

# PROPUESTA DE SOLUCIÓN AL RETO: ARCA CONTINENTAL

---

**A00828513, Alejandro Murcia Alfaro**

**A00831568, Ana Daniela López Dávila**

**A01177643, Andrea Galicia Jimenez**

**A01412476, Iván Leobardo Hernández Buda**

**A01367547, Niza Alvarado Rendón**

**A010125281, Renata Vargas Caballero**

## INTRODUCCIÓN

El socio formador, Arca Continental, nos plantea el desafío de crear una propuesta que apoye al departamento de marketing de la empresa de la siguiente manera: Crear un programa que desarrolle sketches de pósters para promocionar sus artículos (incluyendo promociones, descuentos, etcétera) con modelos generativos de imágenes.

### Tecnologías open-source de generación de texto

Las tecnologías open-source de generación de texto se refieren a herramientas y software de inteligencia artificial que tienen la capacidad de generar contenido escrito, como texto, artículos, descripciones, resúmenes y otros tipos de texto a partir de datos de entrada o instrucciones. Lo que las distingue es que son de código abierto, lo que significa que su código fuente está disponible públicamente, permitiendo a cualquier persona ver, modificar y distribuir el software de forma gratuita.

Un ejemplo de estos modelos es el siguiente:

### 1. LLaMa2

El modelo lingüístico Llama 2 representa un notable avance en la tecnología de procesamiento de lenguaje natural de Meta AI. En comparación con su predecesor, Llama 1, este nuevo modelo es considerablemente más grande y potente, ofreciendo un aumento del 40% en su rendimiento y en el tamaño de los datos con los que trabaja.

Llama 2 se presenta en tres tamaños de parámetros distintos: 7B, 13B y 70B, lo que permite a los usuarios elegir la escala de modelo que mejor se adapte a sus necesidades específicas. La disponibilidad de estas diferentes configuraciones facilita la adaptación de Llama 2 a una amplia gama de aplicaciones y requisitos.

Un aspecto especialmente destacado de Llama 2 es su compromiso con el código abierto. Esto significa que el código fuente del modelo está disponible públicamente, lo que permite a investigadores y desarrolladores explorar, modificar y personalizar el modelo de manera libre y creativa. Además, Meta AI ofrece Llama 2 de forma gratuita tanto para fines de investigación como para su uso en aplicaciones comerciales. [2]

## Tecnologías open-source de generación de imágenes

Las tecnologías open-source de generación de imágenes son herramientas de inteligencia artificial que generan imágenes realistas a partir de texto o datos. Son de código abierto, lo que significa que su código fuente es público y pueden ser utilizadas, modificadas y compartidas libremente. Estas tecnologías son ampliamente utilizadas en campos como el arte generativo y la investigación científica debido a su accesibilidad y capacidad de personalización.

Algunos ejemplos de estas tecnologías que se tienen pensadas a utilizar en este reto son los siguientes:

### 1. DALL-E 2

DALL-E 2 es un generador de arte de IA de texto a imagen basado en el aprendizaje automático que ha causado sensación en la web. Esta innovadora herramienta es obra de OpenAI, una empresa líder en inteligencia artificial. DALL-E 2 opera en el ámbito de la generación artística, lo que implica su capacidad de crear imágenes desde cero, así como de generar ediciones y variantes de obras preexistentes. Es importante destacar que este sistema no "comprende" de manera literal lo que crea, sino que realiza suposiciones basadas en una extensa base de datos

que comprende 650 millones de combinaciones de imágenes y descripciones que ha procesado previamente. [1] Canva proporciona acceso gratuito al uso de esta herramienta en donde cada usuario tiene un límite de 30 prompts de manera mensual. Cada prompt resulta en dos imágenes, lo que nos habilita un gran repertorio para experimentar los parámetros y la manera en la que funciona su algoritmo.

Las potenciales complicaciones que se pueden presentar en este modelo son fallos en la generación de imágenes en elementos específicos como formas de humanos o texto exactamente replicado. A pesar de eso, DALL-E 2 cuenta con licencias de comercialización completamente abiertas: *“Sujeto al cumplimiento de estos términos y nuestra Política de Contenido, puedes utilizar “Generations” para cualquier propósito legal, incluido el uso comercial. Esto significa que puedes vender los derechos sobre las generaciones (imágenes) que crees, incorporarlos a obras como libros, sitios web y presentaciones, y comercializarlos de otra manera”*.

