



Quién soy yo

- Senior Data Scientist en Cognizant
- Especializado en Deep Learning y Computer Vision
- [LinkedIn](#)
- [GitHub](#)



Índice de la sesión

- Deep Learning para Texto a Imagen
- Modelos de Texto a Imagen disponibles
- Stable Diffusion
- Texto a ...
- Práctica

1. Deep Learning para Texto a Imagen

Redes Neuronales entrenadas para comprender la idea de un texto y generar una imagen que sea representativa de esa idea.

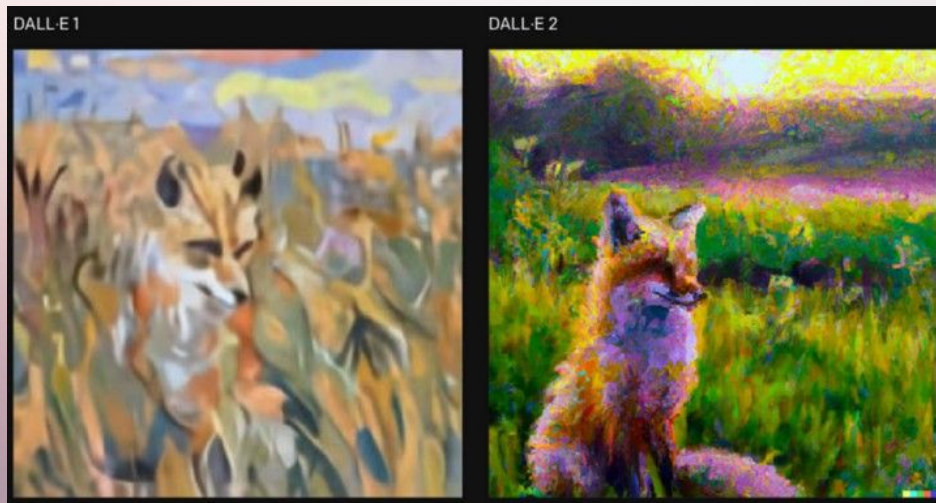
StackGAN en 2017 fue un gran hito para este campo:



2. Deep Learning para Texto a Imagen

En Enero de 2021, OpenAI sacó su primera versión del modelo DALL·E.

En Abril de 2022, OpenAI sacó su segunda versión del modelo DALL·E.



3. Deep Learning para Texto a Imagen

En Mayo de 2022, Google sacó su modelo de texto a imagen llamado: Imagen.

Mejor calidad y capacidad para escribir textos.



DallE tiene problemas para escribir "Deep Learning" en un cartel. Imagen no tiene problemas escribiendo "Text to Image" en una tienda.

4. Deep Learning para Texto a Imagen

En Junio de 2022, Google sacó Parti. Diferente arquitectura + muchos más parámetros = mejor calidad



D. A giant cobra snake made from X . $X \in \{\text{"salad", "pancakes", "sushi", "corn"}\}$

E. A wombat sits in a yellow beach chair, while sipping a martini that is on his laptop keyboard. The wombat is wearing a white panama hat and a floral Hawaiian shirt. Out-of-focus palm trees in the background. DSLR photograph. Wide-angle view.

F. The saying "BE EXCELLENT TO EACH OTHER" ..., (a) brick wall and alien (b) driftwood. (c) old wooden boat with reflection. (d) stained glass. (See text for full prompts.)

5. Deep Learning para Texto a Imagen

Estos modelos poseen en su mayoría una tendencia clara que comparte con el país donde han sido creados.

Es un problema difícil de evitar ya que siempre que hay un punto de referencia, va existir una tendencia que parta de ahí.

El modelo siempre refleja la tendencia de los datos que se le dan. Nada más.

an attractive person



an emotional person



an exotic person



a poor person



a terrorist



a thug



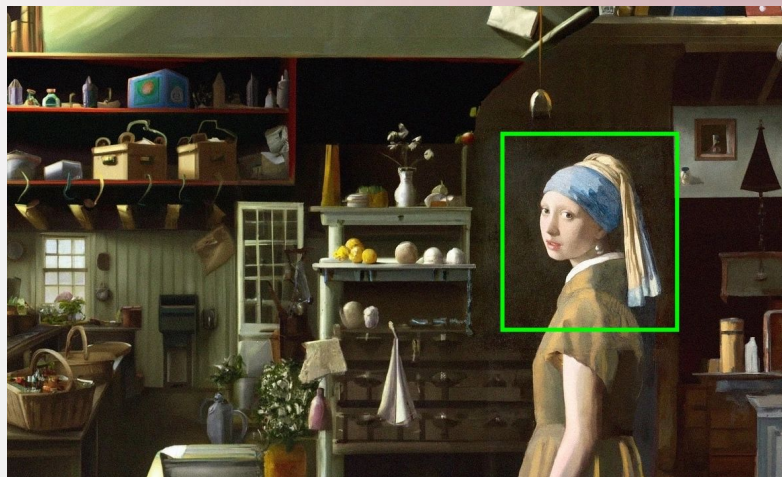
6. Modelos disponibles

Modelos que usan algunas de las arquitecturas anteriores y que podemos usar fácilmente.

Algunos son 100% de pago, otros dejan probarlo gratis y otros son 100% gratis.

7. Modelos disponibles

DALL-E 2 es 100% de pago sin prueba gratuita.
115 imágenes por 15\$.
Permite crear, modificar y rellenar imágenes.



8. Modelos disponibles

Midjourney es de pago pero ofrece unas demos gratis. Su uso es de Bot de Discord. Desde 10\$ por 200 imágenes. Ha sido entrenado conteniendo artistas digitales famosos. Puede crear y aumentar la resolución de sus creaciones.



9. Modelos disponibles

Stable Diffusion es 100% gratis y de código abierto. Los modelos se pueden descargar y usar de forma personal o usarlos online.



10. Stable Diffusion

Al ser lanzado con código abierto, la comunidad ha creado herramientas de todo tipo:

- **Lexica:** Es una plataforma donde los usuarios suben generaciones junto con el texto usado.
- **DreamStudio:** Usar el modelo online junto con la capacidad de editar imágenes usando texto.
- **PromptMania:** Permite crear textos para generar buenas imágenes.
- **Photoshop:** Hay implementaciones para usarlo como una herramienta de edición dentro de Photoshop.

11. Texto a ...

Existen otros modelos de generación sintética de datos a partir de texto.

- **Audio:** Meta creó un modelo llamado AudioGen, capaz de generar audio a partir de texto.
- **Video:** Google creó Phenaki, un modelo de generación de video a partir de texto. Este texto puede ir cambiando y siguiendo un guión y la generación de video se adapta de forma realista:

<https://phenaki.video/>

12. Práctica

Generar imágenes a partir de texto usando Stable Diffusion.

[LINK](#)



HUGGING FACE



PyTorch

¡Muchas gracias por
uniros a esta W3 Drop!