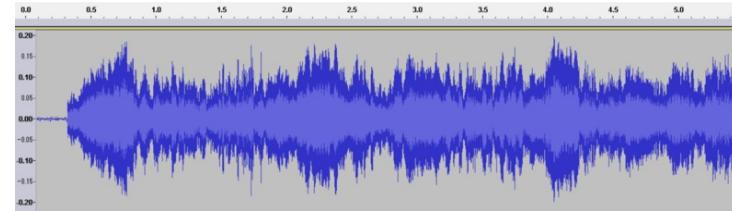


# Introducción

**Desafío:** *entender* datos (entradas) complejas, no estructuradas



Computer Vision



Computational Speech

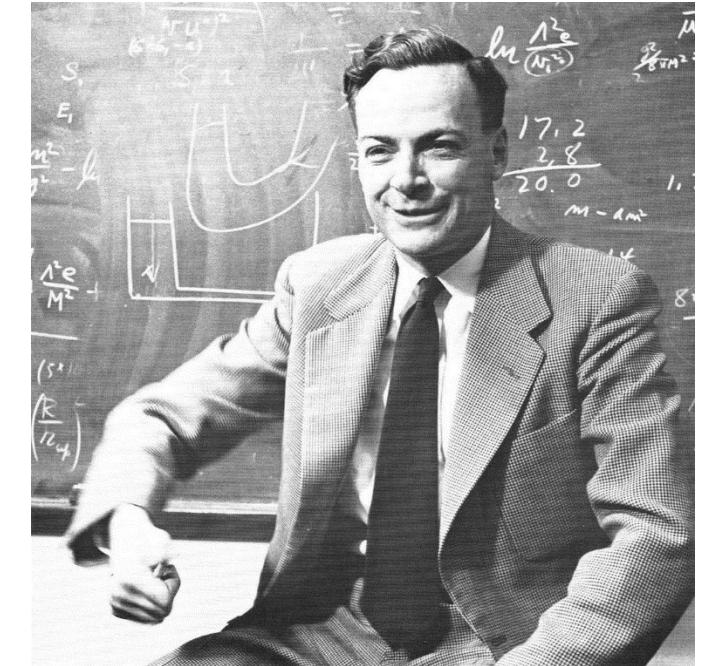


Natural Language Processing



Robotics

# Introducción



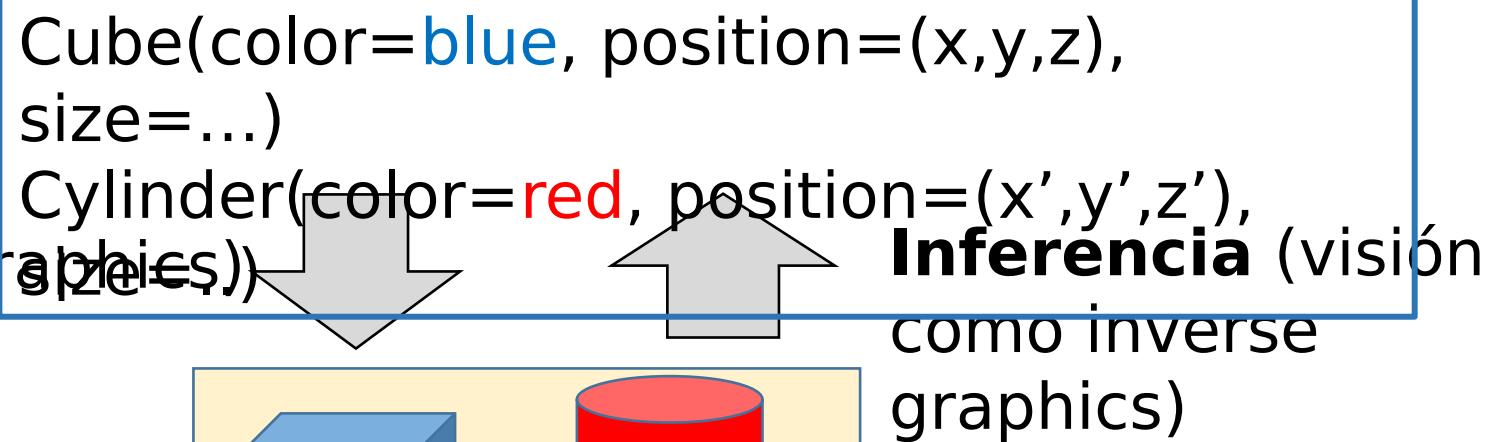
Richard Feynman: “*What I cannot create, I do not understand*”

Modelado generativo: “Lo que comprendo, lo puedo **crear**”

# Modelado generativo: computación gráfica

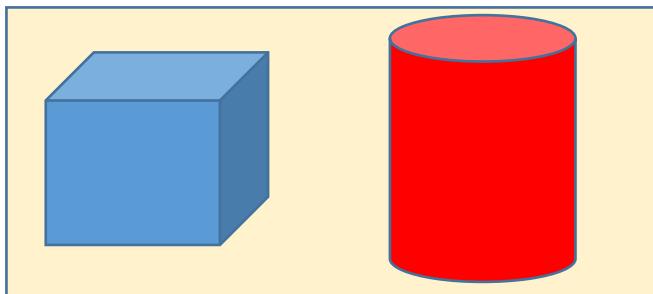
¿Cómo generar imágenes naturales con una computadora?

Descripción  
n  
alto nivel  
**Generación** (graphics)

Cube(color=blue, position=(x,y,z),  
size=...)  
Cylinder(color=red, position=(x',y',z'),  
size=...) 

**Inferencia** (visión  
como inverse  
graphics)

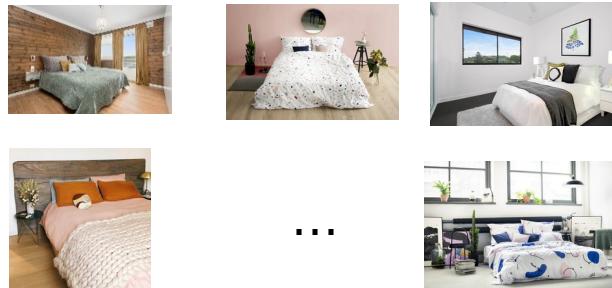
Salida



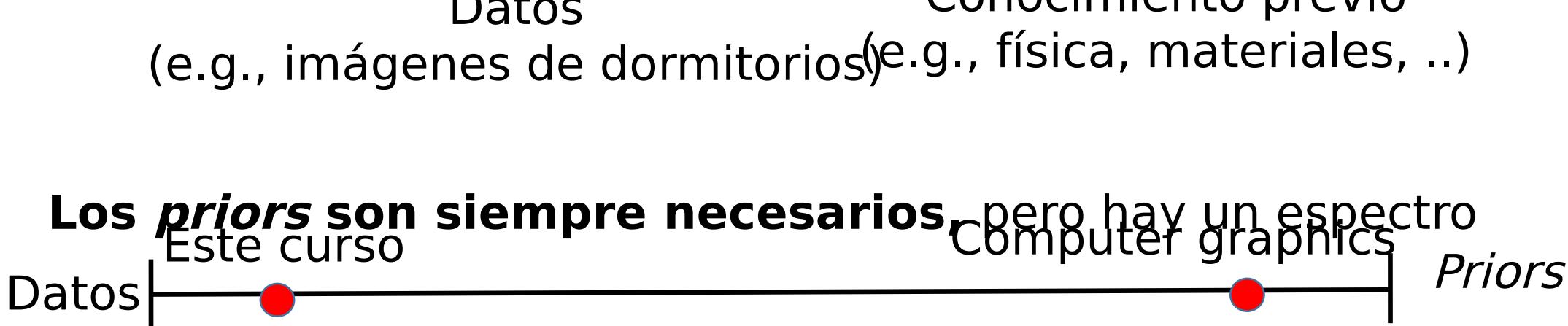
La mayoría de los modelos que veremos tienen una **estructura similar**  
**(generación + inferencia)**

# Modelos generativos estadísticos

Los modelos generativos **estadísticos** se aprenden de los datos



Datos (e.g., imágenes de dormitorios) Conocimiento previo (e.g., física, materiales, ...)



