
Herramientas de generación de imágenes con inteligencia artificial

Este informe, tiene como meta compartir algunas herramientas de generación de imágenes con inteligencia artificial (IA) que pueden resultar útiles en actividades educativas, para el desarrollo de material docente, o como opción para que el alumno genere información visual.

Estas herramientas permiten crear imágenes a partir de texto o comandos, brindando opciones creativas y diversas.

BlueWillow: es un sistema de creación de imágenes por IA que, aunque no es muy conocido, se encuentra en fase beta y es totalmente gratuito. Sin embargo, no se sabe cuánto tiempo durará la fase beta ni si eventualmente se convertirá en un servicio de pago similar a Midjourney.

Para utilizar BlueWillow, es necesario conectarse a su servidor de Discord y enviar un comando para obtener los resultados deseados.

Sitio web: <https://www.bluewillow.ai/>

Canva: es una plataforma en línea que ofrece múltiples herramientas de diseño, incluido un generador de imágenes mediante IA. Para utilizarlo, simplemente se debe acceder a un nuevo proyecto o uno existente, y en la columna izquierda la opción "Text to image". Allí se podrá escribir el comando o texto descriptivo de lo que se desea y Canva generará cuatro imágenes en función de ello.

También se puede elegir el estilo y la relación de aspecto que se prefiera.

Sitio web: <https://www.canva.com/>

Craiyon (anteriormente Dall-E Mini): Este proyecto gratuito tiene como objetivo proporcionar una IA de calidad para generar imágenes a partir de texto. Ofrece una versión gratuita que genera imágenes ilimitadas con marca de agua, con tiempos de espera de 1 a 2 minutos.

También hay opciones de suscripción paga para obtener resultados más rápidos y sin marca de agua.

Sitio web: <https://www.craiyon.com/>

DALL-E 2: es uno de los sistemas de IA más conocidos para generar imágenes. Hay que registrarte en OpenAI para acceder a este servicio gratuito. En el sitio web de DALL-E, se puede encontrar ejemplos destacados creados por la comunidad. Solo se debe ingresar un comando o texto y en poco tiempo se obtienen cuatro resultados.

Tener en cuenta que algunas palabras o peticiones pueden estar censuradas según las normas de la comunidad.

Sitio web: <https://labs.openai.com/>

DeepAI: Esta sencilla página web permite crear imágenes a partir de texto utilizando algunos diseños gratuitos. Sin embargo, las mejores opciones están reservadas para suscriptores de pago. Algunos comandos pueden estar censurados si incluyen nombres o términos considerados dañinos. Además, cada vez que se genera una imagen, esta será diferente, pero existe la opción de mejorarla si no se está satisfecho con el resultado inicial. La resolución de las imágenes generadas no es alta, pero se ofrecen diferentes formatos.

Sitio web: <https://deepai.org/>

Deep Dream Generator: Desarrollado por Google, Deep Dream Generator es un sistema de IA gratuito que genera imágenes a partir de peticiones textuales. Para comenzar a utilizarlo, es necesario crear una cuenta de usuario. Ofrece la posibilidad de generar imágenes con distintos estilos.

Sitio web: <https://deepdreamgenerator.com/>

Dreamlike: es un sistema de creación de imágenes basado en Stable Diffusion con distintos modelos entrenados de maneras diferentes. Además del Stable Diffusion 1.5 básico, también ofrece modelos de fotorrealismo o anime para generar imágenes.

Sitio web: <https://dreamlike.art/create>

InstantArt: es una página web que permite generar imágenes mediante inteligencia artificial de forma totalmente gratuita. Su principal fortaleza radica en los 26 modelos de IA altamente afinados entre los cuales se puede elegir al momento de crear imágenes. Algunos de estos modelos se basan en Midjourney o Stable Diffusion, mientras que otros se basan en diferentes sistemas.

Sitio web: <https://instantart.io/>

Mage.space: es otro generador de imágenes en línea donde se puede utilizar modelos como Stable Diffusion, Analog y muchos otros. También existe la opción de seleccionar la cantidad de pasos, el ratio de aspecto y la diferencia del dibujo en relación al comando que fue ingresado.

