



UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

# Módulo 6 IA Generativa

Diploma en Inteligencia Artificial



informatica.usm.cl  
@informaticausm





UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

## Acerca del relator

## Pablo Álvarez González

- Profesor Diploma Inteligencia Artificial de la UTFSM
- Socio Director Sistemas Openb SpA
- Ingeniero Informático de la USACH, Master en Finanzas Aplicadas de la UDD, Diplomado en Data Analytics de la Universidad de Chile/MIT Sloan y Master en Inteligencia Artificial de la UAI.
- Más de 15 años en roles gerenciales , supervisando la implementación de soluciones personalizadas de Infraestructura TI y proyectos de analítica de datos e Inteligencia Artificial.
- Fundador de una ONG enfocada en analítica de datos para impacto social.
- Durante este año ha impartido cursos de "IA Generativa con Python" y "Programación de LLM usando Python".





# Contenido Módulo 6: IA Generativa

1. Bienvenida y Marco de Trabajo
2. Introducción
3. Modelos Probabilísticos, HMM, GMM, VAE
4. Transformers, LLM, GPT
5. Modelos de Difusión
6. Aplicaciones de LLM
7. Desafíos





# Contenido Módulo 6: IA Generativa

1. Bienvenida y Marco de Trabajo
2. Introducción
3. Modelos Probabilísticos, HMM, GMM, VAE
4. Transformers, LLM, GPT
5. Modelos de Difusión
6. Aplicaciones de LLM
7. Desafíos





## Protocolo de uso de IA para el Módulo de Planificación

### ✓ LO QUE SÍ HAREMOS:

- ✓ Los alumnos podrán usar modelos LLM o Generativos, indicándolo explícitamente
- ✓ Generar borradores de documentos
- ✓ Analizar y mejorar textos existentes
- ✓ Brainstorming
- ✓ Validar consistencia en planes
- ✓ Anexar pantallas, modelos usados

### ✗ LO QUE NO HAREMOS:

- ❖ Reemplazar el criterio profesional con respuestas automáticas
- ❖ Copiar y pegar sin revisar ni entender el contenido
- ❖ Usar IA para decisiones críticas sin validación humana
- ❖ Compartir información confidencial de proyectos reales
- ❖ Delegar completamente la responsabilidad profesional a la IA



UNIVERSIDAD TÉCNICA  
FEDERICO SANTA MARÍA

DEPARTAMENTO  
DE INFORMÁTICA

## Trabajo Personal

- Trabajo Modulo 6 IA Generativa
- Fecha de Inicio: 9-12-2025
- Fecha de Entrega: 20-12-2025 23:59 hrs
- Restricciones:
  - Trabajo Individual
  - Si piden ayuda a otro compañero deben citarlo
  - Si usan material de otro compañero deben citarlo



UNIVERSIDAD TÉCNICA  
FEDERICO SANTA MARÍA

DEPARTAMENTO  
DE INFORMÁTICA

# Capítulo 1: Introducción

- ¿Qué es la IA Generativa?
- Inteligencia Artificial
- Contexto de la IA Generativa (GenAI)
- Capacidades y Aplicaciones de GenAI
- Marco Ético y Riesgos de la Implementación



**¿Qué es la IA Generativa**





UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

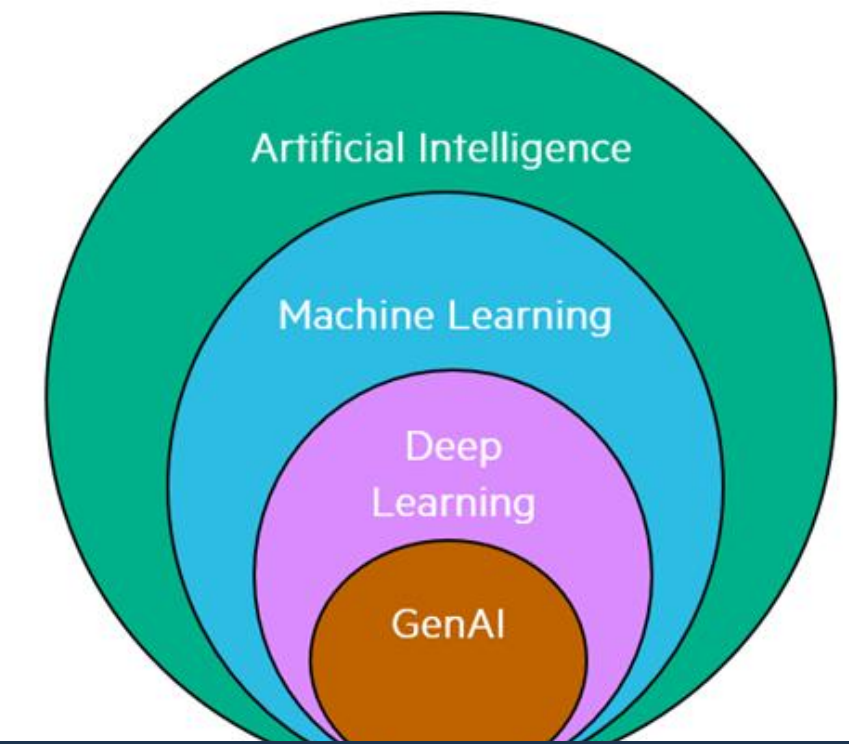
DEPARTAMENTO  
DE INFORMÁTICA

# Inteligencia Artificial

La meta básica es replicar el comportamiento humano o desarrollar máquinas inteligentes

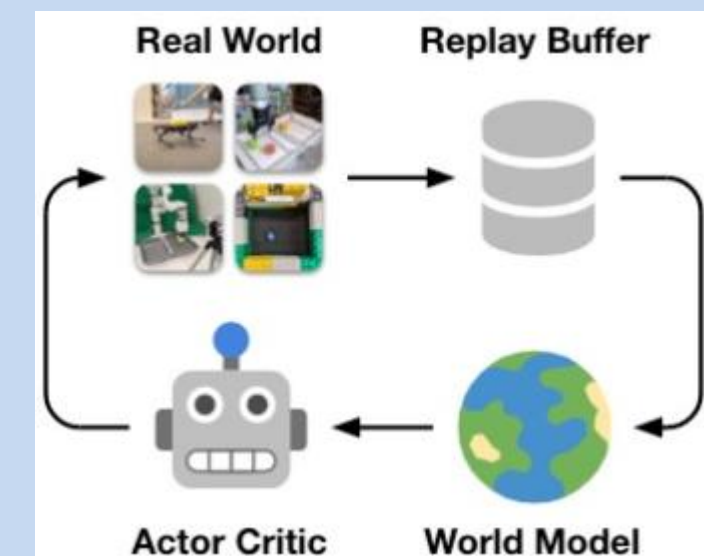
Las sub metas son:

- Percepción
- Razonamiento
- Control/Movimiento
- Planificación
- Comunicación
- Creatividad
- Aprendizaje



- **DayDreamer aprende un modelo generativo de experiencia para RL, ej. Modelo de mundo sin simulación.**

- **Robots cuadrúpedos aprenden a caminar en menos de 1 hora**



**¿De qué es capaz la IA  
Generativa?**



UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

# Generación de Texto

**Prompt:** Debes escribir la prueba matemática de que hay infinitos números primos.

Requerimiento de Estilo: Hazlo en el estilo del Antipoema Chileno Nicanor Parra. Usa un tono irónico, coloquial, directo y con un poco de humor negro o escepticismo

EL ANTIPOETA: Mira, Científico, con ese aire de quien ha resuelto el misterio del universo, que no es más que un desorden elegante, ¿vienes a contarme la patraña de los infinitos primos? Por favor. Los números son como los políticos: una plaga que se multiplica, pero ¿infinitos? Eso suena a propaganda divina. ¿Cuál es el truco de magia de hoy, Euclides?

EL CIENTÍFICO: No es magia, Antipoeta, es matemática pura. Es una contradicción al absurdo, que es como decir: "Si asumo que la vida tiene sentido, el resultado es que me muero de risa, luego la vida no tiene sentido". Aquí es igual, pero con menos drama y más precisión.

EL ANTIPOETA: Adelante con tu farsa. Ya me preparo para bostezar.

El "Artefacto" de la Demostración (A la Chilena)

EL CIENTÍFICO:

Asumiremos, por el bien de la tertulia y la irritación general, que la lista de números primos es FINITA. Que tienen un final, como un carrete que se acaba o una botella de pisco.

Si la lista es finita, podemos hacer una lista completa y ordenada:

$$N = (p_1 \times p_2 \times p_3 \times \dots \times p_n) + 1$$

Donde  $p_n$  es el último primo. El cabo de la cuerda. El fin de la historia.

EL ANTIPOETA:

Ya,  $p_n$ . ¿Y ese pobre diablo qué culpa tiene de ser el último?

EL CIENTÍFICO:

Ninguna. Pero ahora viene la mala noticia para él. Vamos a construir un número, llamémosle  $N$ , el Número-Bomba, que es tan pretencioso como tú:

Multiplicamos todos los primos de la lista y le sumamos uno, por pura insolencia:

$$N = (p_1 \times p_2 \times p_3 \times \dots \times p_n) + 1$$

EL ANTIPOETA:

¡Un número nuevo! ¡Qué original! Como si no hubiera suficientes. ¿Y eso qué prueba, que la estupidez es contagiosa?