## 2. Stable-Diffusion

El modelo "Stable Diffusion" representa una nueva frontera en el ámbito de los sistemas generativos, y en particular, en la conversión de texto a imagen. Su esencia radica en utilizar la difusión como medio para simular la generación de imágenes, tomando inspiración de cómo las sustancias se propagan en un medio a lo largo del tiempo. La difusión en el aprendizaje profundo intenta emular el proceso de corrupción y restauración de una imagen. Imagina que tienes una fotografía clara y la sometes a un proceso donde se va degradando hasta convertirse en puro ruido visual.

El modelo intenta hacer lo contrario: a partir de ese ruido, busca restaurarla a su estado original. Durante el entrenamiento, el modelo aprende cómo realizar este proceso de restauración de la manera más precisa posible. Para generar imágenes a partir de texto, el "Stable Diffusion" no opera solo. Necesita comprender primero el texto proporcionado, para lo cual se apoya en modelos de procesamiento de lenguaje natural avanzados, que traducen las palabras en una representación numérica. Esta representación es la que alimenta al modelo de difusión, proporcionando un marco sobre cómo debería lucir la imagen final basada en la descripción textual.

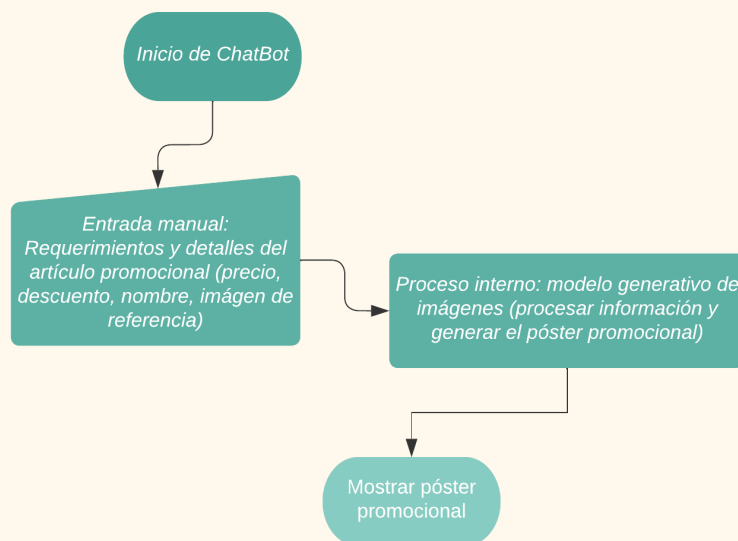
El proceso de entrenamiento es crucial. Se necesita un conjunto de datos vasto y variado que conste de descripciones textuales junto con sus imágenes correspondientes. Al exponer al modelo a numerosos ejemplos, este ajusta y adapta sus parámetros internos para lograr una correlación más precisa entre el texto y la imagen. Así, con cada ciclo de entrenamiento, se

vuelven más adeptos a generar visualizaciones que se alinean estrechamente con las descripciones dadas.

Este modelo es sumamente robusto, ya que combina la potencia del procesamiento de lenguaje natural con técnicas avanzadas de generación de imágenes para convertir descripciones textuales en visualizaciones detalladas y precisas. Aunque su implementación puede parecer desafiante al principio, el potencial y las aplicaciones de este modelo en campos como la generación de contenido, diseño y animación son verdaderamente prometedoras.

## PROPUESTA

### Diagrama de flujo



Nuestra propuesta se basa en la creación de un Chatbot donde se pueda introducir información relevante de acuerdo al artículo promocional (precio, descuento, nombre del artículo, imagen del artículo o artículos) como input, y el póster promocional relativo al artículo como output. Como se muestra en los siguientes pasos:

**Paso 1:** Hacer un prompt. Ejemplo: “Haz un promocional sobre {object} de {feature 1, feature 2, feature 3} considerando lo siguiente: {promotion}”/ “Haz un anuncio publicitario sobre

coca-cola de 600ml considerando lo siguiente: 15% de descuento en la compra de unas donitas bimbo de 80g”.

**Paso 2 (opcional):** Cargar imágenes de los productos a usar en el promocional.

**Paso 3:** Copiar el diseño proporcionado por el programa.

## REFERENCIAS

1. Yadav, A. (2023, March 2). *Cómo usar DALL-E 2 para crear increíble arte con inteligencia artificial*. Blog Español | Simplified; Blog Español | Simplified.  
<https://simplified.com/es-blog/ia-generadora-de-imagenes/como-usar-dall-e/#:~:text=DA LL%2DE%202%20es%20un,ha%20provocado%20asombro%20en%20internet>.
2. Ertugrul Portakal. (2023, July 31). *How to Access and Use Llama 2?* Textcortex.com.  
<https://textcortex.com/es/post/how-to-access-and-use-llama-2>