



# Construcción de estados del arte con Inteligencia Artificial (IA)

Wilfrido Gómez  
Social Data IBERO



Estudios en Ingeniería  
para la Innovación

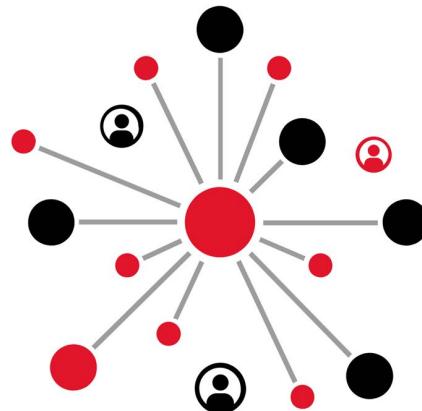


Instituto de Investigaciones para  
el Desarrollo de la Educación



E Q U I D E

Investigación con Impacto Social



# Social Data IBERO



Estudios  
Internacionales /



Formación e Incidencia Ignacianas /  
Programa de Derechos Humanos /



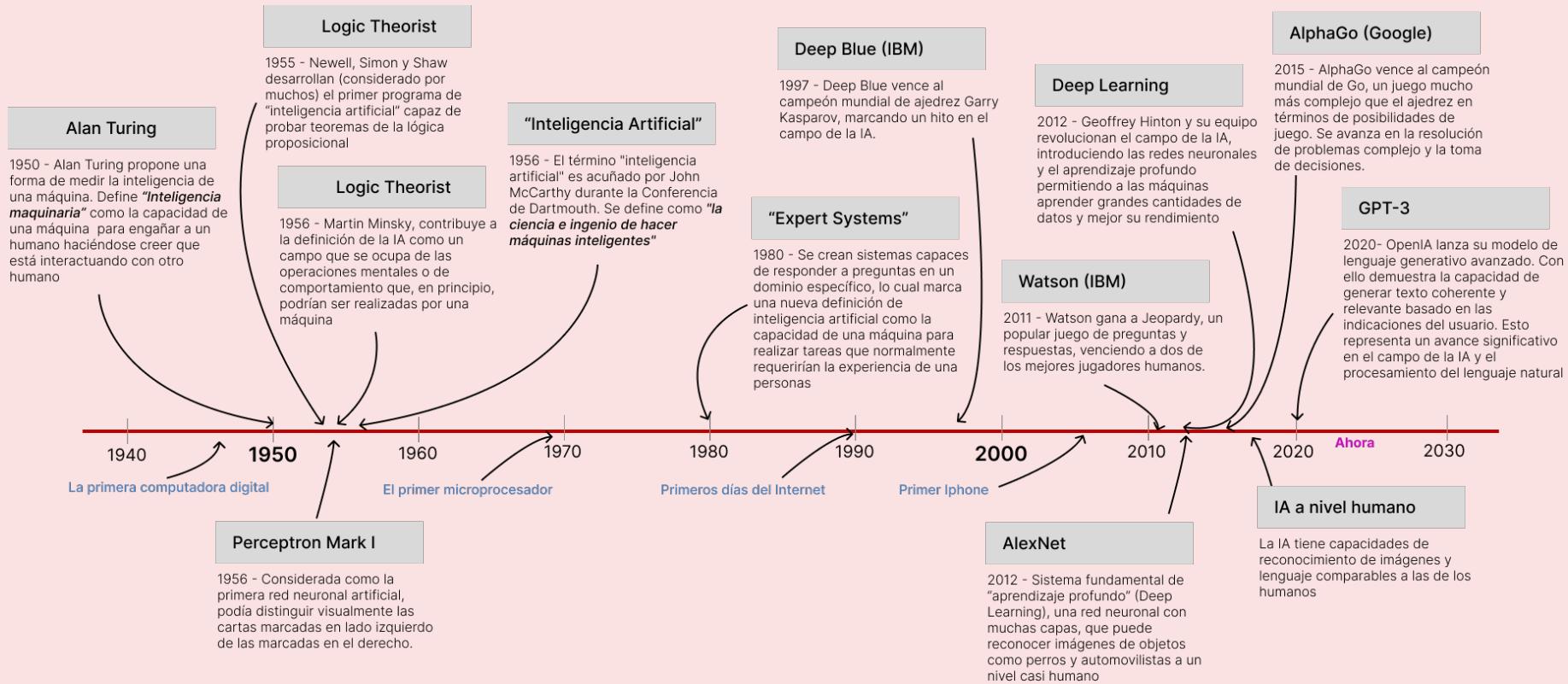
CENTRO TRANSDISCIPLINAR  
UNIVERSITARIO PARA  
LA SUSTENTABILIDAD

- 1. Breve historia de la Inteligencia Artificial**
- 2. ¿Qué es y como funciona la Inteligencia Artificial?**
- 3. Revisión de la Literatura impulsadas por Inteligencia Artificial**
- 4. Debates Actuales**



# Breve historia de la Inteligencia Artificial (IA)

# Línea del tiempo



# Language and image recognition capabilities of AI systems have improved rapidly

Test scores of the AI relative to human performance

+20

0 ← Human performance, as the benchmark, is set to zero.

-20

-40

-60

