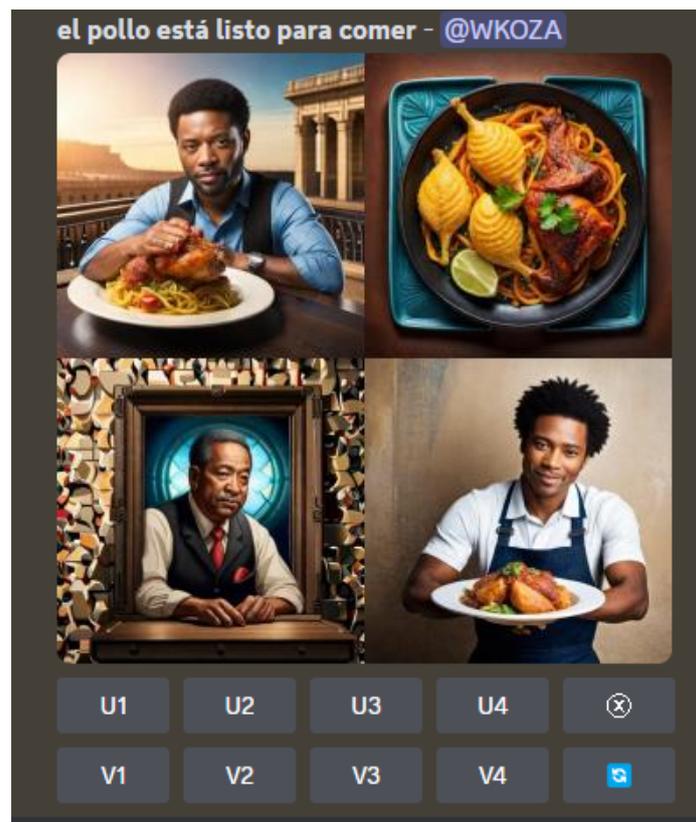


FIGURA 16 Imágenes generadas en Bluewillow AI a partir del *prompt* ‘el pollo está listo para comer’. (consultado el 6/6/2023)

En las Figuras 12 y 13 observamos que los algoritmos tomaron el sentido de ‘banco’ como elemento para sentarse, descartando alusiones a la institución bancaria. Si bien el análisis de los sesgos no forma parte del objetivo de este trabajo, cabe destacar que Dall.E 2 generó en la primera, tercera y cuarta, señoras, en apariencia, de mayor edad, respecto de las obtenidas por Bluewillow AI. Faltaría explicar por qué el primer dispositivo

generó un varón en la segunda imagen y, en el segundo, una mujer de pie.

En el caso de ‘el pollo está listo para comer’, se le solicitó a Dall.E 2 que generase dos veces, en los dos casos (Figuras 14 y 15), los resultados fueron similares.

Se obtuvo la misma interpretación del *prompt* en Bluewillow AI, como se puede apreciar en la Figura 16.

Ambas IA generaron imágenes tomando en consideración el significado de *pollo* como alimento. Entre los problemas que exceden al ámbito lingüístico se encontraría por qué Dall.E 2 genera la imagen de un pollo, aparentemente, crudo y por qué Bluewillow AI incluye en su output un hombre sentado a una mesa (posiblemente, listo para comer, es decir, para que le sirvan la comida).

5 | ALGUNAS REFLEXIONES FINALES

Las IA generadoras de contenido han tenido un gran impacto en la sociedad y es posible acceder a ellas con relativa facilidad. La popularidad de estos sistemas deriva de la calidad de lo generado, donde se pueden observar imágenes con alto grado de detalle y complejidad. No obstante, esto trae aparejado el riesgo de banalizar artes como la fotografía, la ilustración, la pintura, etcétera, dado que se puede caer en el error de suponer que el esfuerzo y el aprendizaje que suponen dominar estas disciplinas se evitarían gracias a estos algoritmos.

Otro aspecto a considerar es que, como se mencionó más arriba, para lograr esta clase de dispositivos, se debió recurrir a una gran masa de datos extraídos de internet, para desarrollar las técnicas de aprendizaje automático pertinentes. Con esto se evidencia un carácter parasitario en estas tecnologías, en la medida en que deben servirse de productos elaborados previamente. A tales efectos, algunos colectivos de artistas como la ADA, desde hace tiempo, vienen estudiando los impactos negativos que pueden surgir tanto para el campo profesional de la ilustración como para el arte en general, si las IA no son reguladas mediante un marco legal.

