



**Data Science**  
Modelos  
generativos



StyleGAN2 (Karras et al.)

<https://thispersondoesnotexist.com/>

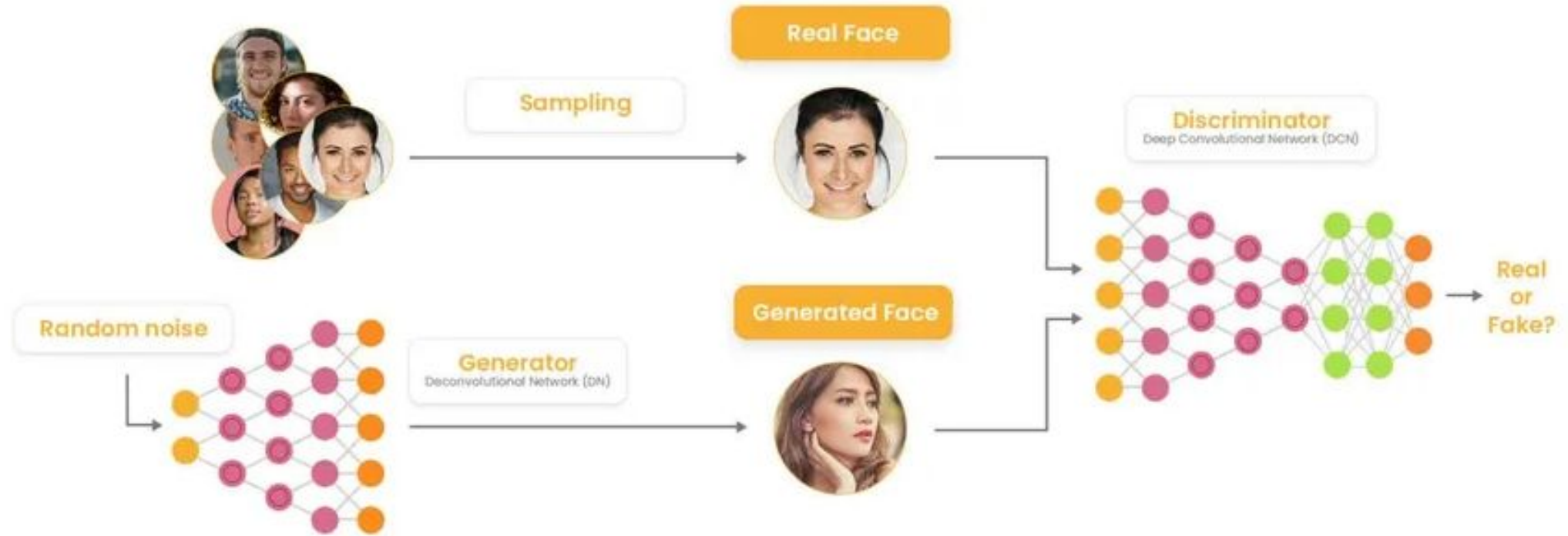
Los modelos generativos nos permiten **crear muestras** que no existen pero que resultan plausibles.

Además, podemos realizar una generación **condicionada**. Es decir, que basado en las indicaciones que le demos, la muestra generada se acota a un subconjunto.



*Ardilla que Picasso nunca pintó*

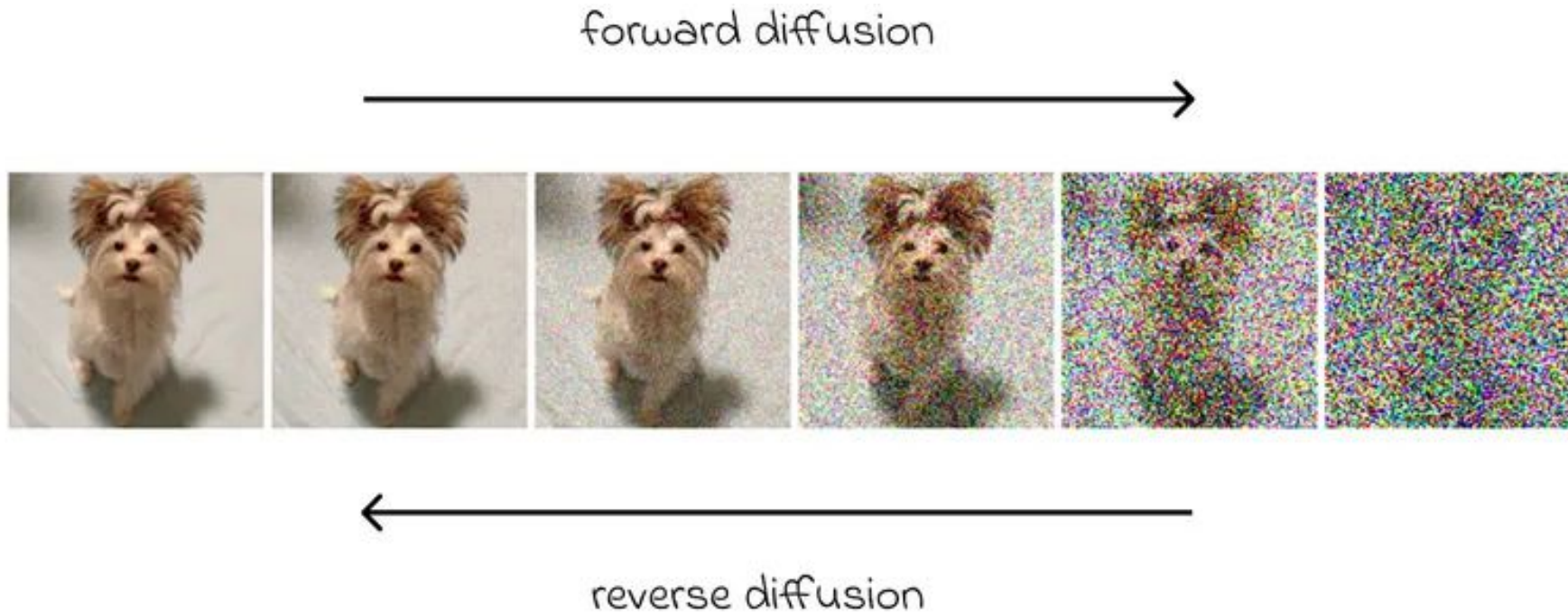
Una de las nuevas aproximaciones a este reto fueron las redes neuronales adversariales generativas o GANs (Goodfellow 2014) que nos permitían entrenar modelos capaces de generar muestras realistas.