# Modelos generativos estadísticos

Un modelo generativo es una distribución de probabilidad  $p(x)$

- **Datos:** muestras (e.g., imágenes de dormitorios)
- **Conocimiento previo:** forma paramétrica (e.g., Gaussiana?), función de costo (e.g., máxima verosimilitud?), algoritmo de optimización, etc.

Distribución

Imagen  $x$  →  $\rightarrow$  probabilidad  $p(x)$



de  
probabilida  
d  
 $p(x)$

→ probabilidad  $p(x)$

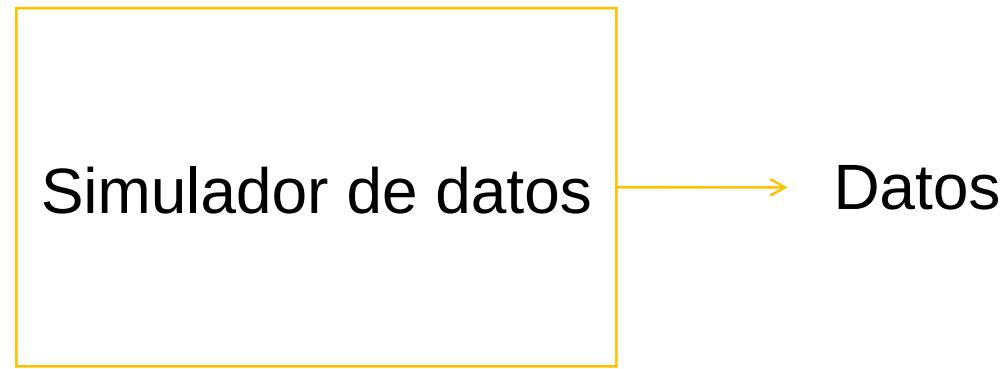
Es generativo porque **muestreo de  $p(x)$  genera nuevas imágenes**



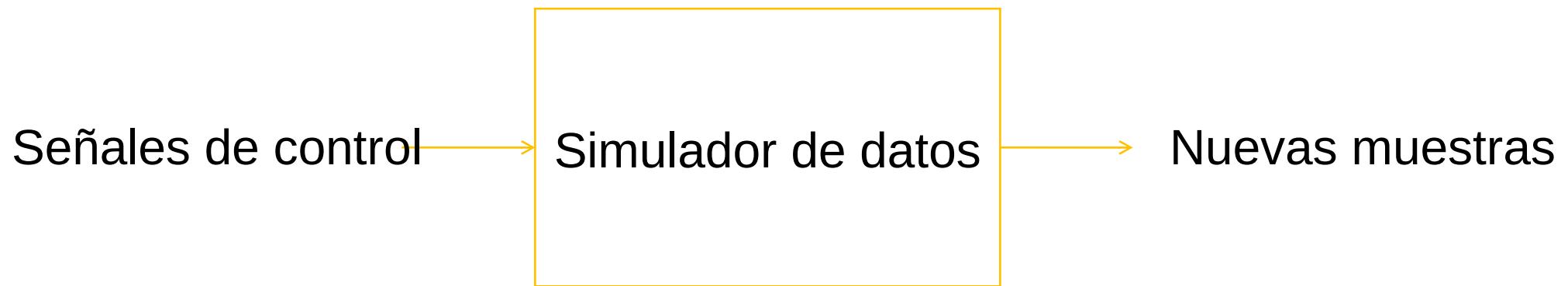
...



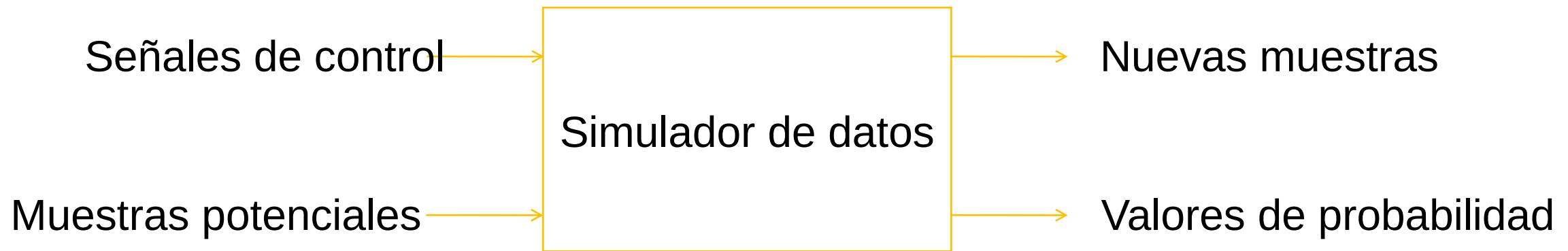
# **Construyendo un simulador para el proceso de generación de datos**



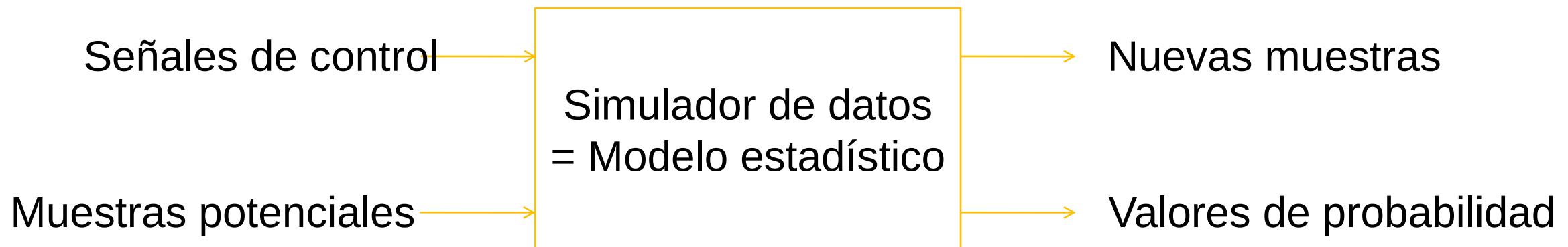
# Construyendo un simulador para el proceso de generación de datos



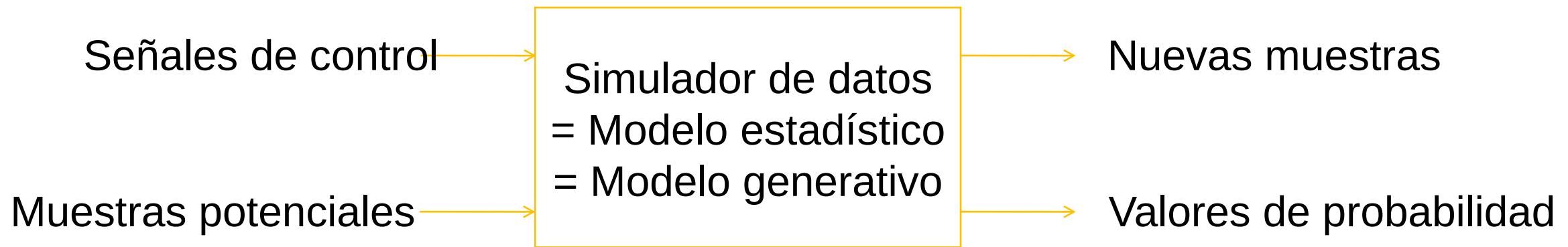
# Construyendo un simulador para el proceso de generación de datos