Por último, ya dijimos que las IA trabajan mediante instrucciones declaradas por el usuario. Ahora bien, el grado de satisfacción va a depender de qué tan compleja es la imagen que se quiere lograr y, por ende, del grado de explicitación que presente el *prompt*. Esto último implica un desafío para el usuario, dado que se trata de una situación comunicativa con un algoritmo computacional, que lo obliga a evitar posibles ambigüedades. Con motivo de esto, existen numerosos estudios que reflexionan sobre cuál de la manera más adecuada de *promptear*, a la vez que es posible acceder a varios sitios donde hay recomendaciones para la redacción de prompts como *prompthero*¹⁸, e incluso cursos, como el proporcionado por *promptingguide*¹⁹.

En relación con esto último, y más allá de que en este recorrido fuimos proponiendo ejemplos ad hoc, vale la pena hacerse una serie de preguntas respecto de hacia dónde derivará esta relación entre pensamiento, *prompt* y resultado de la IA. ¿Bastará con corpus de entrenamientos más robustos?, ¿habrá que desarrollar una nueva modalidad de escritura? o ¿es el sistema de la lengua el que terminará delimitando el terreno?

¹⁸<https://prompthero.com/>

¹⁹<https://www.promptingguide.ai/es>

ALGO DE LO QUE LEÍ SOBRE EL TEMA

| Polisemia:

Adelstein, Andreína, Marina Berri & Victoria Boschioli (2012). Polisemia regular y representación lexicográfica: los nombres locativos en español. *Terminalia*, 5, 33-41.

Apresjan, Juri (1974). Regular Polysemy. *Linguistics* 12(142), 5-32.

Hidalgo, Danilo (1978). La ambigüedad lingüística. *Filología y lingüística*, 4(2), 39-45.

Pustejovsky, James (1995). *The generative lexicon*. Cambridge-Massachusetts: MIT Press.

Real Academia Española & Asociación de Academias de la Lengua Española (2009). *Nueva gramática de la lengua española*. Madrid: Espasa.

| Inteligencia, inteligencia artificial, aprendizaje automático y cuestiones computacionales

Alpaydin, Ethem (2026). *Machine learning: The new AI*. Cambridge-Massachusetts: MIT Press.

Ardilla, Rubén (2011). Inteligencia. ¿Qué sabemos y qué nos falta por investigar? *Revista de la Academia Colombiana de Ciencias*, 35(134), 97-103.

Chowdhury, Mashrur, Amy Apon & Kakan Dey (2017). *Data Analytics for Intelligent Transportation Systems*. Amsterdam: Elsevier.

López, Bruno (2007). Introducción a la inteligencia artificial. Recuperado de: <http://www.itnuevolaredo.edu.mx/takeyas/Articulos/Inteligencia%20Artificial/ARTICULO%20Introduccion%20a%20la%20Inteligencia%20Artificial.pdf> el 6/6/2023.

Mantegna, Micaela (2020). *Inteligencia ARTEficial: creatividad computacional, inteligencia artificial generativa y derechos de autor*. Tesis de maestría, Universidad de San Andrés.

Noy, Shakked & Whitney Zhang (2023). Experimental evidence on the productivity effects of the generative artificial intelligence. Recuperado de https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4375283 el 6/6/2023.

OpenAI. (2023). ChatGPT: Optimizing language models for dialogue. <https://openai.com/blog/chatgpt/>

Sternberg, Robert (2000). *Handbook of intelligence*. Nueva York: Cambridge University Press.

Piaget, Jean (1952). *The origins of intelligence in children*. Nueva York: International Universities Press.

| Relación entre arte e IA y redacción de prompts

McCormack, Jon, Camilo Cruz, Nina Rajcic, Stephen Krol, María Llano & Meng Yang (2023). Is writing prompt really making art? *Lecture notes in computer science*, 13988, 196-211.

O'Hear, Anthony (1995). Art and technology: An old tension. *Royal institute of philosophy supplement*, 38, 143-158.

Zhou, Yongchao, Andrei Muresanu, Ziwen Han, Keiran Paster, Silviu Pitis, Harris Chan & Jimmy Ba (2023). Large language models are human-level prompt engineers. *Eleventh conference on learning representation*. Recuperado de <https://arxiv.org/pdf/2211.01910.pdf> el 6/6/2023.

| Sesgos y estereotipos

Bianchi, Federico, Pratyusha Kalluri, Esin Durmus, Faisal Ladhak, Myra Cheng, Debora Nozza, Tatsunori Hashimoto, Dan Jurafksy, James Zou, & Aylin Caliskan (2023). Easily accessible text-to-image generative amplifies demographic stereotypes at large scale. En *The 2023 ACM conference on Fairness, Accountability, and Transparency*. ACM FAccT 2023. Recuperado de <https://dl.acm.org/doi/pdf/10.1145/3593013.3594095> el 6/6/2023.