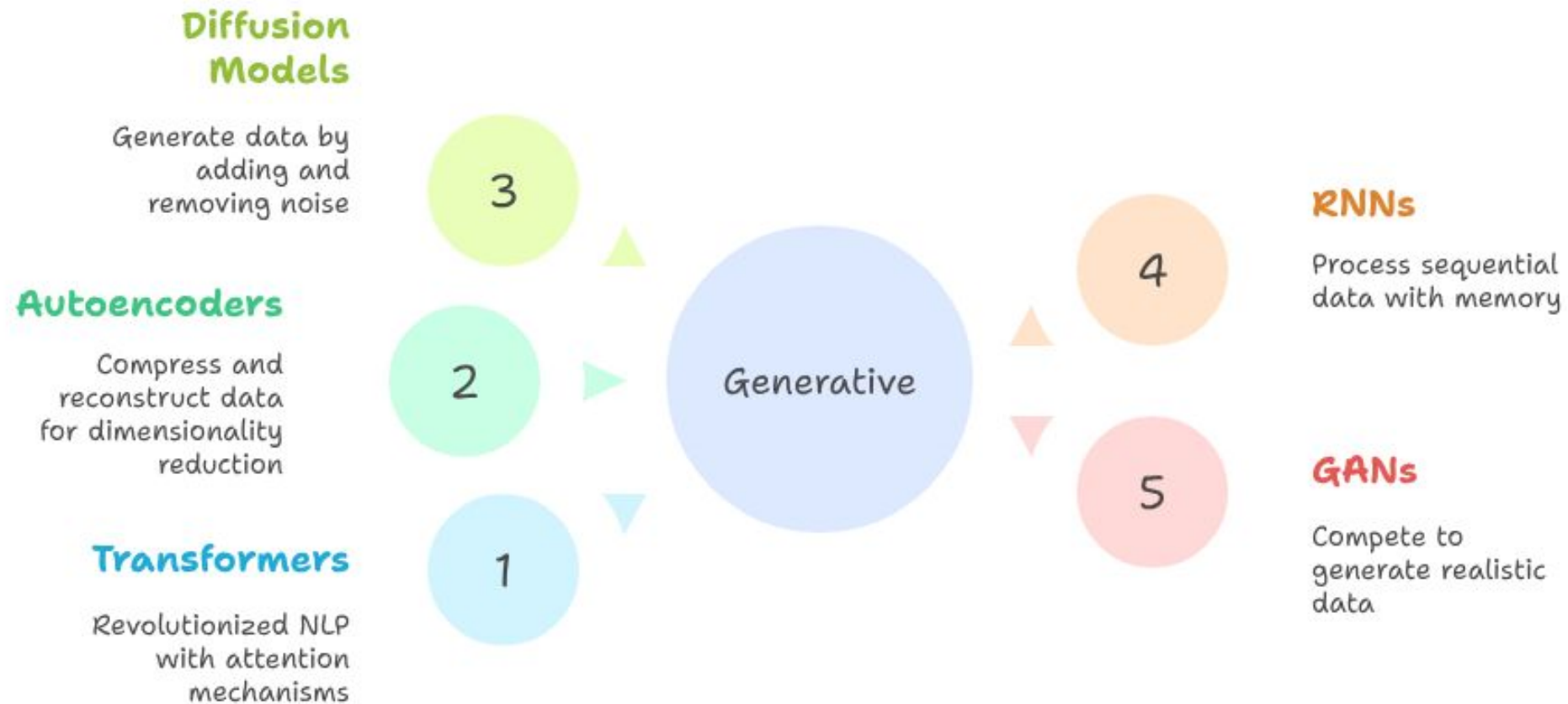


Pero hemos encontrado técnicas específicas para cada caso. Cada tipo de datos (imagen, audio o texto, por ejemplo) debemos tratarlo de formas concretas para poder sacarle el máximo provecho.