EL CIENTÍFICO:

Paciencia. Ahora, este número  $N$  debe cumplir una de dos cosas, porque según la aritmética, todo número mayor a 1 es o primo o tiene un factor primo. Es la ley, y la ley es dura.

Caso 1: ¡El Nuevo Primo Incómodo! EL CIENTÍFICO: Supongamos que  $N$  es primo. Si  $N$  es primo, felicidades, hemos encontrado un primo más grande que  $p_n$ , que se suponía era el último. EL

ANTIPOETA: ¡Contradicción! Has roto tu propio juguete. Es como si el cura jurara que Dios no existe.

Caso 2: El Factor Primo Escondido EL CIENTÍFICO: Supongamos que  $N$  no es primo. Entonces debe tener un factor primo, llamémosle  $q$ . Es decir,  $q$  debe dividir a  $N$ .

- Como  $q$  es primo, debe estar en nuestra lista inicial  $P$  (porque supusimos que la lista era completa, ¿recuerdas?):

$$q \in \{p_1, p_2, \dots, p_n\}$$

- Si  $q$  es uno de esos primos, entonces  $q$  también debe dividir el producto de todos ellos:

$$q \text{ divide a } (p_1 \times p_2 \times \dots \times p_n)$$

- Pero ¡espera! Sabíamos que  $q$  también divide a  $N$ .

$$q \text{ divide a } N$$

- Si  $q$  divide a  $N$  y también divide a  $N-1$  (el producto), entonces  $q$  debe dividir la diferencia entre ellos:

$$\text{Diferencia} = N - (p_1 \times p_2 \times \dots \times p_n) = 1$$

Por lo tanto,  $q$  debe dividir a 1.

EL ANTIPOETA: Un momento. Un número primo  $q$  no puede dividir a 1. ¿Quién tiene el descaro de ser primo y dividir a la unidad? Solo el 1 mismo, y el 1 no es primo por decreto universal, una de las pocas cosas que tienen claras estos burócratas de los números. EL CIENTÍFICO: Exacto.

Hemos llegado a otra Contradicción. Si  $N$  no era primo, el factor primo  $q$  que encontramos no puede existir en nuestra lista finita.





UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

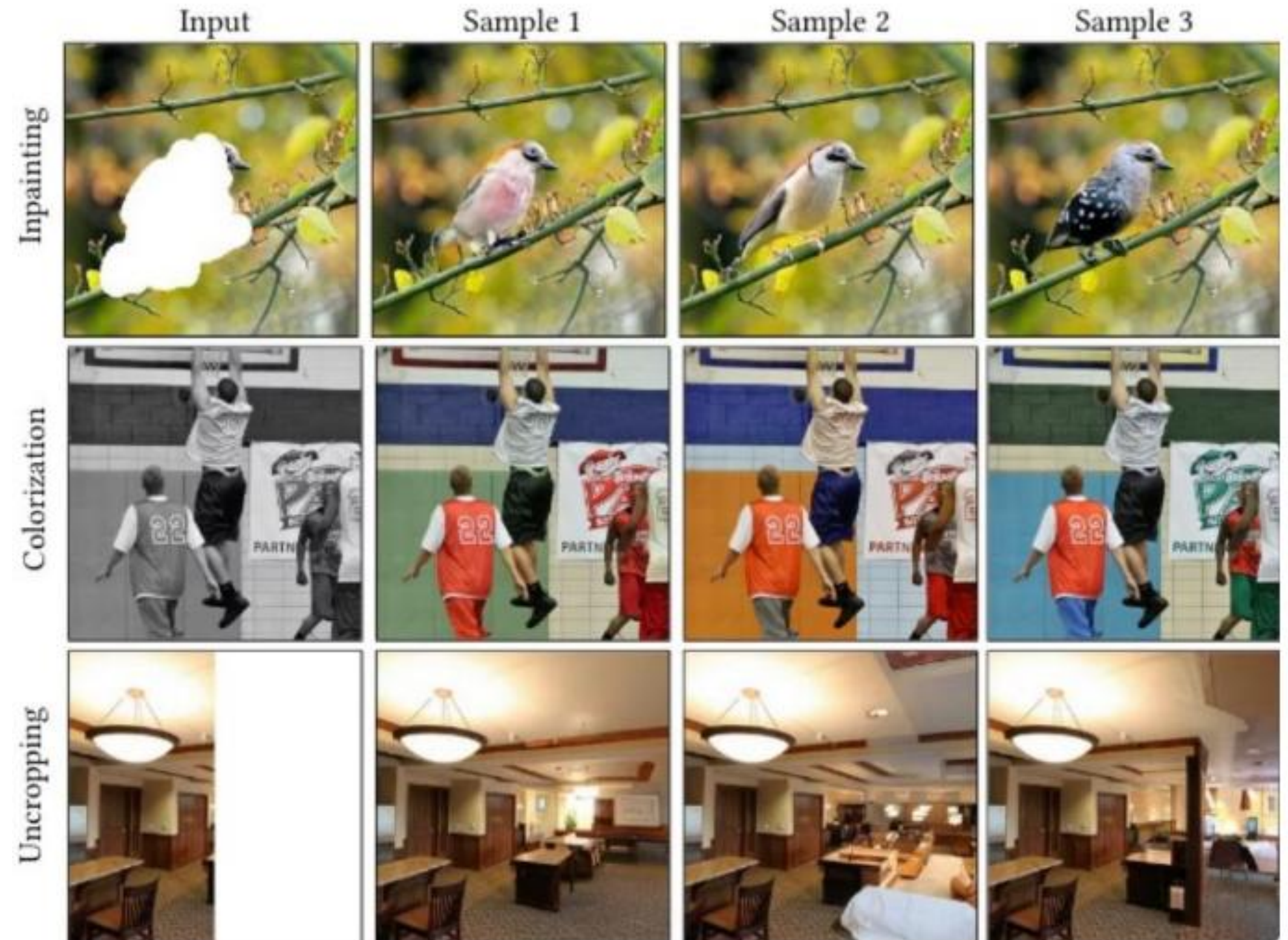
# Edición de Imágenes

Una variedad de tareas implican la edición automática de una imagen:

- Inpainting (relleno inteligente) completa los píxeles faltantes (pre-especificados)
- Colorization (colorización) restaura el color a una imagen en escala de grises
- Uncropping (desrecorte) crea una reconstrucción fotorrealista de un lado faltante de una imagen

Figura que ilustra las tareas de edición de imágenes:

- Inpainting: Imagen de entrada que muestra un pájaro en una rama con un área blanca grande. Las muestras 1, 2 y 3 muestran el pájaro siendo restaurado.
- Colorization: Imagen de entrada que muestra una escena de baloncesto en escala de grises. Las muestras 1, 2 y 3 muestran la escena coloreada.
- Uncropping: Imagen de entrada que muestra una vista parcial de una habitación. Las muestras 1, 2 y 3 muestran la habitación siendo reconstruida.