-80

-100

2000

2005

2010

2015

2020

The capability of each AI system is normalized

to an initial performance of -100.

Data source: Kiela et al. (2021) – Dynabench: Rethinking Benchmarking in NLP

OurWorldinData.org – Research and data to make progress against the world's largest problems.

Licensed under CC-BY by the author Max Roser

Handwriting recognition

Speech recognition

Image recognition

Reading comprehension

Language understanding

AI systems perform better than the humans who did these tests

AI systems perform worse

# Insertar página web

Esta aplicación permite insertar páginas web seguras que comiencen con https:// en la lista de diapositivas. Por motivos de seguridad, no se admiten páginas web no seguras.

Escriba la dirección URL a continuación.

https://

ourworldindata.org/grapher/artificial-intelligence-training-computation

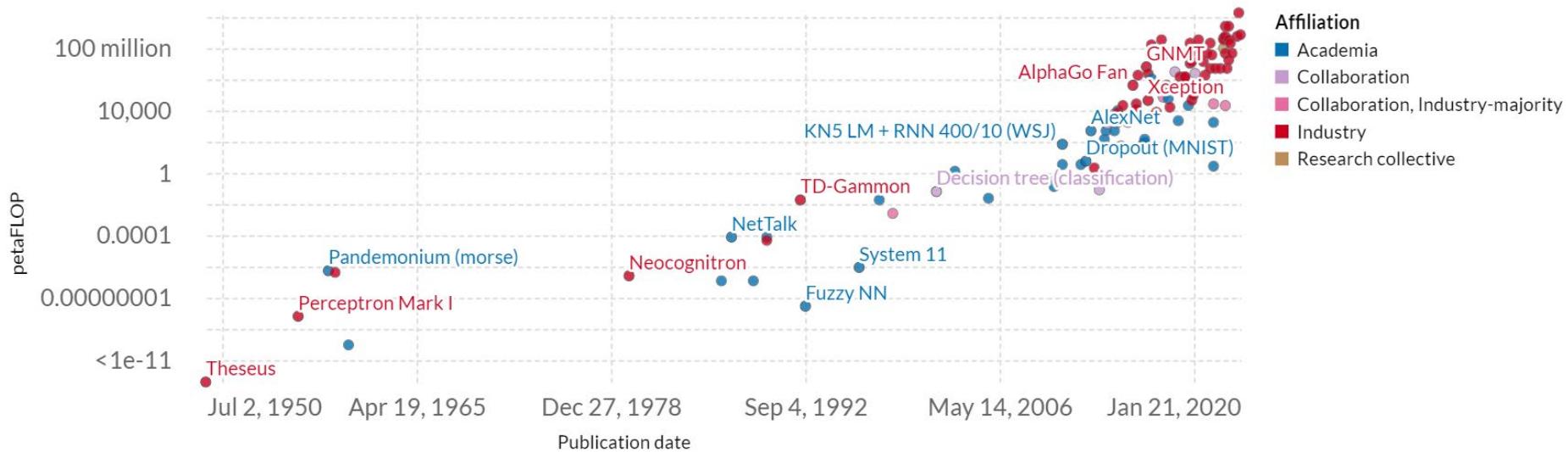
Nota: muchos sitios web famosos permiten el acceso seguro. Haga clic en el botón de vista previa para asegurarse de que la página web es accesible.