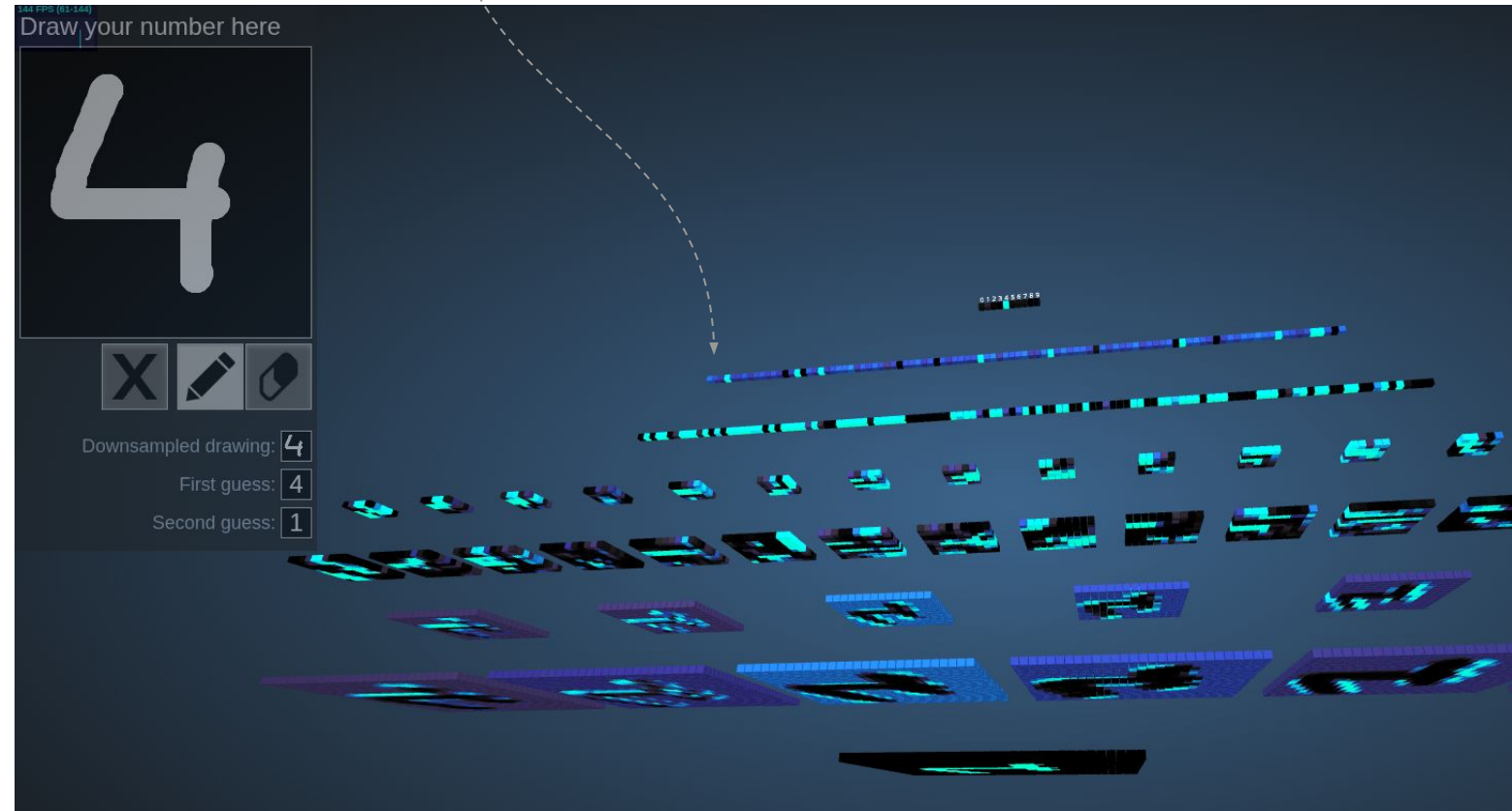
Los modelos de difusión aprenden a generar imágenes definidas de ruido.



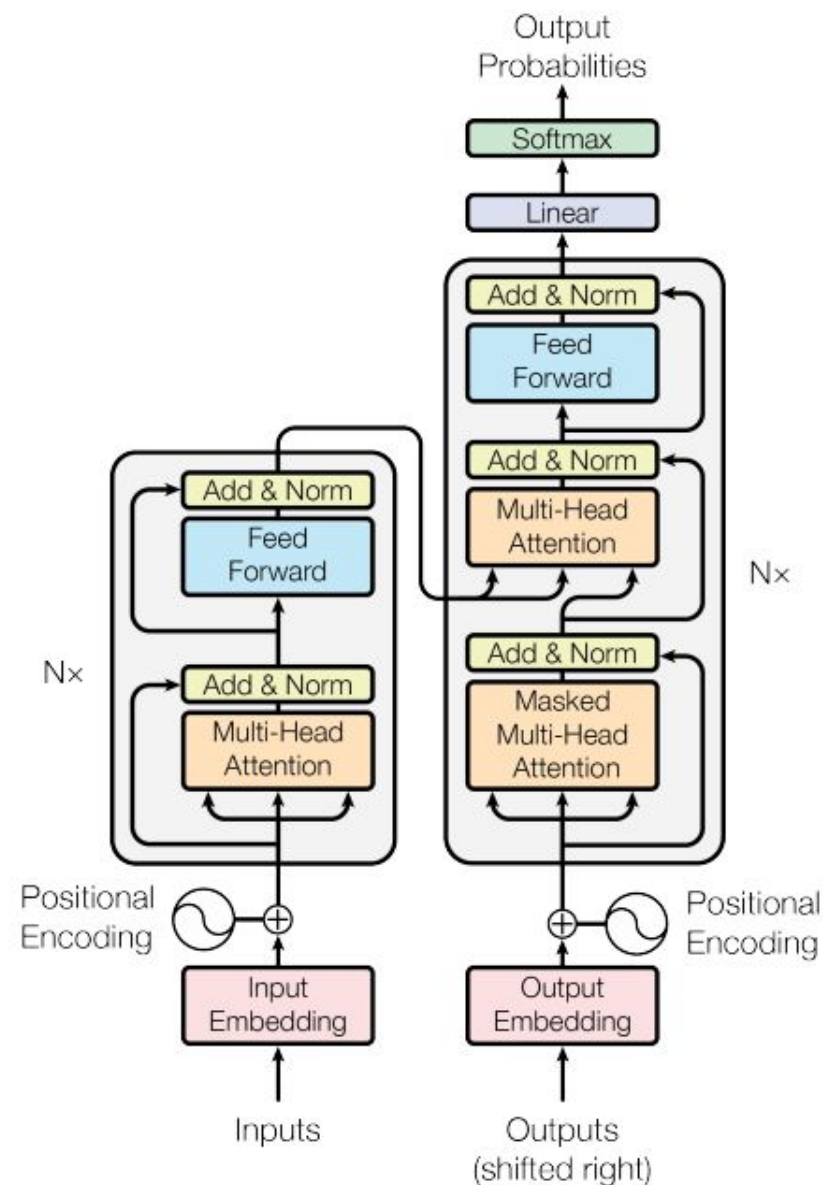
No existe solo un tipo de modelo generativo, si no que tenemos distintas opciones dependiendo del dato en cuestión que nos ocupe. Y lo bueno es que hemos encontrado el medio para **combinarlos**.