Sitio web: <https://www.mage.space/>

NightCafé: es una aplicación web progresiva que se puede utilizar desde cualquier dispositivo de forma gratuita. Sus creadores buscan democratizar el uso de sistemas de inteligencia artificial como el de Stable Diffusion. Esta IA funciona a través de créditos: si se genera solo una imagen por comando, no se gasta ningún crédito, pero si se producen varias, se consumirán. Sin embargo, está la opción de adquirir más créditos si así se desea. Además del comando, NightCafé permite elegir entre diferentes tipos de diseño que influirán significativamente en el resultado final. Sitio web: <https://creator.nightcafe.studio/>

Stable Diffusion: es una inteligencia artificial de código abierto que permite crear imágenes a partir de texto. Con esta herramienta, se puede controlar aspectos como la separación entre el comando que se ha dado y lo que genera la IA, así como la cantidad de pasos para generar el dibujo. La disponibilidad del código abierto ha permitido que la comunidad lo adopte y lo lleve a múltiples plataformas. Incluso es posible entrenar el modelo con datos propios, como imágenes del rostro, para crear imágenes personalizadas. Además, cuenta con una web oficial llamada DreamStudio donde se puede utilizar de forma gratuita. Sitio web: beta.dreamstudio.ai

Midjourney: es un destacado sistema de IA para generar imágenes y posiblemente uno de los mejores disponibles. Funciona a través de Discord y se puede utilizar en su servidor oficial o invitando a su bot a tu propio servidor. En este momento es una de las plataformas más utilizada para el desarrollo de imágenes. Es necesario suscribirse por un costo al mes para continuar generando imágenes. Sitio web: <https://www.midjourney.com/>

Estas herramientas ofrecen diversas opciones para generar imágenes atractivas utilizando inteligencia artificial. Cada una tiene sus propias características y enfoques, por lo que recomendamos explorarlas y probarlas para descubrir cuál se ajusta mejor a sus necesidades y preferencias.

Hay que tener en cuenta que algunas herramientas pueden ofrecer funciones adicionales o versiones premium que requieren suscripción o pago.

Midjourney

MidJourney es un proyecto independiente dedicado a la investigación de la Inteligencia Artificial, y también es el nombre que han dado a su IA especializada en la creación de imágenes a partir de texto, conocido como "text to image".

A diferencia de otras alternativas como DALL-E 2, MidJourney se destaca por crear lienzos con imágenes bien estructuradas y definidas, ofreciendo una gran calidad. Se asemeja más a sistemas como Stable Diffusion y es reconocida por su realismo. Una de las principales ventajas de MidJourney es su capacidad para producir imágenes de alta resolución. De forma predeterminada, puede generar imágenes de hasta 1.792 x 1.024 píxeles, lo que le permite ofrecer un mayor espacio y generar imágenes más grandes.

La modalidad de MidJourney sigue un modelo freemium, lo que significa que ofrece una funcionalidad limitada de forma gratuita, pero también brinda opciones de pago con características adicionales, mayor potencia y capacidad de creación.

En cuanto a su funcionamiento, actualmente MidJourney solo puede ser utilizado a través de un bot en el servidor oficial de Discord del proyecto. Los usuarios interactúan directamente con el bot o lo invitan a sus propios servidores o a servidores de terceros. Todo el proceso se realiza mediante el envío de comandos seguidos del texto descriptivo de lo que se desea visualizar.

Cuando se ingresa un texto descriptivo, la IA de MidJourney primero interpreta el significado de las palabras proporcionadas y luego genera una imagen basada en ese contexto, utilizando el conocimiento y entrenamiento previo que sus creadores le han proporcionado.

Por ejemplo, la IA ha sido entrenada para reconocer que la palabra "coche" se refiere a un vehículo y tiene una forma y características asociadas.

A diferencia de la búsqueda tradicional basada en palabras clave, utilizar MidJourney implica un enfoque opuesto. En lugar de buscar imágenes existentes, se proporciona una descripción textual para que la IA genere una imagen única y personalizada basada en esa descripción.

Perro volando

Perro de la raza pastor alemán, volando en un cielo azul con algunas nubes esponjosas, el perro no tiene alas, pero sí posa como un pájaro, el sol incide perpendicular para hacer la imagen brillante. Tiene una expresión feliz e intrépida en su rostro, mira hacia el horizonte.