# Computation used to train notable AI systems, by affiliation of researchers

Computation is measured in total petaFLOP, which is  $10^{15}$  floating-point operations.

LINEAR   LOG

Select systems



Source: Sevilla et al. (2023)

Note: A research collective is a group of AI researchers who are not organized under an academic or industry affiliation; e.g., [EleutherAI](#). Computation is estimated based on published results in the AI literature and comes with some uncertainty. The authors expect the estimates to be correct within a factor of 2.

[OurWorldInData.org/artificial-intelligence](#) • CC BY

Jul 2, 1950

May 10, 2023

CHART

TABLE

SOURCES

DOWNLOAD

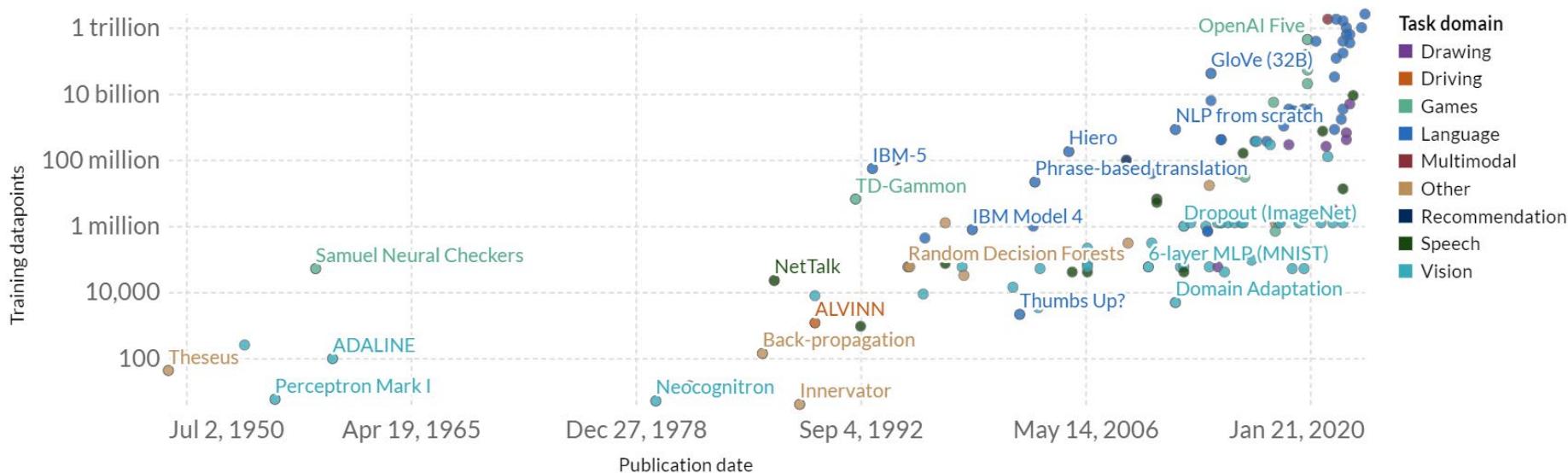


# Number of datapoints used to train notable artificial intelligence systems

Each domain has a specific data point unit; for example, for vision it is images, for language it is words, and for games it is timesteps. This means systems can only be compared directly within the same domain.

LINEAR  LOG

 Select systems



Source: Sevilla et al. (2023)