El gran salto lo hemos dado gracias a los **embeddings** y la capacidad de representar aspectos complejos como pudiera ser el lenguaje o las imágenes como un vector de características con el que operar.



Los **transformers**, gracias al mecanismo de **atención**, permiten generar un texto coherente y muy humano... aunque no siempre fiable.



Google



Buscar con Google

Voy a tener suerte

Ofrecido por Google en: [català](#) [galego](#) [euskara](#)

Hello, Iraitz

Ask Gemini



Deep Research



Image



Canvas



Guided Learning





# Large Language Models

Debemos primeramente crear una representación numérica de los textos.

No son palabras, tampoco letras... son algo intermedio que hemos venido a llamar **tokens** de nuestro **vocabulario**.

User

Hola

×

Add message

```
<|im_start|>user<|im_sep|>Hola<|im_end|>
```

gpt-4o

◇

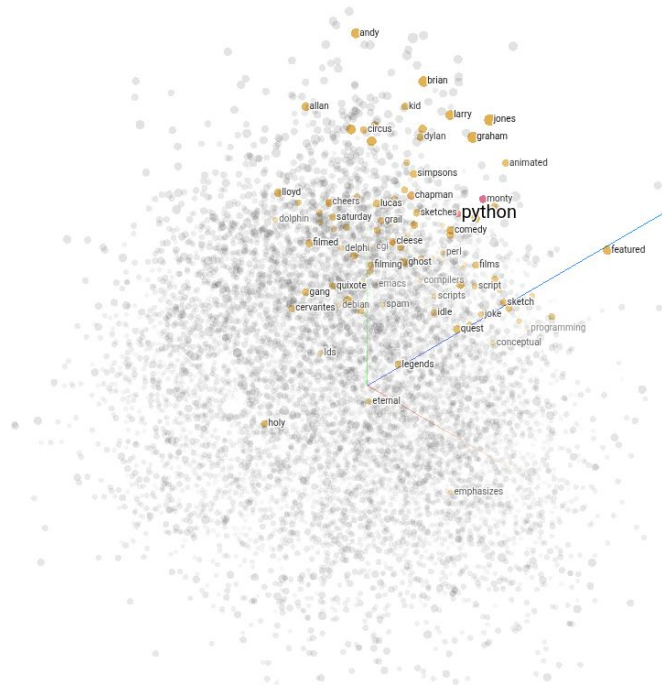
Token count

5

```
<|im_start|>user<|im_sep|>Hola<|im_end|>
```

200264, 1428, 200266, 49864, 200265

Una vez convertidos los datos en números, podemos generar una codificación que represente de lo que habla, su significado. Esto nos permite encontrar términos similares en el espacio de representación.



Search

python

by

word

neighbors

?