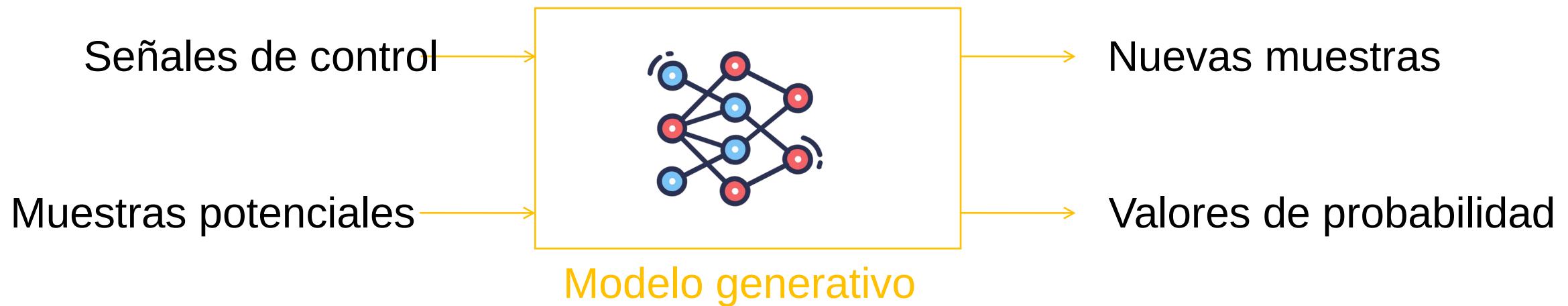
# Construyendo un simulador para el proceso de generación de datos



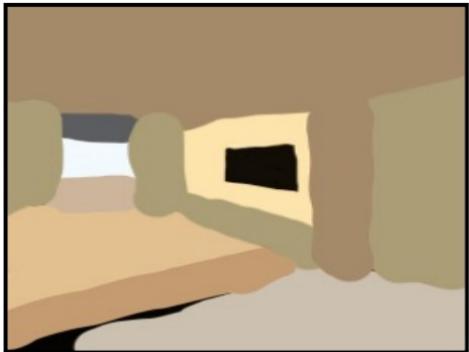
# Construyendo un simulador para el proceso de generación de datos



# Construyendo un simulador para el proceso de generación de datos



# Generación de datos en el mundo real



Generación

Modelo generativo  
de imágenes  
realistas



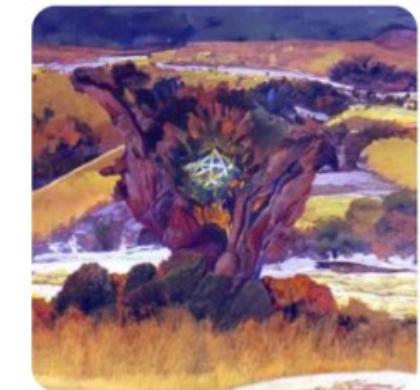
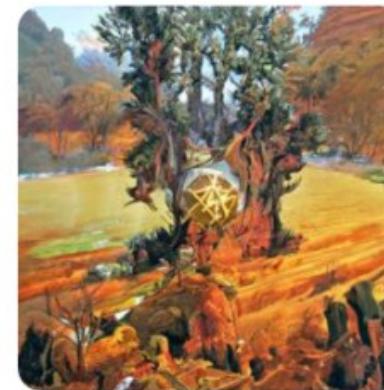
**Stroke paintings to realistic images**  
[Meng, He, Song, et al., ICLR 2022]

“Ace of Pentacles”



Generación

Modelo generativo  
de pinturas



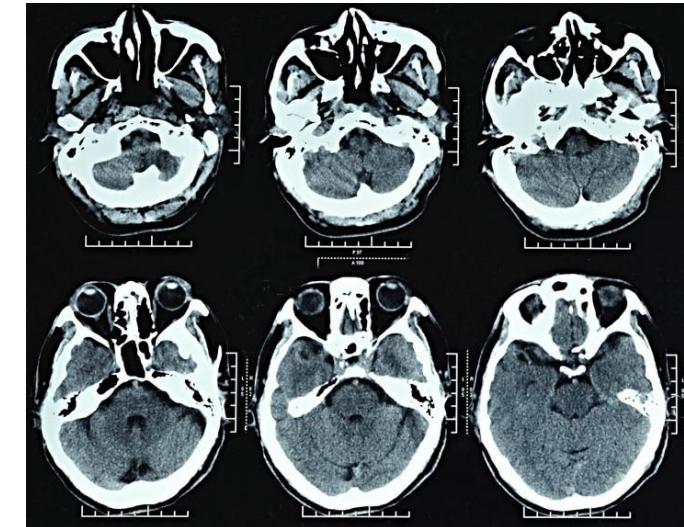
**Language-guided artwork creation**  
<https://chainbreakers.kath.io> @RiversHaveWings

# Resolviendo problemas inversos con modelos generativos



Modelo generativo  
de imágenes  
médicas

Generación

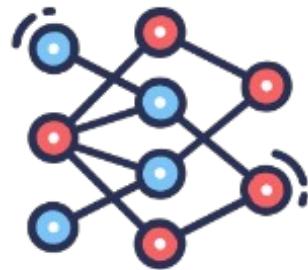


**Medical image reconstruction**  
[Song et al., ICLR 2022]

# Detección de outliers con modelos generativos



Probabilidad  
alta



Probabilidad  
baja



Modelo generativo  
de señales de  
tráfico



**Outlier detection**  
[Song et al., ICLR 2018]

# Progreso en modelos generativos de imágenes - GANs



2014



2015



2016



2017



2018

Ian Goodfellow,  
2019

# Progreso en modelos generativos de imágenes - Modelos de difusión



# Text2Image Diffusion Models

**User input:**

An astronaut riding a horse



# Text2Image Diffusion Models

**User input:**

A perfect Italian meal



# Text2Image Diffusion Models

## User input:

泰迪熊穿着戏服，站在太和殿前唱京剧

A teddy bear, wearing a costume, is standing in front of the Hall of Supreme Harmony and singing Beijing opera



# Dalle3

A minimap diorama of a cafe adorned with indoor plants. Wooden beams crisscross above, and a cold brew station stands out with tiny bottles and glasses