MidJourney ha ganado reconocimiento por su capacidad para crear imágenes detalladas y realistas a partir de texto, y su enfoque basado en Discord ha facilitado su accesibilidad y uso por parte de la comunidad interesada en la generación de imágenes mediante inteligencia artificial.

Aunque la interfaz de MidJourney esté en inglés, la herramienta entiende prácticamente cualquier idioma, pero eso no significa que los resultados que ofrece sean de la misma calidad en uno que en otro.

Las características de los prompts en MidJourney son fundamentales para obtener los resultados deseados en la generación de imágenes.

A continuación, se explican las principales características y cómo utilizarlas:

Proporción: al realizar una solicitud a MidJourney, se generan cuatro imágenes cuadradas por defecto. Sin embargo, se puede especificar si se desea imágenes apaisadas o verticales utilizando el comando "--ar" (aspect ratio) seguido de la relación de aspecto deseada, como "--ar 3:2" o "--ar 16:9". Esto te permite controlar la proporción de las imágenes generadas.

Estilo: MidJourney ofrece una amplia variedad de estilos e iluminaciones para sus imágenes. Se puede especificar el tipo de estilo deseado en tu prompt para obtener resultados acordes a las preferencias de estilo. Permite experimentar con términos relacionados con lentes, óptica, iluminación, efectos visuales y más. Si no se está familiarizado con suficientes términos, se puede buscar vocabulario relacionado para lograr el efecto deseado. MidJourney es lo suficientemente inteligente como para comprender y no generar resultados indeseados cuando se utilizan términos específicos.

Ilustraciones: MidJourney no se limita a generar imágenes realistas, sino que también es capaz de crear ilustraciones en diversos estilos, como arte pictórico, ilustración, cómic, manga, anime, e incluso puede recrear estilos de personajes de animación. Al especificar el estilo deseado en el prompt, MidJourney se adaptará para generar imágenes acordes a esa estética.

Resolución: además de controlar la proporción de la imagen, también se puede especificar el tamaño exacto de la imagen utilizando los comandos "--w" para la anchura y "--h" para la altura. Por ejemplo, si se desea una imagen de 1000 píxeles de ancho y 700 píxeles de alto, puedes agregar "--w 1000 --h 700" al final del prompt.

Calidad: MidJourney ofrece opciones de control de calidad que determinan el tiempo necesario para generar una imagen. Se ajusta la calidad de la imagen utilizando los comandos "--quality" seguidos de valores como ".25", ".5", "1" o "2".

Un valor más alto implica mayor tiempo empleado en generar la imagen y, en general, resultará en una mejor calidad. Por defecto, si no se especifica ningún comando de calidad, se aplicará "--quality 1".

Múltiples imágenes: Partir de una imagen previa o mezclar dos, si se desea que una imagen se base en una imagen previa, se debe incluir su URL como parte del prompt. MidJourney utilizará esa imagen como base para generar la nueva imagen, aunque no necesariamente la reinterpretará de manera fiel. También es posible combinar dos imágenes diferentes para crear un híbrido de ambas. En este caso, se deben incluir las URL de ambas imágenes. Tener en cuenta que la complejidad de la mezcla puede afectar el resultado final.

Evitar presencias habituales: si hay elementos que se desea evitar en las imágenes generadas, se puede utilizar el comando "--no" seguido de la palabra clave para indicar lo que no se desea que aparezca. Por ejemplo, "--no text" para excluir texto de las imágenes generadas.

Caos: MidJourney brinda una opción que permite introducir "caos" en las imágenes generadas. Esta característica otorga una mayor libertad creativa y la posibilidad de explorar inventivas artificiales a partir del prompt. Al utilizar MidJourney sin especificar este parámetro, el valor predeterminado es cero, pero se puede aumentar el nivel de caos, utilizando el comando "--chaos" o "--c" seguido de un número del 0 al 100. Cuanto mayor sea el número, más "caótica" será la imagen generada.

Además, MidJourney ofrece una lista de palabras clave que se pueden incluir en el prompt para obtener resultados precisos en una dirección concreta. Estas palabras clave están relacionadas con estilos artísticos, técnicas fotográficas y efectos visuales.