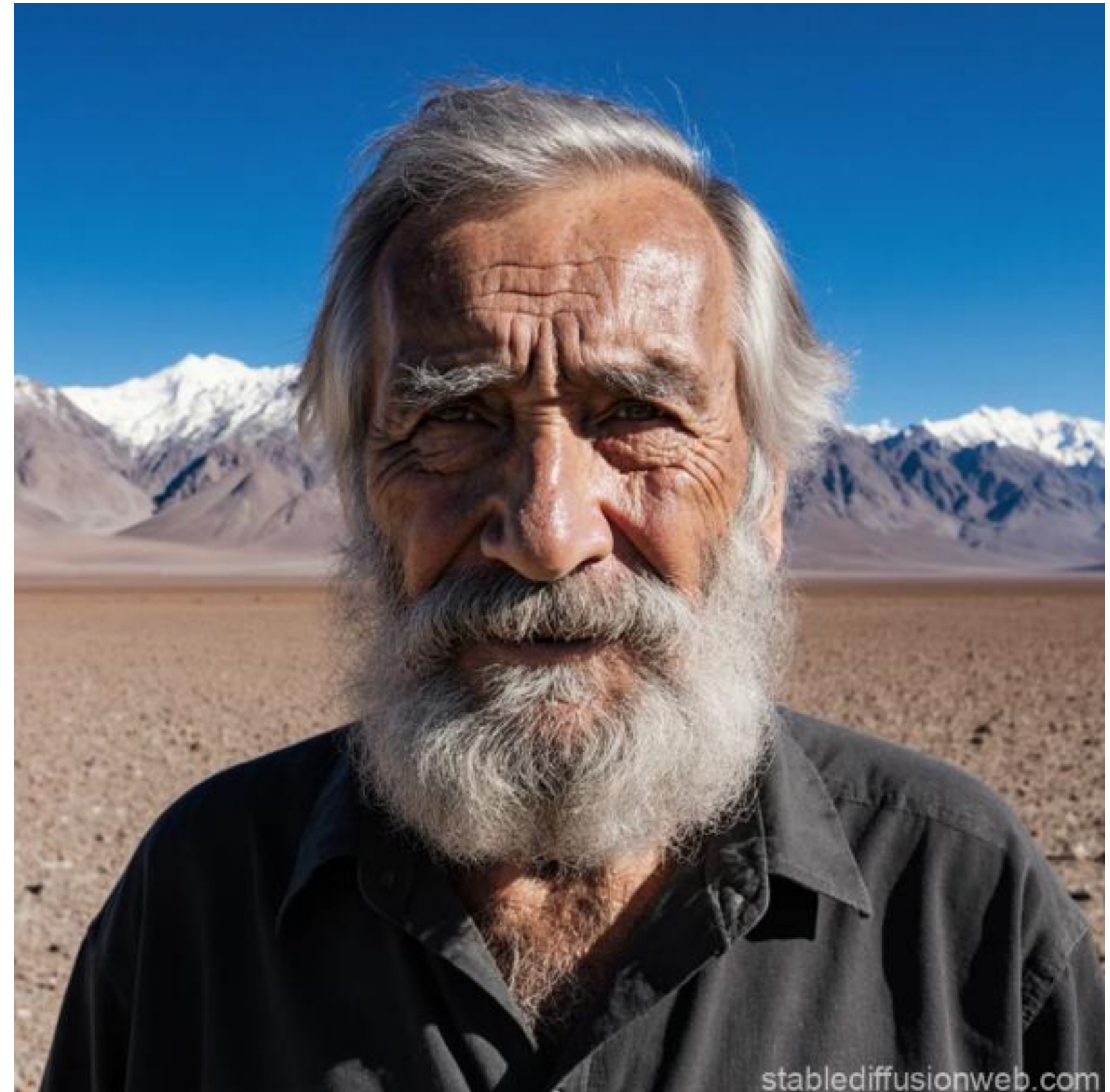
UNIVERSIDAD TÉCNICA  
FEDERICO SANTA MARÍA

DEPARTAMENTO  
DE INFORMÁTICA

Dada una descripción, generar una imagen.

Prompt: Retrato en primer plano de un anciano chileno de la región de Atacama, con piel curtida por el tiempo, ojos hundidos y una expresión amable. Tiene cabello entrecano, barba completa y bigote. Al fondo, el vasto y árido paisaje del desierto de Atacama se extiende, con montañas escarpadas y sin nieve bajo un cielo azul claro e intenso. La iluminación es dura y natural, enfatizando las texturas de su rostro y el ambiente seco. Fotorrealista, muy detallado, 8k, cinematográfico.

## Generación de Imágenes







UNIVERSIDAD TÉCNICA  
FEDERICO SANTA MARÍA

DEPARTAMENTO  
DE INFORMÁTICA

# Generación de Imágenes

Dada una descripción, generar una imagen.

Prompt: Retrato en primer plano de un anciano chileno de la región de Atacama, con piel curtida por el tiempo, ojos hundidos y una expresión amable. Tiene cabello entrecano, barba completa y bigote. Al fondo, el vasto y árido paisaje del desierto de Atacama se extiende, con montañas escarpadas y sin nieve bajo un cielo azul claro e intenso. La iluminación es dura y natural, enfatizando las texturas de su rostro y el ambiente seco. Fotorrealista, muy detallado, 8k, cinematográfico.







UNIVERSIDAD TÉCNICA  
FEDERICO SANTA MARÍA

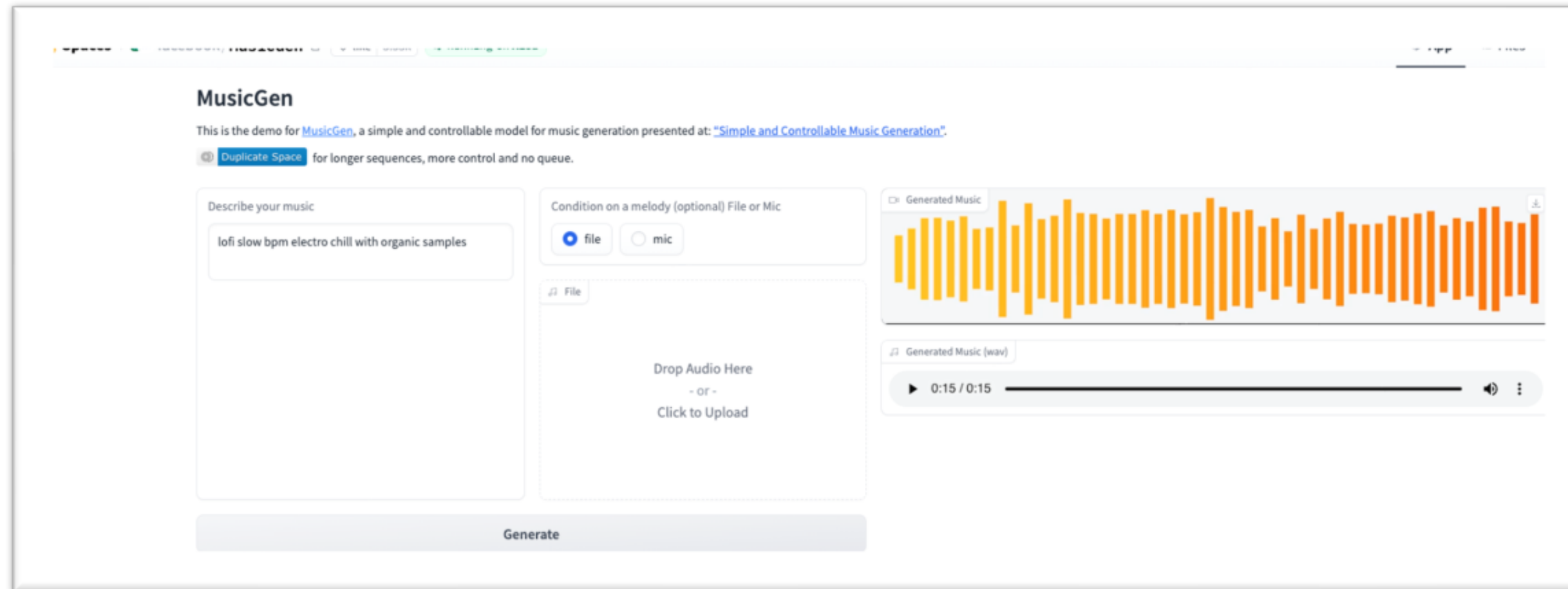
DEPARTAMENTO  
DE INFORMÁTICA

# Generación de Videos

- Los modelos de difusión latente utilizan un espacio de baja dimensionalidad latente por eficiencia
- Pregunta clave: ¿cómo generar múltiples frames relacionados?
- ‘Alinea tus latentes’ inserta convolución / atención temporal entre cada convolución / atención espacial
- ‘Preserva tu propia correlación’ / incluye ruido temporalmente correlacionado

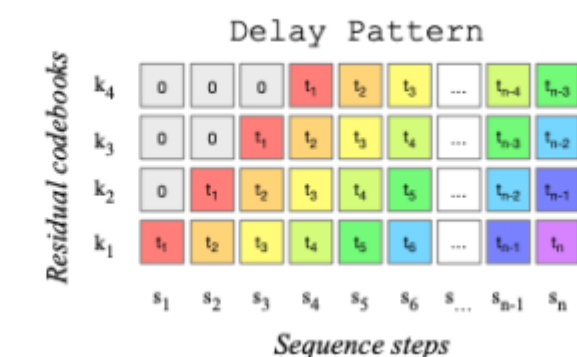
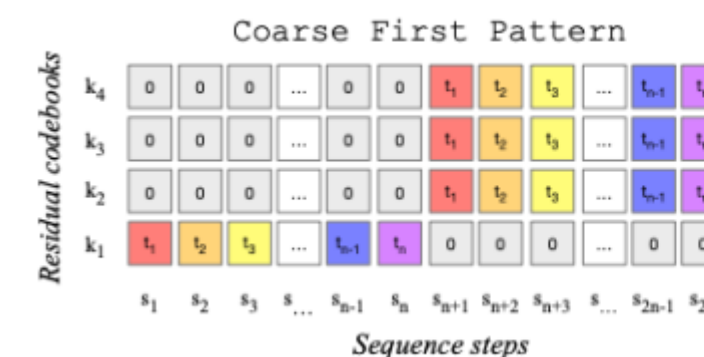
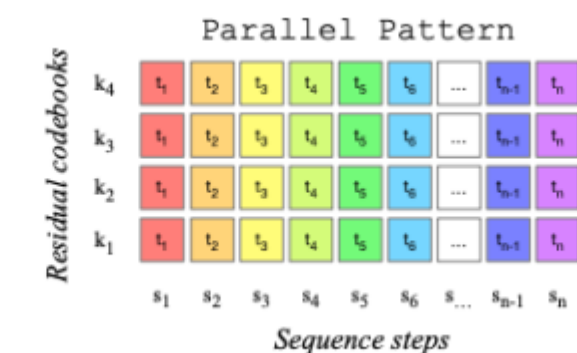
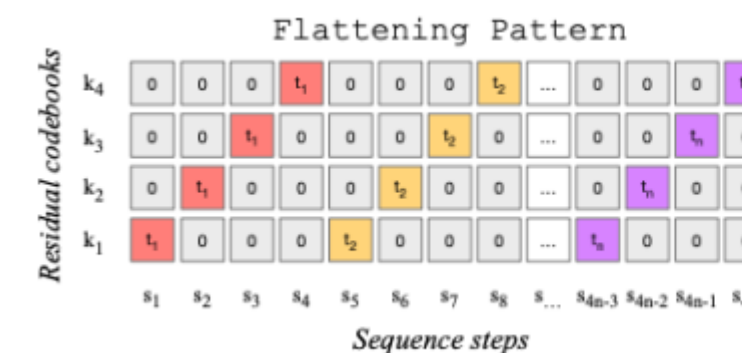