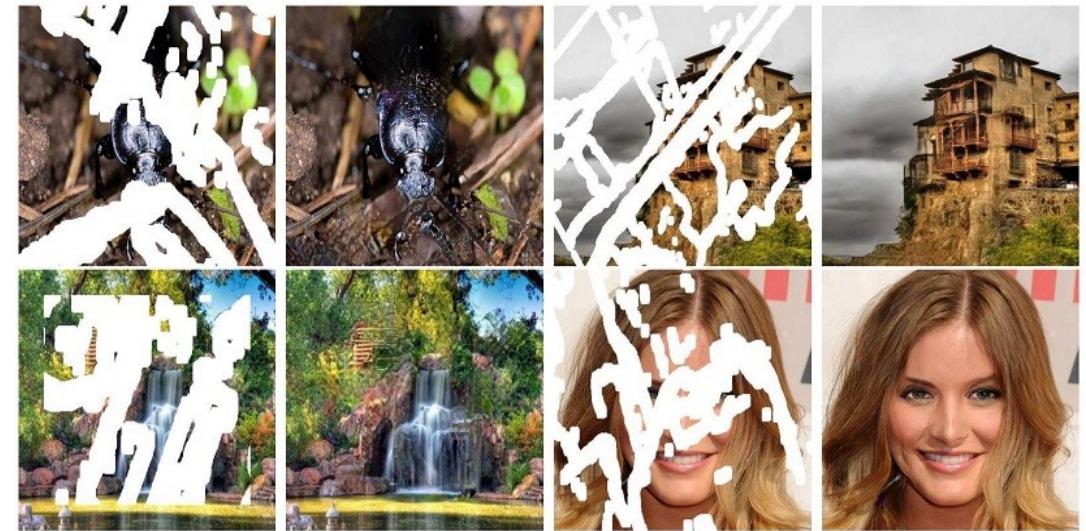
# Progreso en problemas inversos

$P(\text{high resolution} \mid \text{low resolution})$



Menon et al, 2020

$P(\text{full image} \mid \text{mask})$



Liu al, 2018

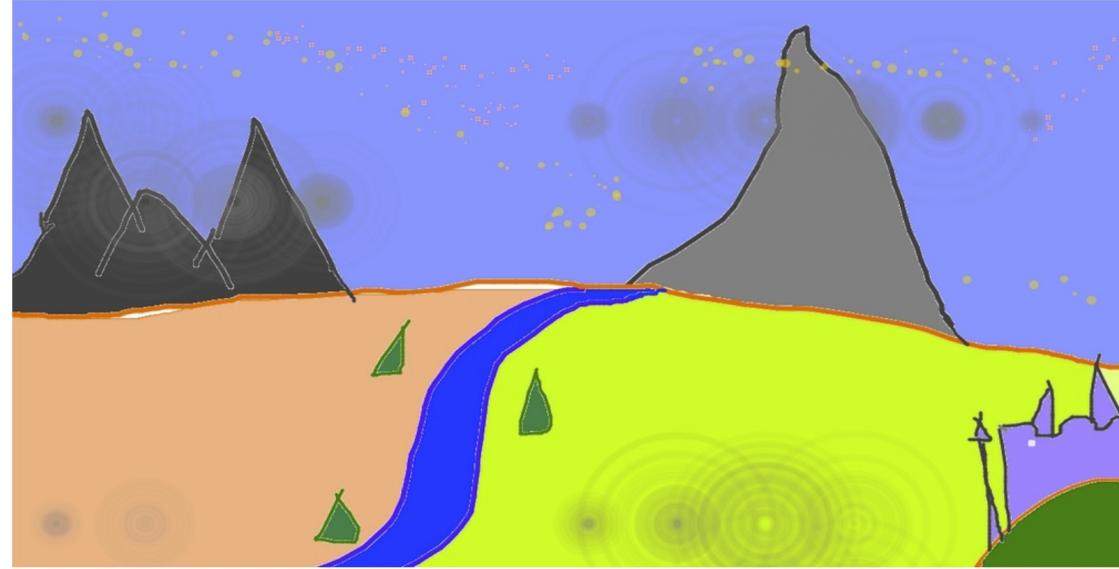
$P(\text{color image} \mid \text{greyscale})$



Antic, 2020

# Progreso en problemas inversos

User input:

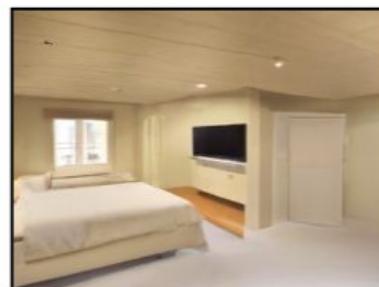
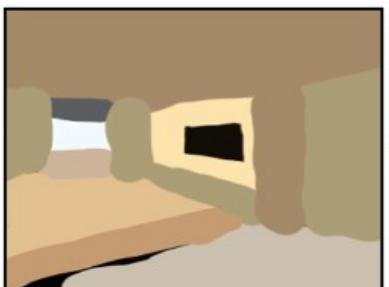