100

distance

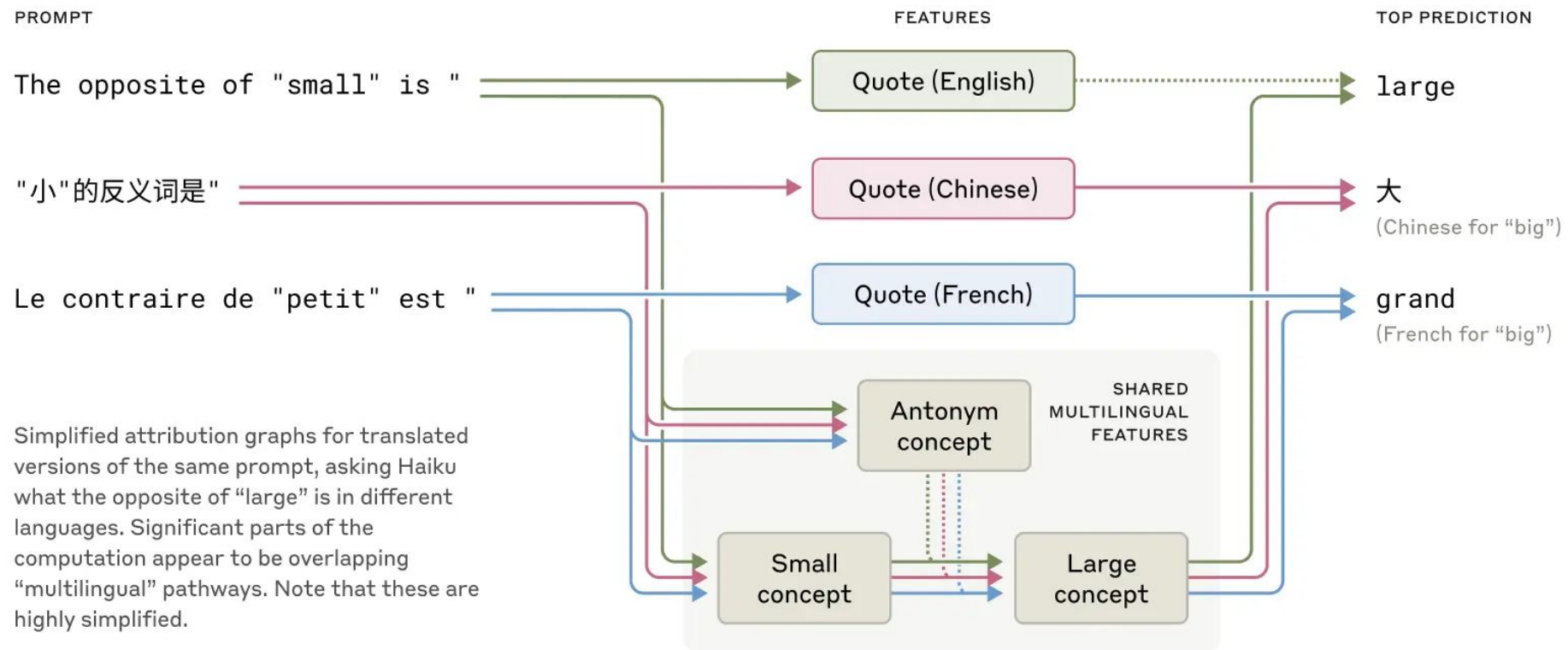
COSINE

EUCLIDEAN

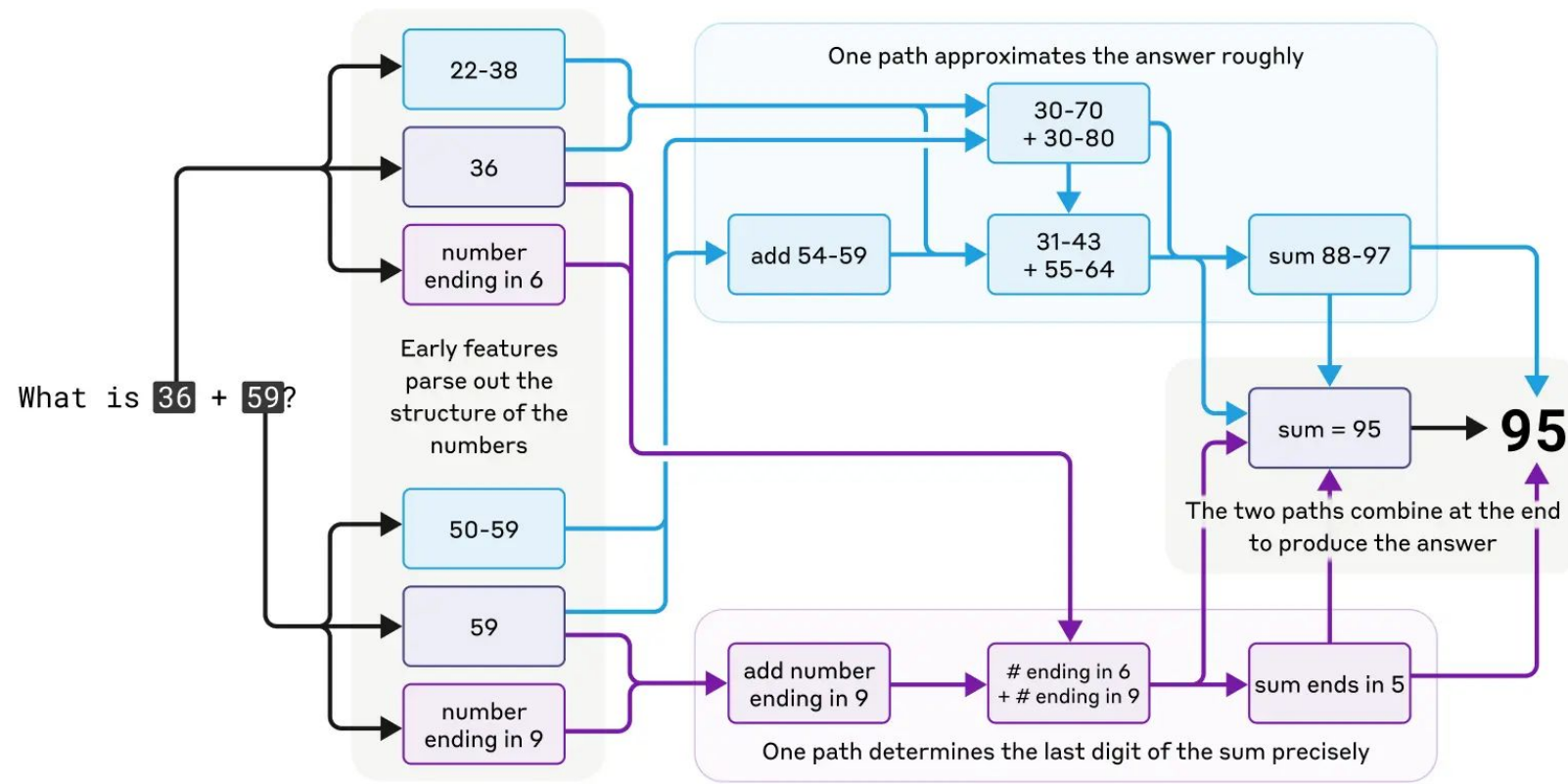
Nearest points in the original space:

monty	0.219
chapman	0.541
cleese	0.560
grail	0.583
circus	0.588
sketch	0.596
idle	0.632
sketches	0.644
scripts	0.681
perl	0.687
brian	0.697
cervantes	0.702
graham	0.704

Esto permite a los modelos traducir o autocompletar sin importar el idioma empleado.