## MusicGen

- Un modelo descodificador transformador sobre **unidades cuantizadas** (elementos discretos de un libro de códigos de *frames* de audio)
- **Entrelaza sonidos** ajustando cómo los libros de códigos se atienden entre sí
- Permite el **condicionamiento** en texto y/o muestras de audio





UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

# Generación de Código

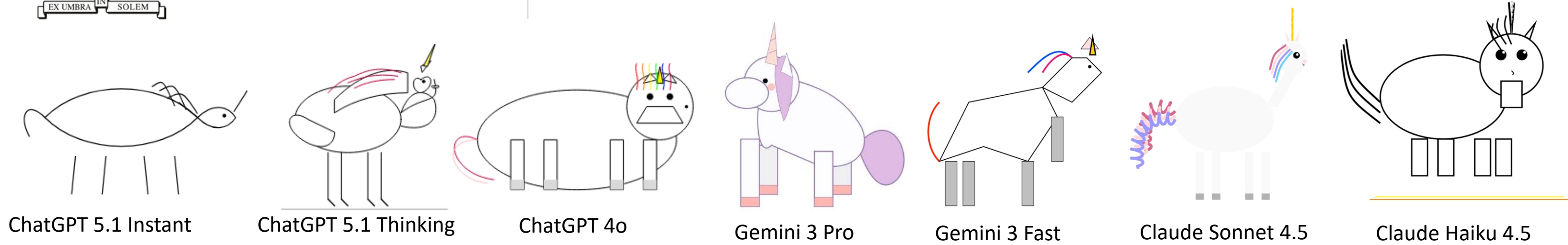


Figura: Código latex que representa varios intentos con el prompt “Genera un unicornio en TikZ”



**Ampliando**



UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

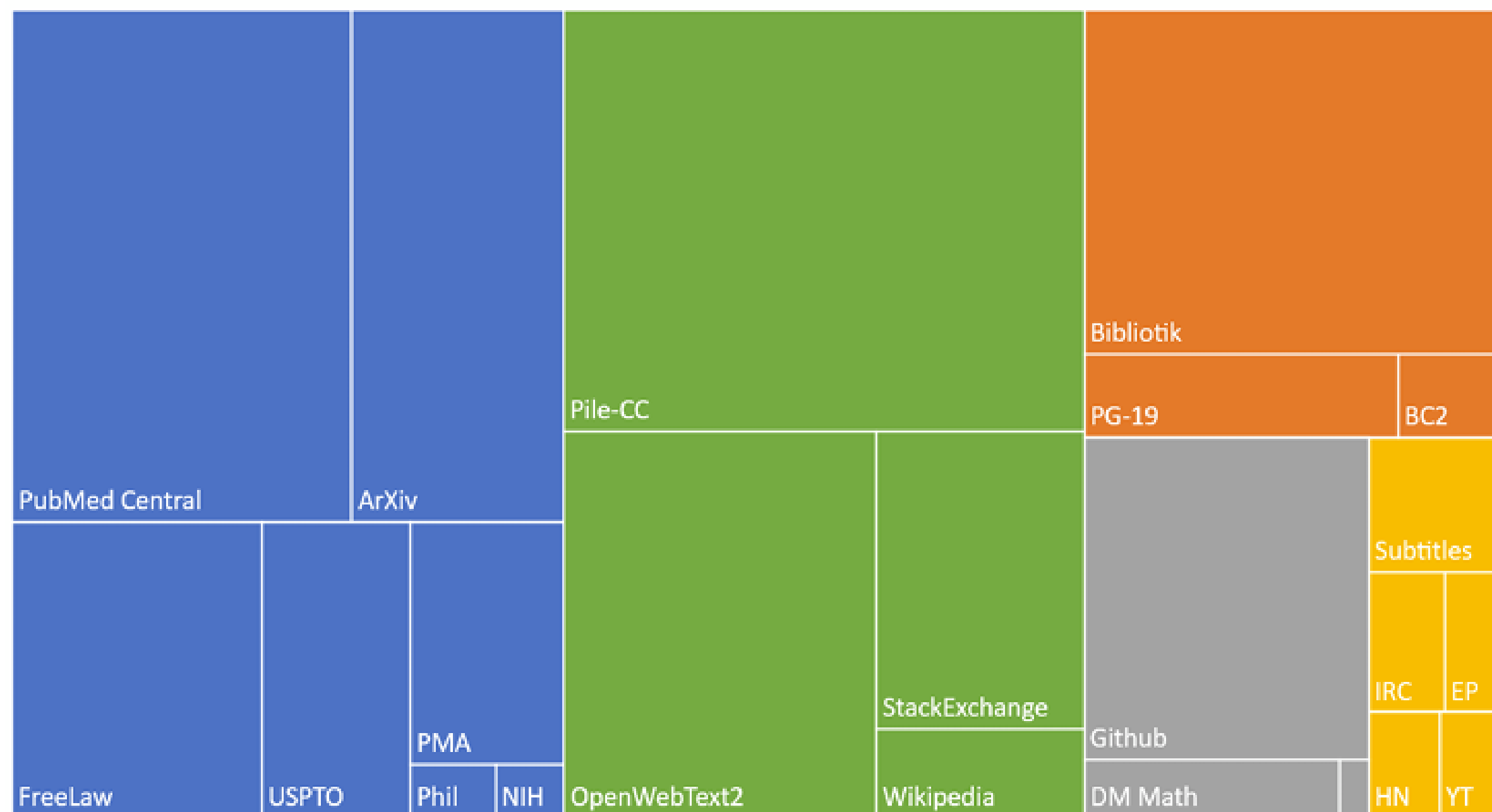
# Datos de Entrenamiento para LLMs

## La Pila:

- Dataset Open source para entrenamiento
- Compuesto por 22 pequeños datasets
- Se privilegian textos de alta calidad
- 825 GB eq 1.2 trillones de tokens

## Composition of the Pile by Category

■ Academic ■ Internet ■ Prose ■ Dialogue ■ Misc





- **InstructGPT** utiliza Reinforcement Learning with Human Feedback(RLHF) para ajustar un modelo GPT preentrenado.
- Del artículo: “In human evaluations on our prompt distribution, outputs from the 1.3B parameter InstructGPT model are preferred to outputs from the 175B GPT-3, despite having 100x fewer parameters.”

**La Figura ilustra los 3 pasos del método:**

- Paso 1: Recopilar datos de demostración y entrenar una política supervisada. Se toma un sample de un prompt de nuestro conjunto de datos de prompts. Un etiquetador demuestra el comportamiento de salida deseado. Estos datos se utilizan para ajustar GPT-3 mediante aprendizaje supervisado.
- Paso 2: Recopilar datos de comparación y entrenar un modelo de recompensa. Se toma un sample de un prompt y varias salidas del modelo. Un etiquetador clasifica las salidas de mejor a peor. Estos datos se utilizan para entrenar nuestro modelo de recompensa.
- Paso 3: Optimizar una política contra el modelo de recompensa usando Reinforcement Learning. Se toma un sample de un nuevo prompt del conjunto de datos. La política genera una salida. El modelo de recompensa calcula una recompensa para la salida. La recompensa se utiliza para actualizar la política usando PPO.

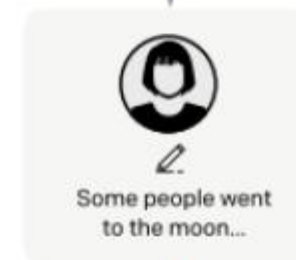
**Step 1**

**Collect demonstration data,  
and train a supervised policy.**

A prompt is  
sampled from our  
prompt dataset.



A labeler  
demonstrates the  
desired output  
behavior.



This data is used  
to fine-tune GPT-3  
with supervised  
learning.



**Step 2**

**Collect comparison data,  
and train a reward model.**

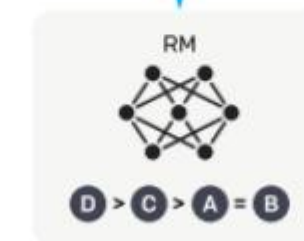
A prompt and  
several model  
outputs are  
sampled.



A labeler ranks  
the outputs from  
best to worst.



This data is used  
to train our  
reward model.



**Step 3**

**Optimize a policy against  
the reward model using  
reinforcement learning.**

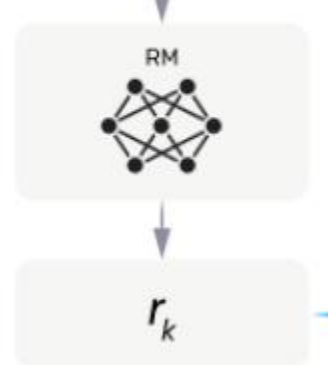
A new prompt  
is sampled from  
the dataset.



The policy  
generates  
an output.



The reward model  
calculates a  
reward for  
the output.