# Progreso en problemas inversos

Stroke Painting to Image



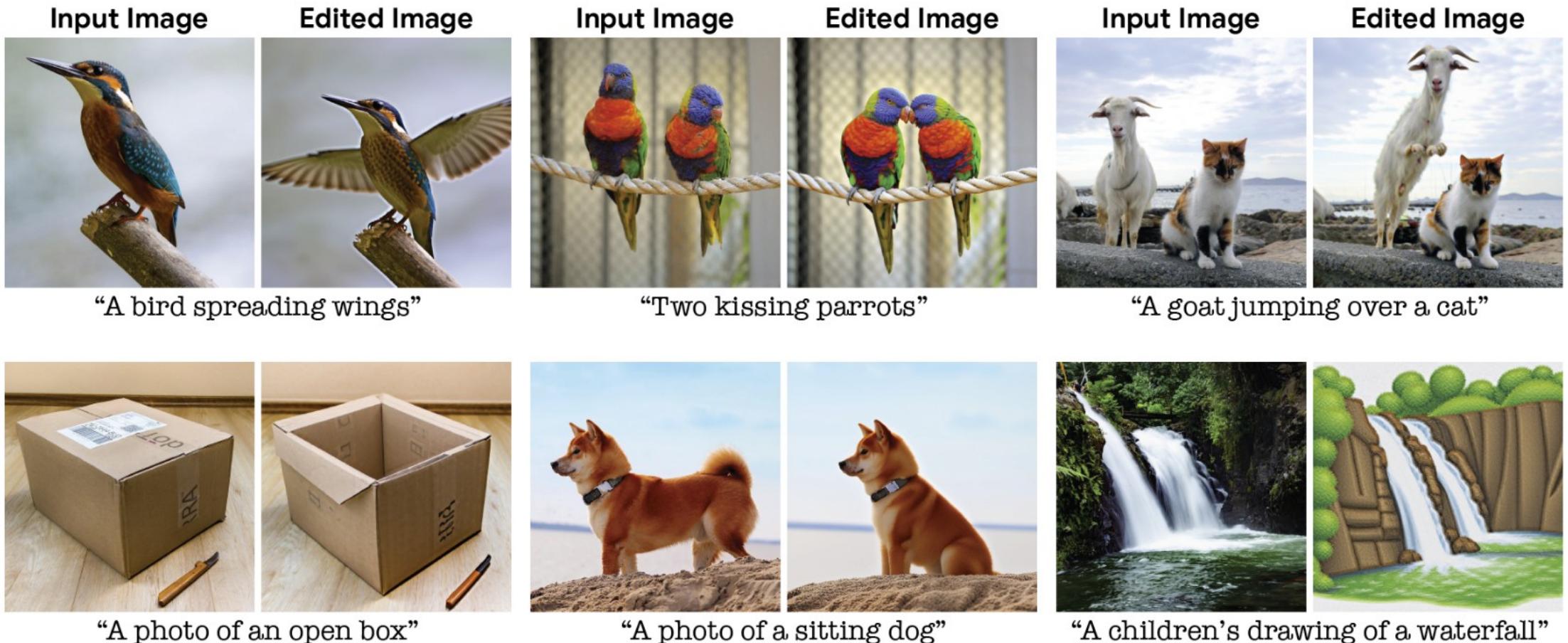
Stroke-based Editing



Input

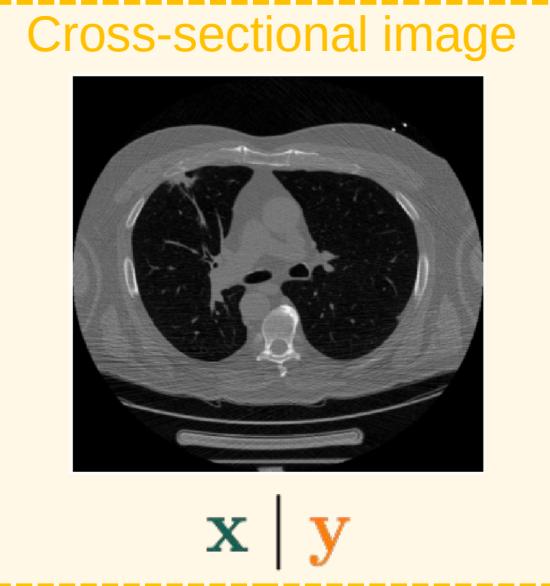
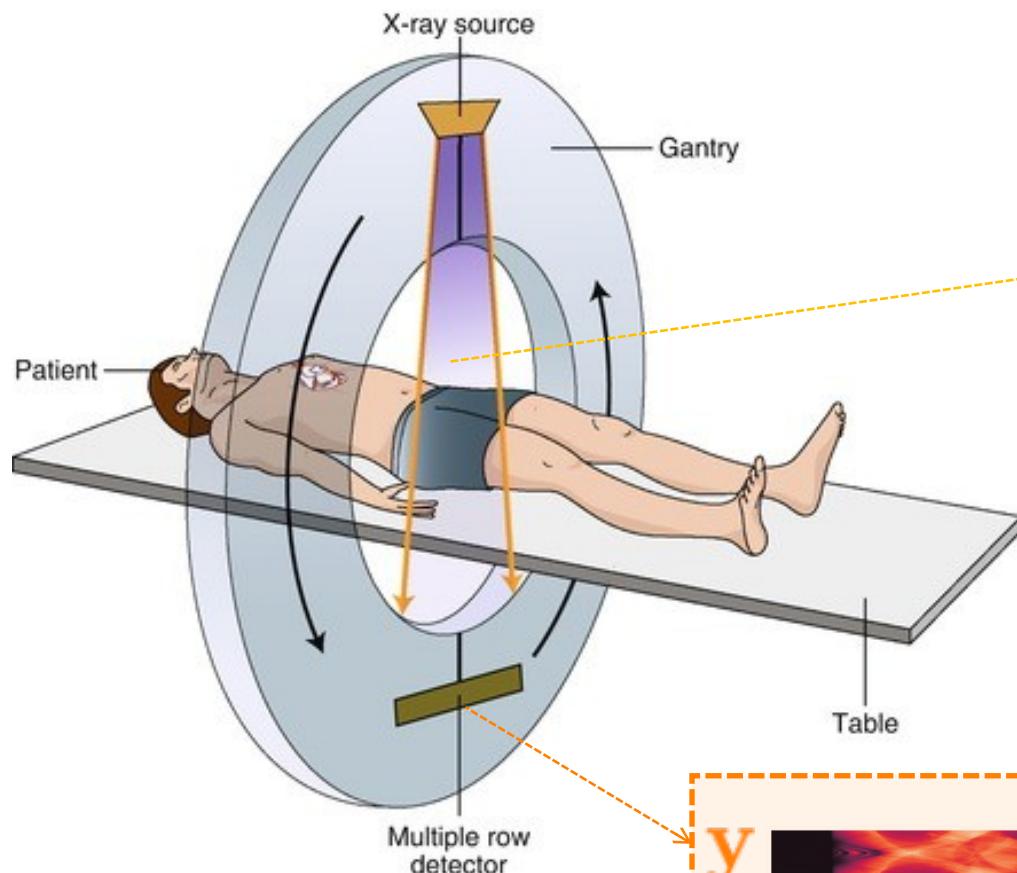
Output

# Progreso en problemas inversos



Kawar et al., 2023

# Reconstrucción de imágenes médicas

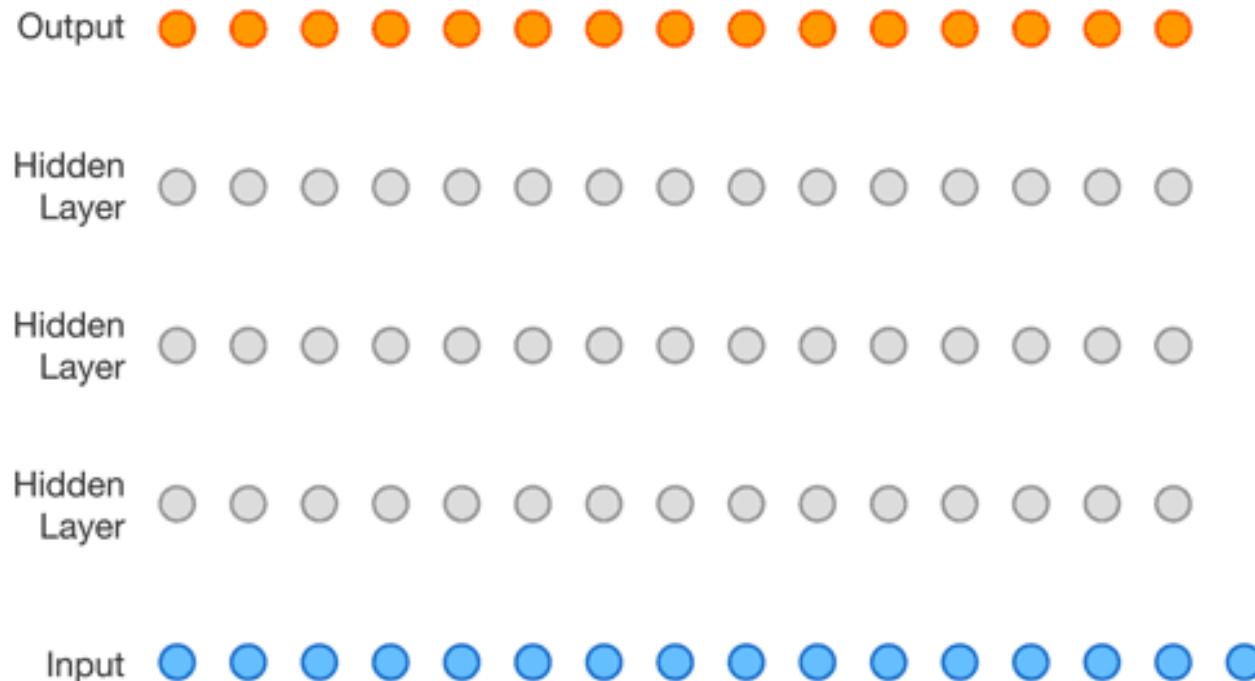


Sparse-view  
computed  
tomography  
(CT)