Al agregar estas palabras clave en el prompt, MidJourney tomará en cuenta las indicaciones y generará imágenes acordes a los estilos y características solicitados. Algunos ejemplos de estas palabras clave incluyen "Aerial photography" para obtener vistas cenitales de paisajes o ciudades, "Astrophotography" para paisajes nocturnos con estrellas y nebulosas, "Cinematic" para lograr un aspecto cinematográfico, "Closeup shot" para generar primeros planos realistas, "Chromatic aberration" para

simular una distorsión fotográfica, "Cyberpunk" para recrear paisajes urbanos con iluminaciones y neones, y muchas más que se detallan a continuación.

Aerial photography: para ver un paisaje, una ciudad, etc. desde una vista cenital

Astrophotography: paisajes nocturnos que incluyan estrellas, nebulosas, etc. en el cielo.

Cinematic: aspecto cinematográfico

Closeup shot: genera primeros planos para imágenes realistas

Chromatic aberration: simula esta distorsión fotográfica

Cyberpunk: estilo de este subgénero de la ciencia ficción (paisajes artificiales, iluminaciones urbanas nocturnas, neones...)

Depth of field: profundidad de campo, fondo desenfocado

Diffraction grating: red de difracción, mejora las iluminaciones

Diffuse back light: simula la retroiluminación artificial mediante un foco difuso

Expressionism: simula arte expresionista

Fauvism: estilo artístico que usa el color de forma provocativa

Glitch art: estética de arte tradicional "roto" de forma digital

Glowing: crea una imagen brillante

HDR: alto rango dinámico

Hyper realistic: genera imágenes ultra realistas

Intricately-detailed: gran nivel de detalle

Lo-fi photography: fotografías hechas a propósito con aspecto de mala calidad, pero con el tipo de estética buscada (distorsionadas, borrosas, con fugas de luz, viñetas...)

Oil painting: genera imágenes que simulan pinturas al óleo

Pixel art: un poco de nostalgia por lo retro

Pop art: imágenes de este estilo artístico

Ray-traced / Ray tracing: empleo del algoritmo especializado en la generación de iluminaciones y sombras

Realistic lense flare: genera un destello de lente realista, como en una fotografía

Rococo: imágenes que representen este movimiento artístico

Scattering: dispersión, un efecto popular para ciertas fotografías o composiciones

Screen space reflections: técnica para calcular de forma óptima los reflejos sobre una superficie

SSAO: técnica de oclusión ambiental para mejorar las zonas sombreadas

Sunset / Sunset lightning: recrea la luz de un atardecer en la imagen

Ukiyo-e: genera imágenes que simulan esta técnica de estampación japonesa

Vintage: imágenes en sepia, de baja calidad, emulando fotografías muy antiguas

Watercolor: emula un cuadro pintado con acuarelas

X architecture (siendo 'X' un tipo de arquitectura): genera imágenes con este estilo arquitectónico. Por ejemplo, 'stalinist architecture', 'scandinavian architecture', 'traditional japanese architecture' o 'minimalist architecture'.

XXXXs fashion (siendo 'XXXX' un número): genera imágenes donde las personas aparecen vestidas y peinadas con la moda de la década que le indiquemos. Por ejemplo, '1970s fashion'.

XX mm lens (siendo 'XX' un número): genera imágenes que recrean fotos con una lente de la distancia focal que le indiquemos. Por ejemplo, '55 mm lens'.

Con estas características y palabras clave, MidJourney amplía las posibilidades creativas. Al redactar prompts adecuados y utilizar términos precisos, se aprovecha al máximo esta herramienta y se pueden explorar un sinnúmero de posibilidades visuales. Es importante pensar de forma amplia y experimentar con diferentes combinaciones para obtener diferentes resultados visuales. El tener conocimientos en ilustración, arte y fotografía genera un diferencial al utilizar esta herramienta.

Referencias para la generación de la información

<https://www.xataka.com/basics/19-paginas-servicios-para-crear-imagenes-cero-utilizando-inteligencia-artificial>

<https://www.xataka.com/basics/que-midjourney>

<https://www.xataka.com/robotica-e-ia/guia-practica-para-escribir-mejores-prompts-midjourney-desbloquear-su-verdadero-potencial>