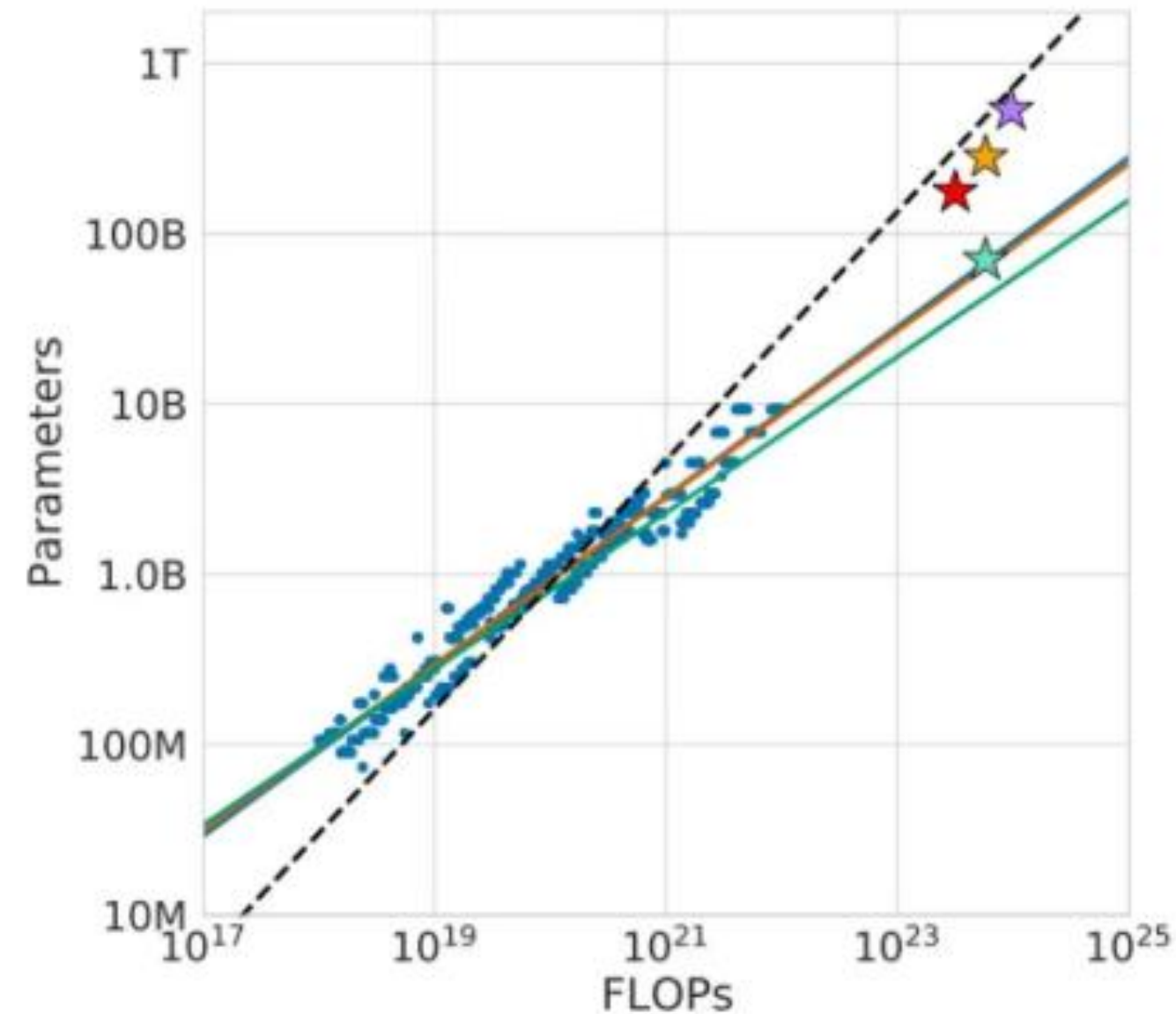


UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

# Costos de Entrenamiento

- Eje Y: Parameters (escala logarítmica,  $10^M$  a  $10^T$  )
- Eje X: FLOPs (escala logarítmica,  $10^{17}$  a  $10^{25}$  )
- Leyenda:
  - Enfoque 1
  - Enfoque 2
  - Enfoque 3
  - Kaplan et al (2020)
  - Chinchilla (70B)
  - Gopher (280B)
  - GPT-3 (175B)
  - Megatron-Turing NLG (530B)

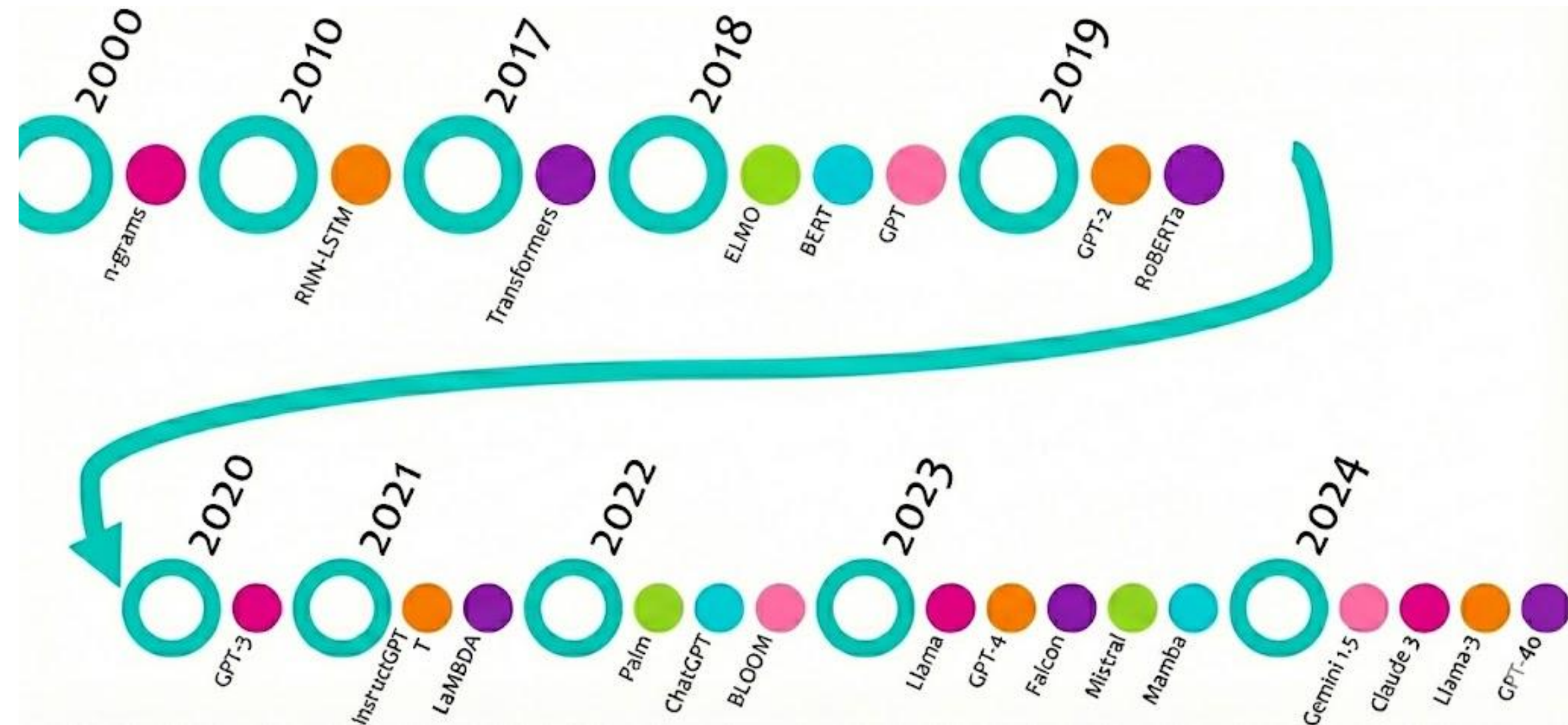


- Approach 1
- Approach 2
- Approach 3
- - - Kaplan et al (2020)
- ★ Chinchilla (70B)
- ★ Gopher (280B)
- ★ GPT-3 (175B)
- ★ Megatron-Turing NLG (530B)



## ■ Cronología de Language Models

- 2000: n-grams
- 2010: RNN-LSTM
- 2017: Transformer LMs
- 2018: ELMO, BERT, GPT
- 2019: GPT-2, RoBERTa
- 2020: GPT-3
- 2021: InstructGPT, LaMDA
- 2022: Palm, ChatGPT, BLOOM
- 2023: Llama, GPT-4, Falcon, Mistral, Mamba
- 2024: Gemini 1.5, Claude 3, Llama-3, GPT-4o