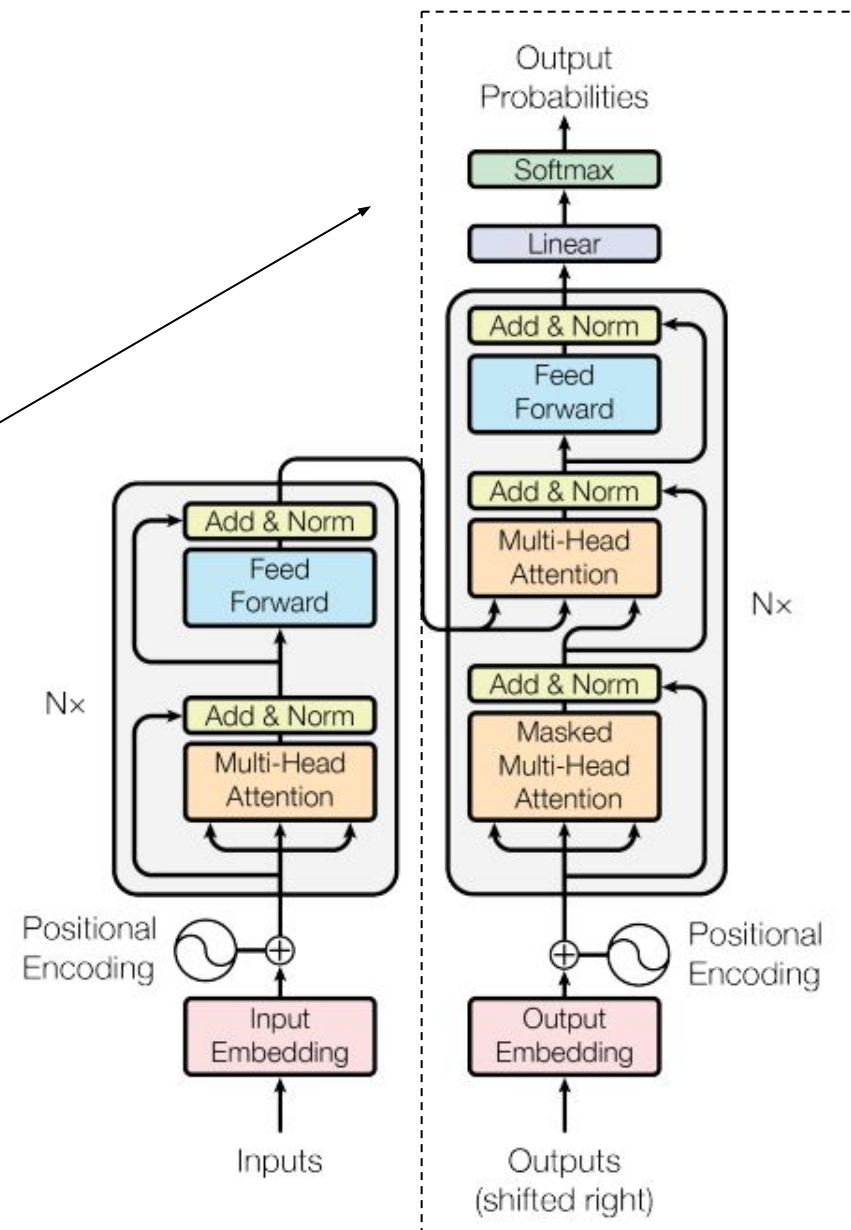
Aunque hay que tener en cuenta que no todas las tareas se les dan bien...



## Large Language Models

Los modelos generativos presentan una arquitectura híbrida de **generación** y **condicionamiento**.

Esta pieza se encarga de generar texto coherente.