Modelo directo  $p(y | x)$  dado por simulación física

# WaveNet

## Modelo generativo de audio



### Text to Speech



Parametric



Concatenative



WaveNet



Unconditional



Music

van den Oord et al, 2016c

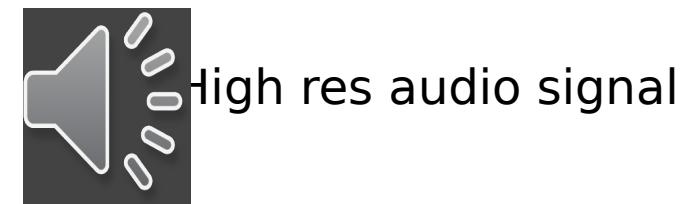
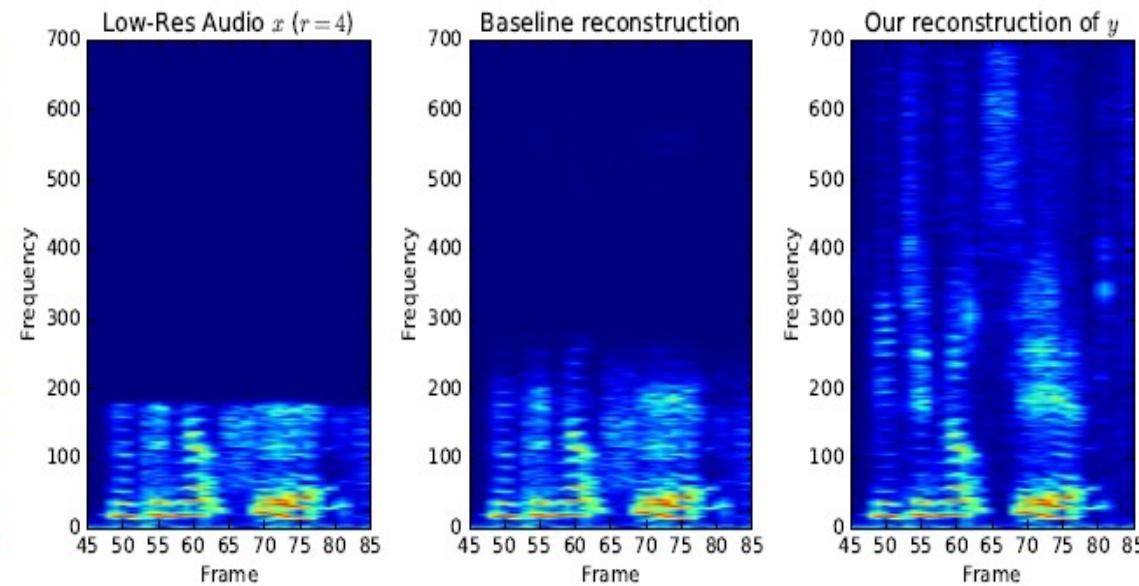
# Diffusion Text2Speech

Modelo generativo de audio



# Superresolución de audio

Modelo generativo condicional  $P(\text{high-res signal} \mid \text{low-res audio signal})$



Kuleshov et al., 2017

# Generación de texto

Custom prompt

To get an A+ in deep generative models, students have to

## Completion

**To get an A+ in deep generative models, students have to** be willing to work with problems that are a whole lot more interesting than, say, the ones that most students work on in class. If you're a great student, the question above can be avoided and you'll be able to do great work, but if you're not, you will need to go beyond the basics before getting good.

Now to be clear, this advice is not just for the deep-learning crowd; it is good advice for any student who is taking his or her first course in machine learning.

The key point is that if you have a deep, deep brain of a computer scientist, that's just as important to you.

P(next word | previous words)

Radford et al., 2019  
Demo from [talktotransformer.com](http://talktotransformer.com)

# Generación de texto - ChatGPT

ER

What should I do to get an A+ in CS236 at Stanford?

To get an A+ in CS236 (Deep Generative Models) at Stanford, you will need to excel in both your understanding of the material and your performance in assignments and exams. Here are some general tips to help you achieve this:

1. **\*\*Attend Lectures and Engage Actively\*\*:** Attend all lectures and actively engage with the material. Take thorough notes, ask questions, and participate in discussions. This will help you understand the concepts better.
2. **\*\*Read the Assigned Material\*\*:** Make sure to read the assigned textbooks, papers, and supplementary materials. Understanding the theoretical foundations is crucial.
3. **\*\*Stay Organized\*\*:** Keep a well-organized notebook or digital notes. This will help you quickly review and understand the material.
4. **\*\*Seek Help When Needed\*\*:** Don't hesitate to ask questions if you're having trouble with a concept. You can ask the professor, teaching assistants, or classmates for clarification.
5. **\*\*Complete Assignments Thoroughly\*\*:** Take your time to complete assignments, ensuring you fully understand the requirements and concepts involved. Start early and seek help if you're stuck. High-quality assignments are often a significant portion of your grade.
6. **\*\*Collaborate, but Don't Plagiarize\*\*:** Collaboration is often encouraged, but make sure you understand your institution's policy on collaboration and plagiarism. Always give credit where it's due and submit original work.
- ...
15. **\*\*Meet with the Professor\*\*:** If you're aiming for an A+, consider scheduling meetings with the professor to discuss your progress and seek feedback.

Remember that getting an A+ can be highly competitive, and the specific grading criteria may vary from one course to another and one professor to another. It's essential to understand the grading policies and expectations of your instructor. Always aim for excellence, but also keep in mind