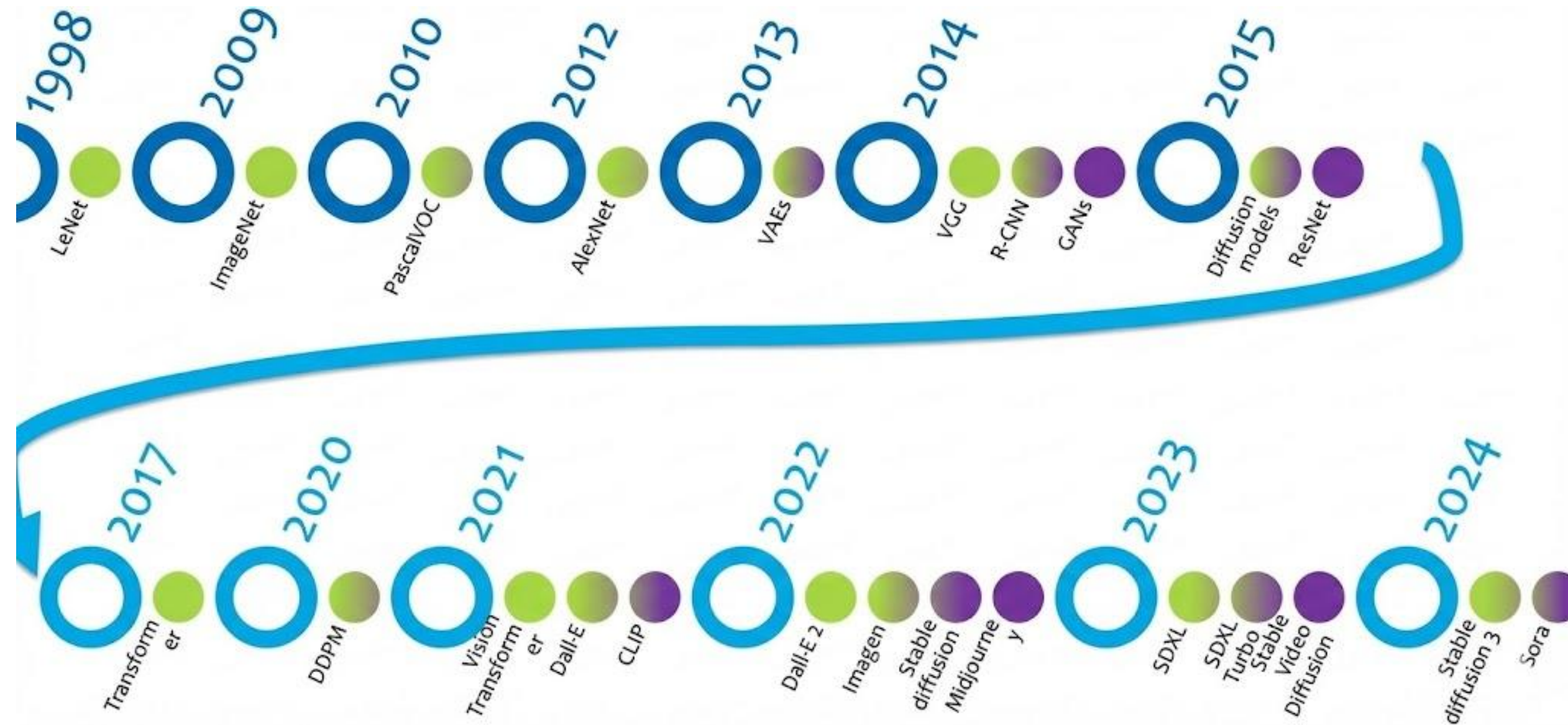
UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

# Timeline Generación de Imágenes

## ■ Cronología de Generación de Imágenes

- 1998: LeNet
- 2009: ImageNet
- 2010: Pascal VOC
- 2012: AlexNet
- 2013: VAEs
- 2014: VGG, R-CNN, GANs
- 2015: Diffusion models, ResNet
- 2017: Transformer
- 2020: DDPM
- 2021: Vision Transformer, Dall-E, CLIP
- 2022: Dall-E 2, Imagen, Stable diffusion, Midjourney
- 2023: SDXL, SDXL Turbo, Stable Video Diffusion
- 2024: Stable diffusion 3, Sora





**¿Por qué Entender  
GenAI?**



UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA







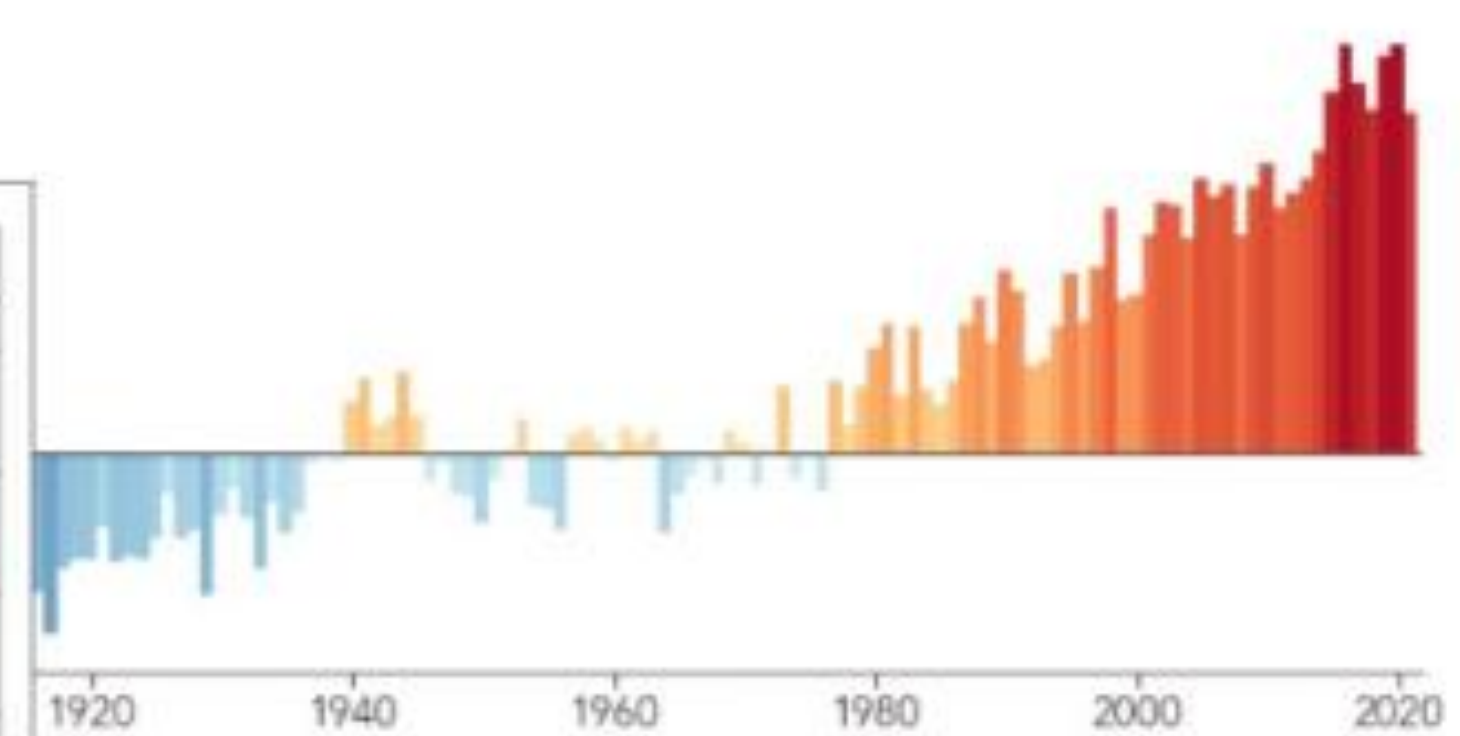
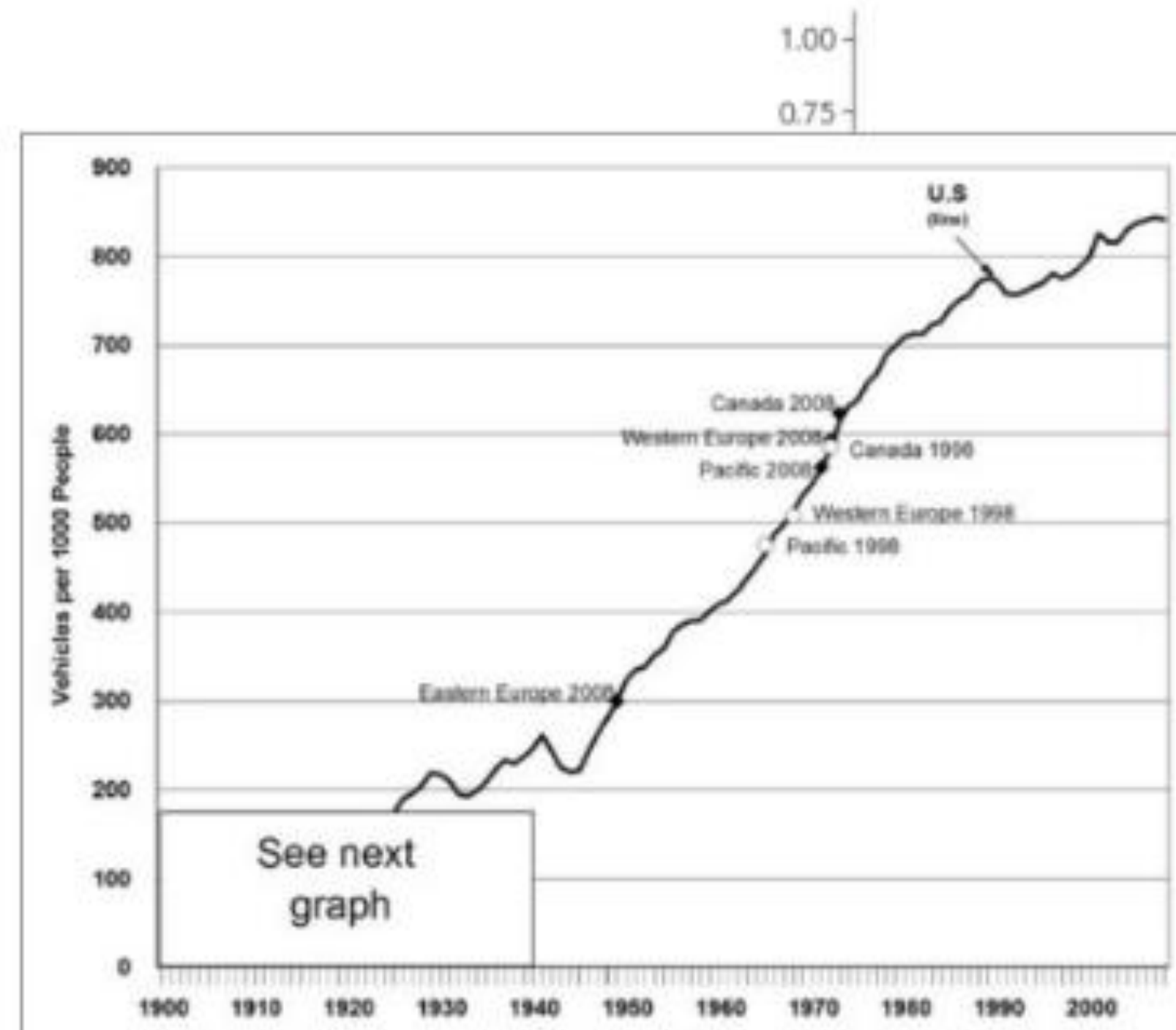
UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA



Imagen desde: <https://daily.jstor.org/the-science-of-traffic/>









Number of Annual U.S. Pedestrian Fatalities, 1980-2022



% of World Population That Uses a Smartphone





UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA





**Pero.....**



UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

# Riesgos en el uso de GenIA

timesofindia.indiatimes.com/technology/tech-news/deloitte-makes-ai-mistake-again-this-time-in-report-for-canadi...



Technology

Tech News

Tech Tips

Mobiles & Tabs

Hack Of Day

AI News

Today's ePaper



## Deloitte makes 'AI mistake' again, this time in report for Canadian government; here's what went wrong

TOI Tech Desk / TIMESOFINDIA.COM / Updated: Nov 27, 2025, 15:43 IST

Share



AA

Select TOI as

Deloitte faces scrutiny again for AI-generated errors in a Canadian government healthcare report, mirroring a similar incident in Australia. The firm refunded \$290,000 for an Australian report with fabricated references. These repeated AI missteps raise serious concerns about accuracy and oversight in critical government projects, potentially impacting policy decisions and eroding public trust.

[Read Less](#)



Deloitte has once again landed in trouble for an apparent AI-related mistake. Deloitte recently prepared a major healthcare report for the

Advertisement

FP&A Software You Don't Hate



Open



Tired of too many ads?

[Go Ad Free Now](#)







UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

## MIT Technology Review

Publicado por Opinión

Temas ▾

Informes ▾

Eventos ▾



# Cómo la IA está introduciendo errores en los tribunales

James O'donnell · 20/05/2025 · 7 min leer



Han sido un par de semanas de historias sobre IA en los tribunales. Es posible que hayas oído hablar de la víctima fallecida en un incidente de violencia en la carretera cuya familia creó un [avatar IA](#) para mostrarlo como declaración de impacto (posiblemente la primera vez que se hace en Estados Unidos). Pero, según los juristas, se está gestando una controversia mayor y de mayores consecuencias. **Las alucinaciones de la IA aparecen cada vez más en los expedientes judiciales.** Y está empezando a

## Artículos Relacionados



INTELIGENCIA ARTIFICIAL

Chatbots de compañía y el futuro de la privacidad

By Eileen Guo Melissa

Helkkilä

26/11/2025



# Riesgos en el uso de GenIA

28 Nov, 2025

Argentina

Colombia

España

REGISTRARME

INICIAR SESION

Colapinto Hace 13 minutos ⚡ Trends Tasas de interés Patricia Bullrich Últimas Noticias Chechu Bonell

ESTADOS UNIDOS >

## Dos abogados fueron multados tras presentar un escrito judicial con información falsa realizado por ChatGPT

Los letrados "obviaron conscientemente" las señales que sugerían que los datos eran irreales y los defendieron, a pesar de las dudas sembradas por las órdenes judiciales

No fue éste el caso de los abogados Steven Schwartz y su compañero Peter LoDuca, quienes se valieron de este chatbot para realizar un escrito judicial pero no se percataron de que el texto contenía datos falsos.

Los letrados pidieron a la aplicación de **OpenAI** que elabore un documento de diez páginas para presentar en defensa de su cliente, quien demanda a la aerolínea Avianca tras haber sufrido una lesión por un golpe con un carrito de servicio durante un vuelo en 2019.





UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

EL PAÍS

1 AÑO, 9,90 €



## Tecnología

¡Oferta especial Black Friday!

TU TECNOLOGÍA · CIBERSEGURIDAD · PRIVACIDAD · INTELIGENCIA ARTIFICIAL · INTERNET · GRANDES TECNOLÓGICAS · ÚLTIMO

PRIVACIDAD >

# Protección de Datos multa con 2.000 euros al difusor de uno de los desnudos falsos de las chicas de Almendralejo

La resolución, que se centra en el tratamiento de la imagen personal sin consentimiento, es la primera que penaliza la creación de un 'deepfake', o imagen generada por IA

## Riesgos en el uso de GenIA

28/11/25, 12:28 p.m.

CEO of world's biggest ad firm targeted by deepfake scam | Technology | The Guardian



### Technology

This article is more than 1 year old

## CEO of world's biggest ad firm targeted by deepfake scam

Exclusive: fraudsters impersonated WPP's CEO using a fake WhatsApp account, a voice clone and YouTube footage used in a virtual meet

Nick Robins-Early

Fri 10 May 2024 03:01 EDT

The head of the world's biggest advertising group was the target of an elaborate deepfake scam that involved an artificial intelligence voice clone. The CEO of **WPP**, Mark Read, detailed the attempted fraud in a recent email to leadership, warning others at the company to look out for calls claiming to be from top executives.

Fraudsters created a WhatsApp account with a publicly available image of Read and used it to set up a Microsoft Teams meeting that appeared to be with him and another senior WPP executive, according to the email obtained by the Guardian. During the meeting, the impostors deployed a voice clone of the executive as well as YouTube footage of them. The scammers impersonated Read off-camera using the meeting's chat window. The scam, which was unsuccessful, targeted an "agency leader", asking them to set up a new business in an attempt to solicit money and personal details.





1. **Selecciona una herramienta** (ChatGPT, Gemini, Claude, etc.) y envíale el siguiente *prompt*: "**Desglosa los 5 pasos más importantes que una PyME chilena debe seguir para implementar un asistente conversacional basado en GPT/LLM. Sé realista en términos de presupuesto.**"
2. **Transcribe los 5 pasos que obtuviste.** En la parte superior de tu respuesta, indica claramente:
  - **Modelo Usado y el *Prompt* Exacto.** (Cumpliendo con el protocolo del módulo).
3. **Evalúa la respuesta de la IA (2-3 líneas):** ¿La respuesta es completa y práctica para un gerente (e.g., menciona costos, infraestructura, alineación con el negocio, o solo menciona código)? ¿Qué información esencial le falta?

# Conceptos





## Discriminativos

Aprenden la frontera de decisión entre clases

- Objetivo:  $P(Y|X)$
- Clasifican y predicen etiquetas
- Ejemplos: SVM, regresión logística
- Menor complejidad computacional

## Generativos

Aprenden la distribución completa de los datos

- Objetivo:  $P(X,Y)$  o  $P(X)$
- Crean nuevos datos similares
- Ejemplos: VAE, GAN, GPT
- Mayor poder expresivo





Antes de los modelos neuronales modernos, existían enfoques probabilísticos fundamentales que siguen siendo relevantes hoy.

01

## Mezcla Gaussiana (GMM)

Modela datos como combinación de distribuciones gaussianas múltiples. Útil para clustering y modelado de densidad.

02

## Modelos Ocultos de Markov (HMM)

Captura secuencias con estados ocultos. Fundamental en reconocimiento de voz y análisis de series temporales.

03

## Autoencoders Variacionales (VAE)

Combina redes neuronales con inferencia variacional. Genera datos aprendiendo representaciones latentes continuas.





UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

# Redes Generativas Adversarias (GAN)

## Arquitectura Dual en Competencia

### Generador

Crea datos sintéticos a partir de ruido aleatorio, intentando engañar al discriminador.