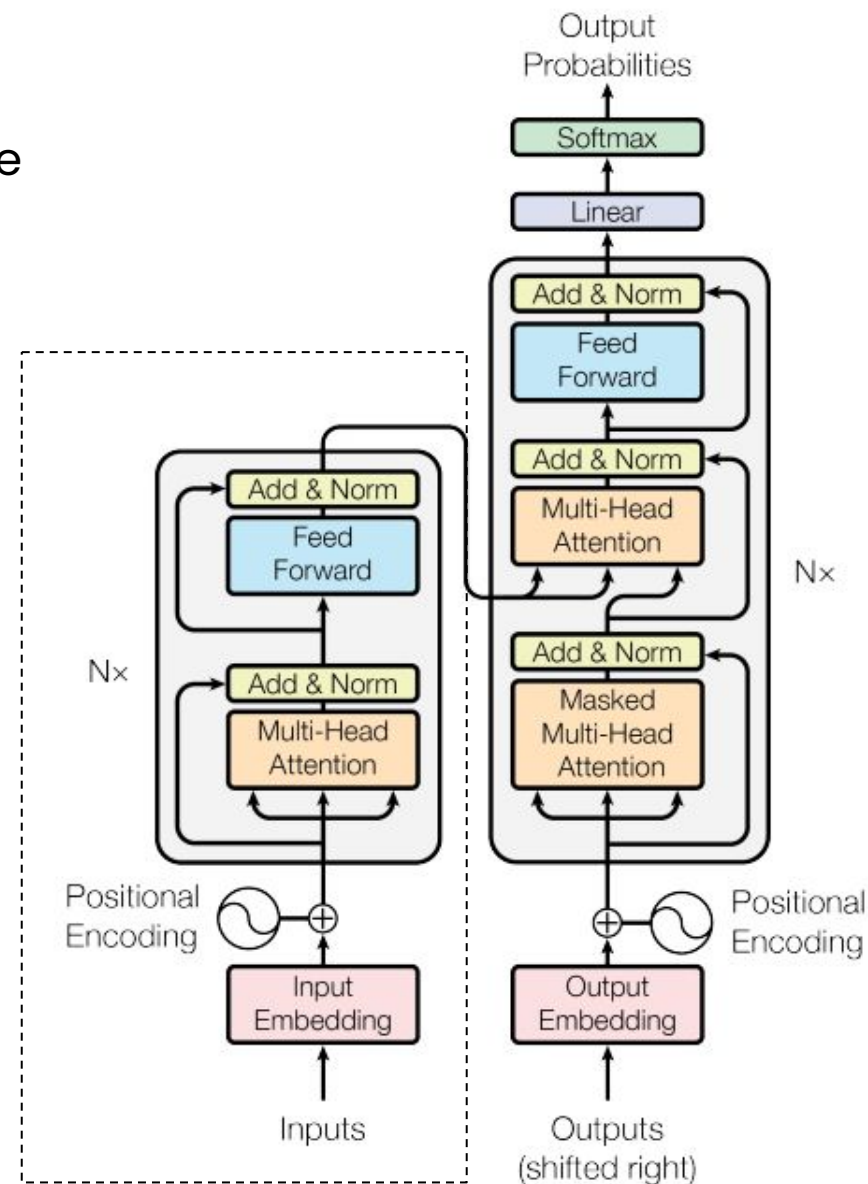




# Large Language Models

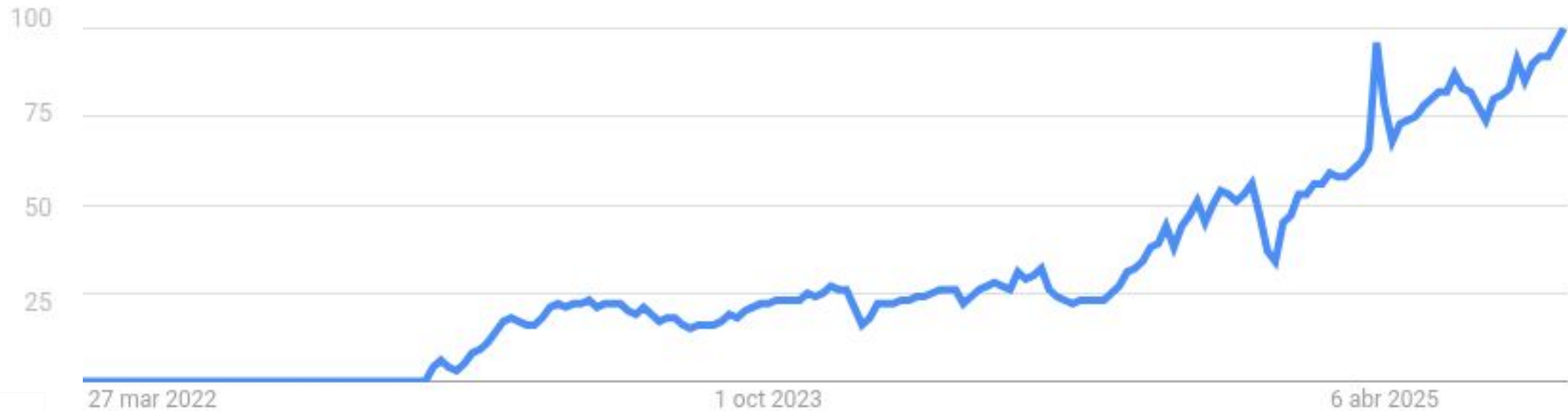
Los modelos generativos presentan una arquitectura híbrida de **generación** y **condicionamiento**.

Esta pieza **instruye** a la generación



# ChatGPT

GPT-2 se liberó en 2019, GPT-3 en 2020... pero el gran momento fue cuando se implementan versiones **ajustadas** a responder, dando lugar a ChatGPT.