# Machine Translation

Modelo generativo condicional  $P(\text{English text} | \text{Chinese text})$

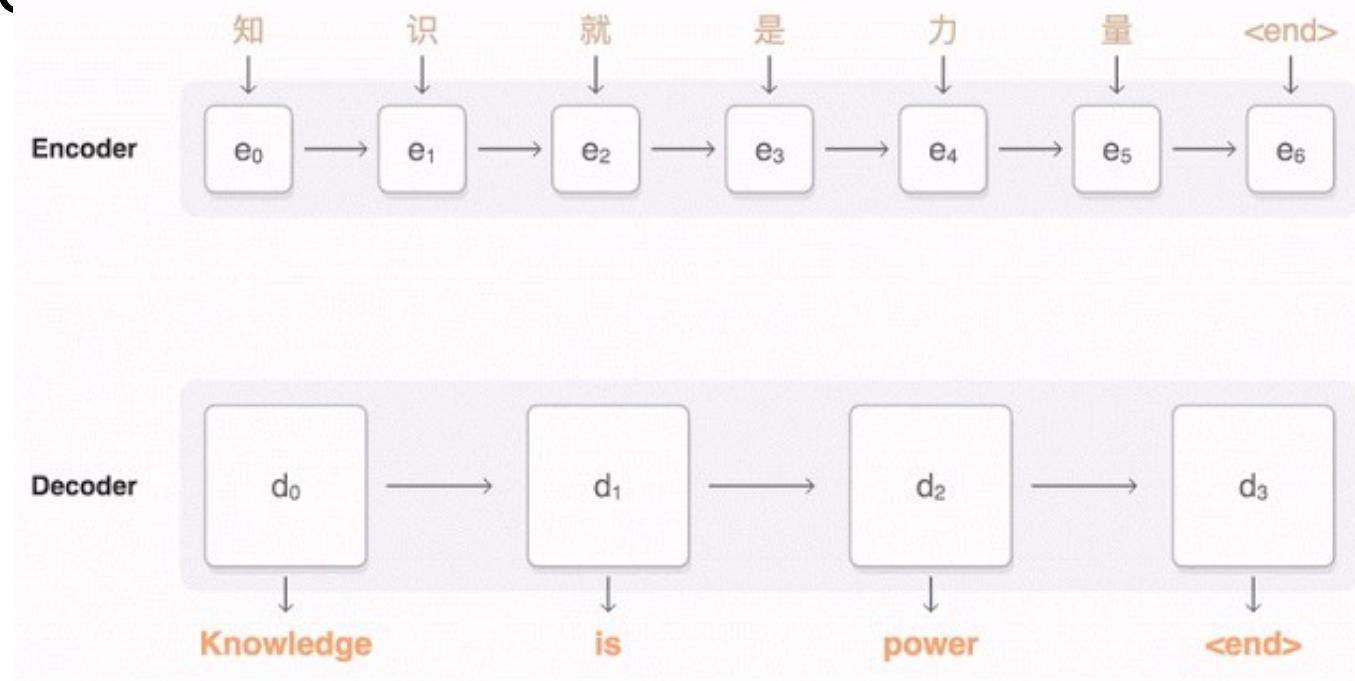
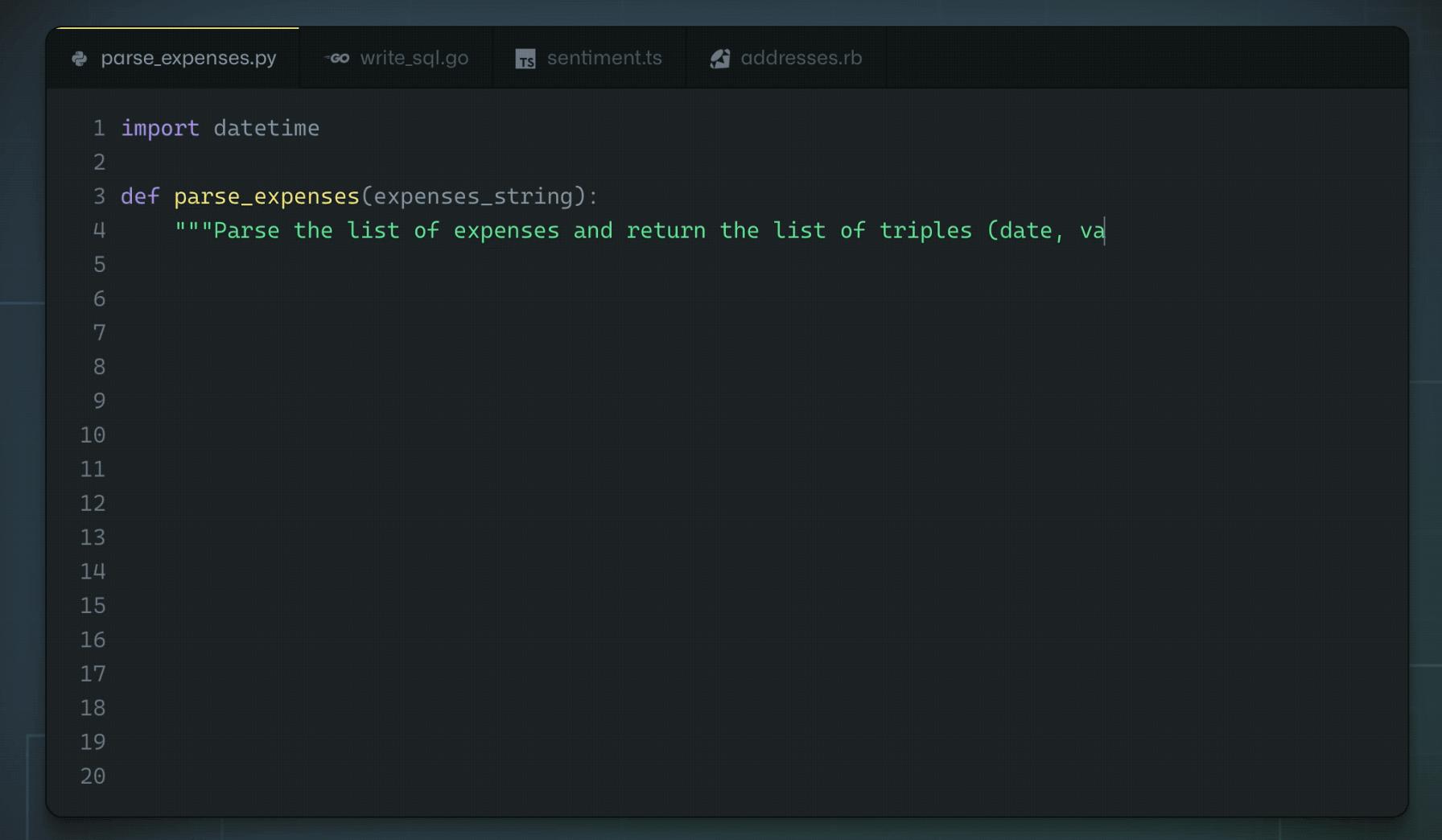


Figura de Google AI research blog.

# Generación de código



A screenshot of a code editor interface showing a Python file named `parse_expenses.py`. The code defines a function `parse_expenses` that takes a string of expenses and returns a list of triples (date, value, category). The code is numbered from 1 to 20. The code editor has a dark theme with syntax highlighting.

```
1 import datetime
2
3 def parse_expenses(expenses_string):
4     """Parse the list of expenses and return the list of triples (date, va|
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
```

# Generación de video

Suddenly, the walls of the embankment broke and there was a huge flood



# Generación de video

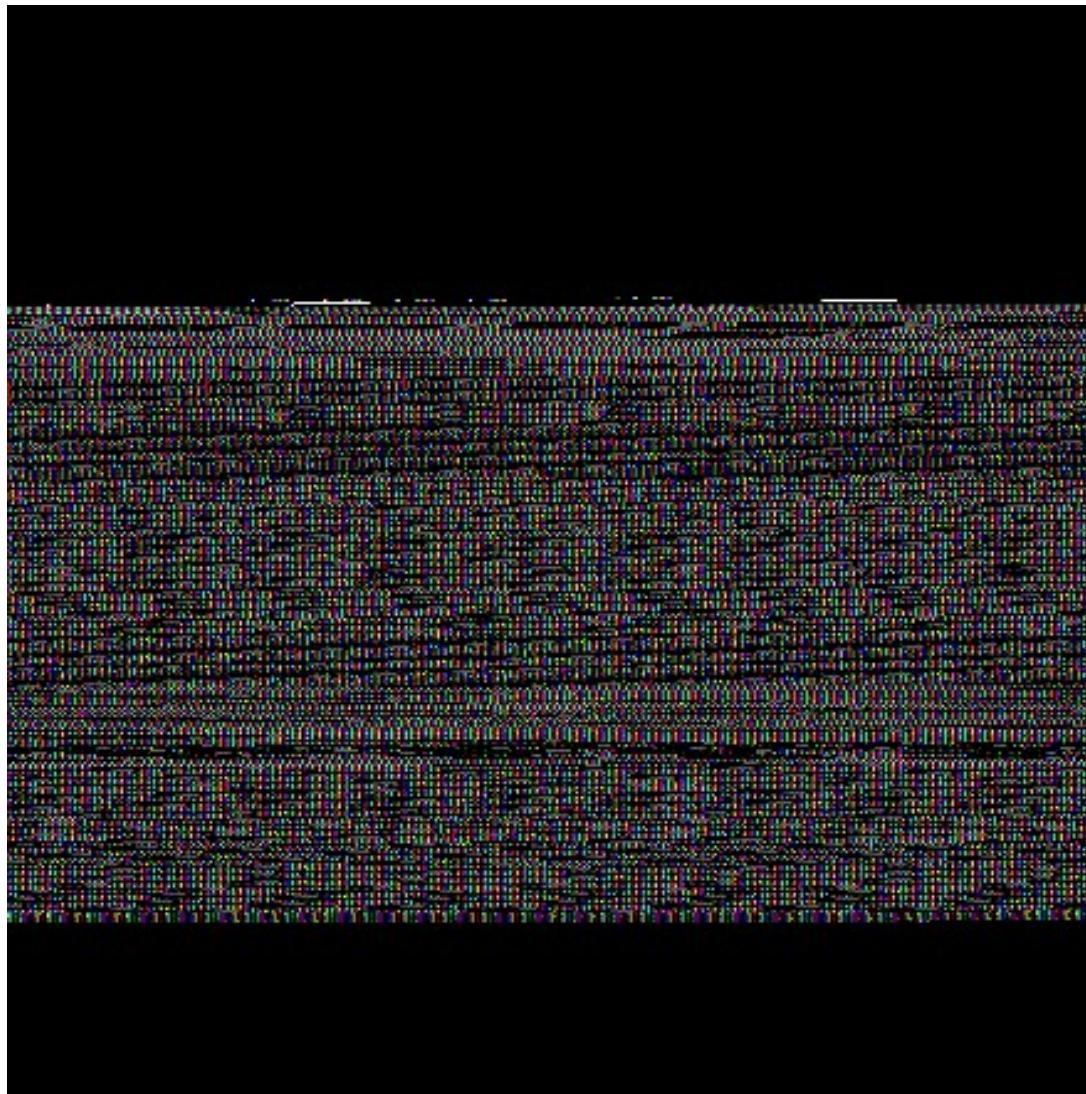
A couple sledding down a snowy hill  
on a tire roman chariot style