### Discriminador

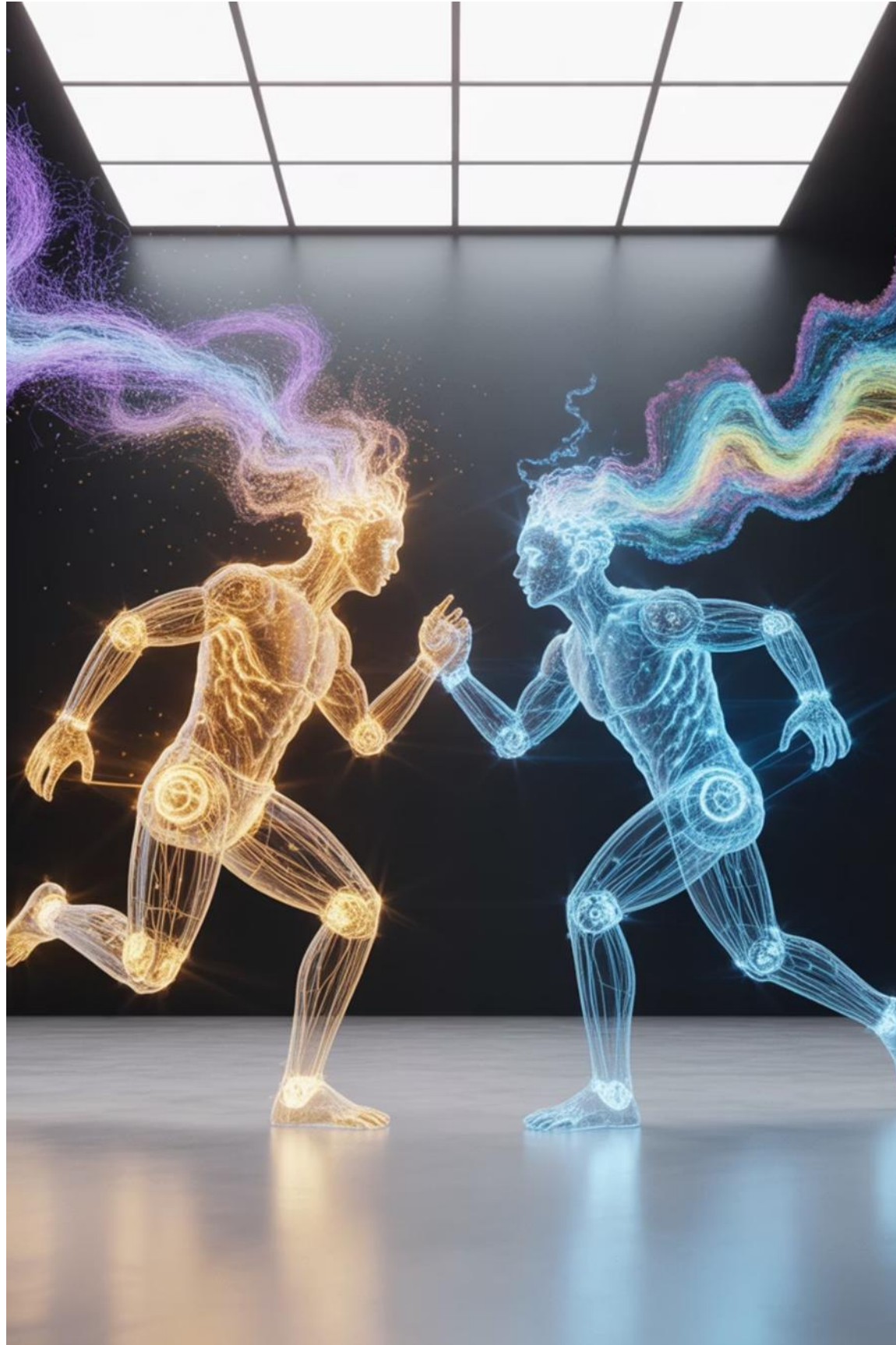
Evalúa si los datos son reales o generados, mejorando constantemente su detección.

### Entrenamiento Adversario

Ambas redes compiten en un juego de suma cero, mejorando mutuamente hasta alcanzar equilibrio de Nash.

### Función de pérdida:

$$\min_G \max_D V(D, G) = E_x[\log D(x)] + E_z[\log(1 - D(G(z)))]$$





UNIVERSIDAD TÉCNICA  
FEDERICO SANTA MARÍA

DEPARTAMENTO  
DE INFORMÁTICA

## Redes Generativas Adversarias (GAN)

# Aplicaciones Revolucionarias de GANs

### Síntesis de Rostros

Creación de rostros humanos ultrarrealistas que no existen  
(StyleGAN, ThisPersonDoesNotExist)

### Traducción de Imágenes

Conversión entre estilos, épocas del año, día/noche con CycleGAN y Pix2Pix

### Super-Resolución

Mejora de calidad y resolución de imágenes mediante técnicas de upscaling inteligente





UNIVERSIDAD TÉCNICA  
FEDERICO SANTA MARÍA

DEPARTAMENTO  
DE INFORMÁTICA

**ThisPersonDoesNotExist**







UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

# Arquitectura Transformer: La Revolución

Introducido en 2017 con "Attention is All You Need", el Transformer eliminó la recurrencia y cambió para siempre el procesamiento de lenguaje natural.

## Mecanismo de Auto-Atención

Permite que cada palabra "atienda" a todas las demás en la secuencia, capturando dependencias a largo plazo sin importar la distancia.

## Paralelización Masiva

A diferencia de RNNs, procesa todas las posiciones simultáneamente, reduciendo drásticamente el tiempo de entrenamiento.

## Encoders y Decoders

Arquitectura modular que permite adaptación flexible: solo encoder (BERT), solo decoder (GPT), o completa (T5).







UNIVERSIDAD TÉCNICA  
FEDERICO SANTA MARÍA

DEPARTAMENTO  
DE INFORMÁTICA

## Modelos Base: BERT vs GPT

### BERT

#### Bidirectional Encoder

Utiliza solo la parte del encoder del Transformer

- Lee contexto bidireccional
- Masked Language Modeling
- Excelente para comprensión
- Clasificación y NER

### GPT

#### Autoregressive Decoder

Utiliza solo la parte del decoder del Transformer

- Generación unidireccional
- Predice siguiente token
- Excelente para generación
- Texto, código, conversación



# La Evolución de GPT

## 1 GPT-1 (2018)

117M parámetros • Demostró el poder del pretraining no supervisado

## 2 GPT-2 (2019)

1.5B parámetros • Generación coherente de texto largo • Controversia inicial por riesgo de desinformación

## 3 GPT-3 (2020)

175B parámetros • Few-shot learning revolucionario • Base de ChatGPT

## 4 GPT-4 (2023)

Multimodal • Razonamiento mejorado • Mayor precisión y seguridad • Capacidades de visión







UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

# Entrenamiento de Modelos de Lenguaje



## Pretraining

Entrenamiento masivo con billones de tokens de texto de internet. Aprende patrones generales del lenguaje.



## Fine-tuning

Ajuste específico con datos de dominio particular. Adapta el modelo a tareas concretas.



## Instruction Tuning

Entrenamiento con ejemplos de instrucciones. Mejora seguimiento de órdenes y alineación con humanos (RLHF).



UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

# Modelos de Difusión: Arte Generativo

Los modelos de difusión aprenden a revertir un proceso gradual de añadir ruido, generando imágenes de calidad excepcional.



## Proceso Directo

Añade ruido gaussiano progresivamente a la imagen original hasta convertirla en ruido puro.



## Proceso Inverso

La red neuronal aprende a eliminar el ruido paso a paso, reconstruyendo la imagen desde el caos.



## Generación Guiada

Condiciona el proceso con texto, permitiendo crear imágenes a partir de descripciones.







UNIVERSIDAD TÉCNICA  
FEDERICO SANTA MARÍA

DEPARTAMENTO  
DE INFORMÁTICA

## Herramientas de Generación Visual



### DALL-E 3

De OpenAI. Integrado con ChatGPT. Excelente comprensión de prompts complejos y coherencia visual.



### Midjourney

Especializado en estética artística. Produce imágenes de alta calidad con estilo único y detallado.



### Stable Diffusion

Modelo open-source. Ejecutable localmente. Gran comunidad y personalización mediante fine-tuning.



### Runway & Pika

Generación de video con IA. Edición avanzada, efectos y creación de clips desde texto o imágenes.



UNIVERSIDAD TECNICA  
FEDERICO SANTA MARIA

DEPARTAMENTO  
DE INFORMÁTICA

## Generación de Contenido

Artículos, código, emails, scripts creativos

## Sistemas Conversacionales

Chatbots, asistentes virtuales, soporte al cliente

## Traducción Automática

Traducción contextual multilingüe en tiempo real

## Asistentes Especializados

Análisis de datos, programación, educación personalizada

# Aplicaciones Prácticas de LLMs







UNIVERSIDAD TÉCNICA  
FEDERICO SANTA MARÍA

DEPARTAMENTO  
DE INFORMÁTICA

## Desafíos de la IA Generativa

### Alucinaciones

Los modelos pueden generar información falsa con alta confianza. Necesario verificación humana crítica.

### Sesgo y Equidad

Reproducen sesgos de datos de entrenamiento. Requiere auditoría constante y técnicas de mitigación.

### Propiedad Intelectual

Debates sobre derechos de autor de contenido generado y datos de entrenamiento utilizados.

### Deepfakes

Potencial para desinformación y suplantación de identidad. Necesidad urgente de detección y regulación.

### Costo Computacional

Entrenamiento e inferencia requieren recursos enormes. Impacto ambiental significativo.

### Desplazamiento Laboral

Automatización de trabajos creativos y cognitivos. Necesario reentrenamiento y adaptación social.