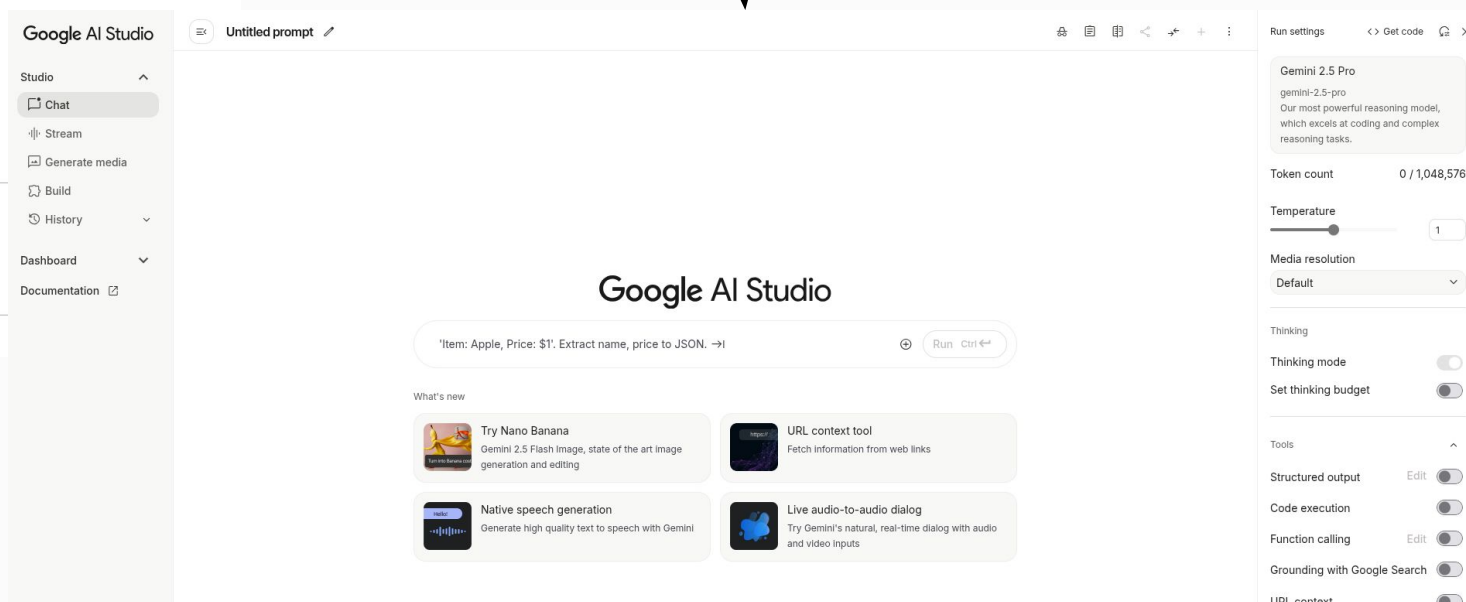
## Playgrounds

Hola, Iraitz

*Playground*

Pregunta a Gemini

+ 🔍 Deep Research 🖼️ Imagen 📄 Pizarra 📖 Aprendizaje guiado



The screenshot displays the Google AI Studio web application. On the left is a sidebar with navigation options: Studio, Chat (selected), Stream, Generate media, Build, History, Dashboard, and Documentation. The main area is titled 'Untitled prompt' and contains a large text input field with the placeholder 'Pregunta a Gemini'. Below the input field are several tool buttons: '+', 'Deep Research', 'Imagen', 'Pizarra', and 'Aprendizaje guiado'. The right sidebar shows the 'Run settings' panel, which includes the selected model 'Gemini 2.5 Pro', token count, temperature slider, media resolution, and various tool toggles like 'Structured output', 'Code execution', 'Function calling', 'Grounding with Google Search', and 'URL context'. The main content area shows a prompt: 'Item: Apple, Price: \$1. Extract name, price to JSON. →I' and a 'Run' button. Below the prompt, there's a 'What's new' section with four featured updates: 'Try Nano Banana', 'URL context tool', 'Native speech generation', and 'Live audio-to-audio dialog'.

[https://aistudio.google.com/prompts/new\\_chat](https://aistudio.google.com/prompts/new_chat)

# Playgrounds

It allows to extract our configuration once ready.

The image shows a screenshot of a Google AI Playground interface. On the left, a 'Get code' window is open, displaying Python code for using the Gemini 2.5 Flash-Lite model. The code includes dependencies for google-genai and google.genai, and a generate() function. A red arrow points from the 'Get code' button in the settings panel to the code window. On the right, the 'Run settings' panel is visible, showing the selected model (Gemini 2.5 Flash-Lite), system instructions, temperature slider, thinking mode toggle, and structured output toggle. The 'Get code' button is circled in red.

**Get code**

Python

< > Code

```
# To run this code you need to install the following dependencies:  
# pip install google-genai  
  
import base64  
import os  
from google import genai  
from google.genai import types  
  
def generate():  
    client = genai.Client(  
        api_key=os.environ.get("GEMINI_API_KEY"),  
    )  
  
    model = "gemini-2.5-flash-lite"  
    contents = [  
        types.Content(  
            role="user",  
            parts=[  
                types.Part.from_text(text="""INSERT_INPUT_HERE"""),  
            ],  
        ),  
    ]
```

Close Copy

Run settings

< > Get code

**Gemini 2.5 Flash-Lite**  
gemini-2.5-flash-lite  
Our smallest and most cost effective model, built for at scale usage.

**System instructions**  
Optional tone and style instructions for the model

**Temperature**  
1

Thinking

Thinking mode ☒

Set thinking budget ☐

Tools

Structured output Edit ☐

Gracias a esta capacidad de condicionar las salidas y trabajar con los embeddings podemos disponer de modelos interoperables entre:

- Texto
- Imagen
- Video
- Audio
- Código
- ...

Siempre podemos encontrar una forma de representar los datos que nos ocupan.

Aunque deberemos dominar:

- La temperatura
- El contexto



Identify the injury shown in this medical image and explain how it can be diagnosed.

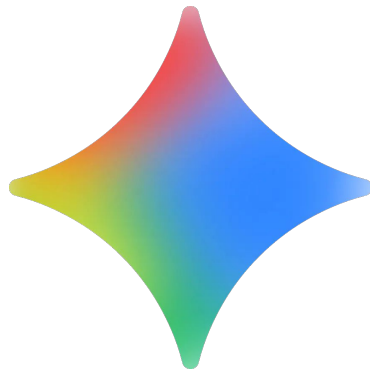
The image shows a Colles fracture, a type of distal radius fracture characterized by:

1. **Fracture of the distal radius:** Visible disruption or break in the distal part of the radius.
2. **Dorsal displacement:** The fractured portion of the radius is displaced toward the back (dorsal) side of the wrist.
3. **Comminution:** Possible multiple bone fragments.
4. **Apex volar angulation:** Angulation of the fracture with the apex pointing toward the palm (volar aspect).



## Proveedores

Por suerte disponemos de proveedores que nos harán la vida más sencilla ya que únicamente deberemos pedir y se nos concederá... si tenemos crédito suficiente.



**GAMMA**



**ElevenLabs**