# Generación de video



# Generación de moléculas



# DeepFakes

Qué imagen es real?

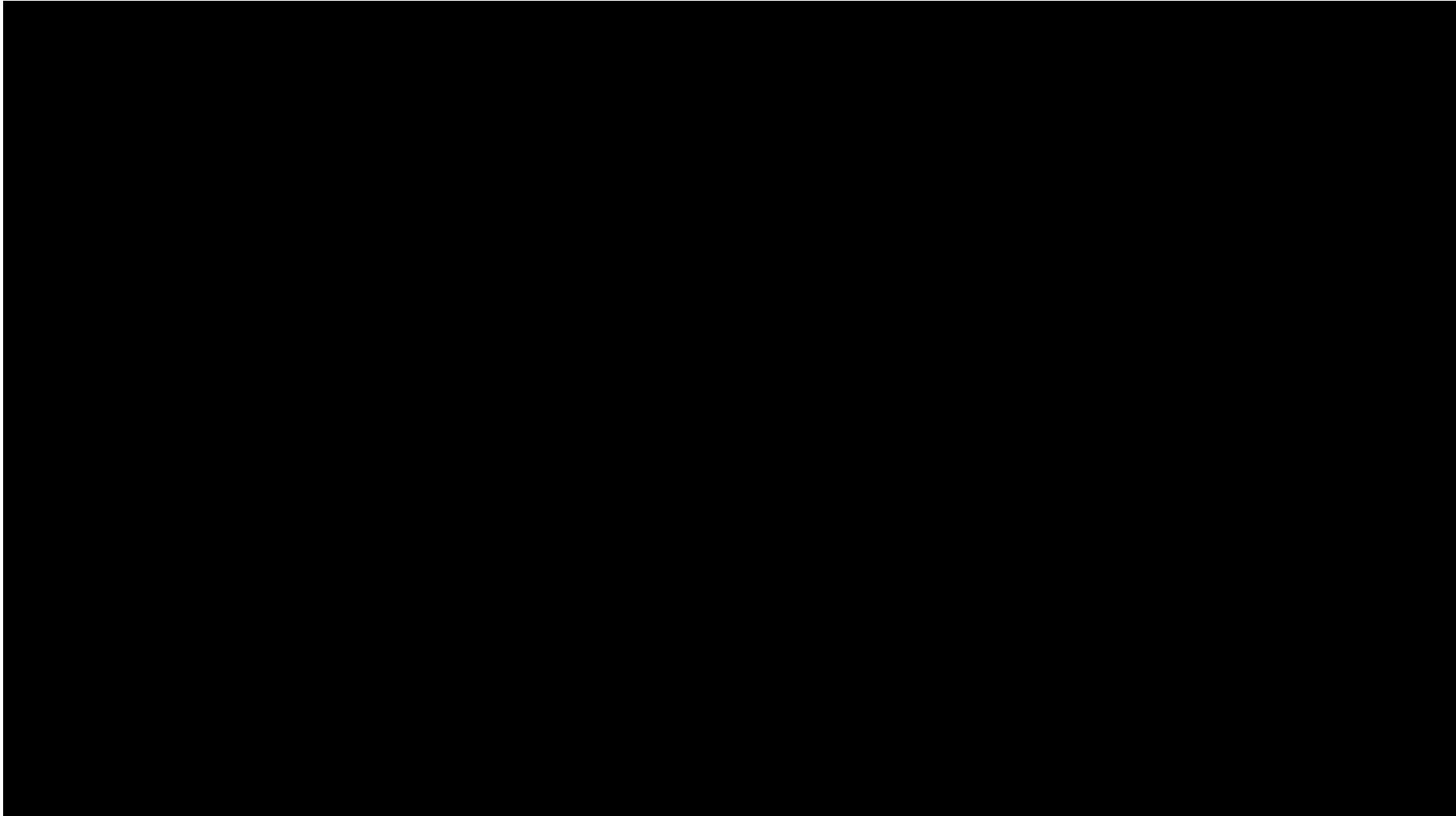


User  
 @StefanoErmon



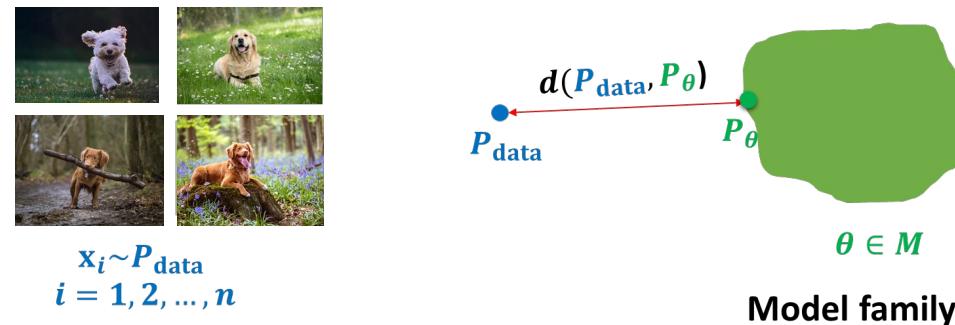
Output

# DeepFakes



# Hoja de ruta y desafíos

- **Representación:** ¿Cómo modelamos la distribución conjunta de muchas variables aleatorias?
  - Necesidad de una representación compacta
- **Aprendizaje:** ¿Cuál es la forma correcta de comparar distribuciones de probabilidad?



- **Inferencia:** ¿Cómo invertimos el proceso de generación? (e.g., visión como computación gráfica inversa)?
  - Aprendizaje no supervisado: inferir descriptores de alto nivel (características) a partir de los datos