OurWorldInData.org/artificial-intelligence • CC BY

 Jul 2, 1950

 May 10, 2023

CHART

TABLE

SOURCES

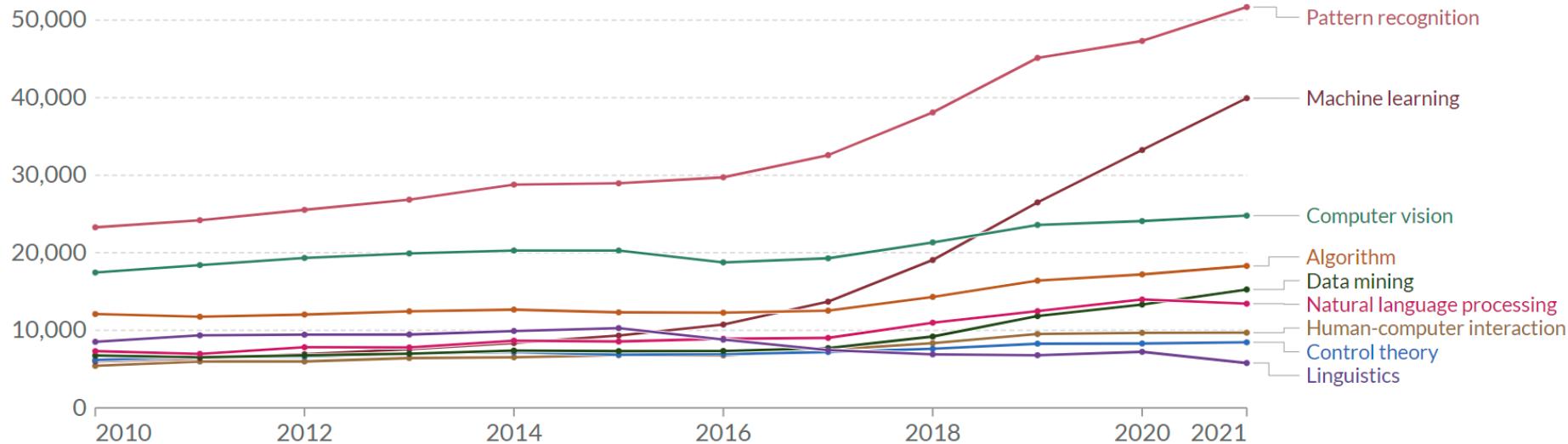
 DOWNLOAD



# Annual number of scholarly publications on AI, by field of study

English-language scholarly publications related to the development and application of artificial intelligence. This includes journal articles, conference papers, repository publications (such as arXiv), books, and theses.

All together ▾



Source: Center for Security and Emerging Technology via AI Index Report (2022)

Note: The reported number of publications in 2021 may be lower than the actual count due to lags in data collection.

OurWorldInData.org/artificial-intelligence • CC BY

► 2010

2021

CHART

TABLE

SOURCES

DOWNLOAD



# Timeline of images generated by artificial intelligence

These people don't exist. All images were generated by artificial intelligence.

2014



Goodfellow et al. (2014) – Generative Adversarial Networks

2015



Radford, Metz, and Chintala (2015) – Unsupervised Representation Learning with Deep Convolutional GANs

2016



Liu and Tuzel (2016) – Coupled GANs

2017



Karras et al. (2017) – Progressive Growing of GANs for Improved Quality, Stability, and Variation

2018



Karras, Laine, and Aila (2018) – A Style-Based Generator Architecture for Generative Adversarial Networks

2019



Karras et al. (2019) – Analyzing and Improving the Image Quality of StyleGAN

2020



Ho, Jain, & Abbeel (2020) – Denoising Diffusion Probabilistic Models

2021



Ramesh et al. (2021) – Zero-Shot Text-to-Image Generation (OpenAI's DALL-E 1)

2022



Saharia et al. (2022) – Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding (Google's Imagen)

## Explaining a Joke

I will explain these jokes:

(1) The problem with Kleptomaniacs is that they always take things literally.

Explanation: This joke is wordplay. Someone who "takes things literally" is someone who doesn't fully understand social cues and context, which is a negative trait. But the definition of kleptomania is someone who literally takes things.

(2) A pessimist is one who expects the worst.

Explanation: Most people expect you to pay them back when you borrow money, however a pessimist is someone who always assumes the worst, so if you borrow money from them, they will expect that you won't pay them back anyways.

**Input:** I was supposed to start writing the paper at 5:00 PM. But then I started playing with this cool new language model for 10 minutes. 10 minutes later, it's suddenly 9:30 PM!

**Model Output:** This joke is about how time flies when you're having fun. The person was supposed to start writing a paper at 5:00 PM, but they started playing with a language model instead. They thought they were only playing with it for 10 minutes, but it was actually 4.5 hours.

**Input:** I was going to fly to visit my family on April 6th. My mom said "Oh great, your step dad's poetry reading is that night!" So now I'm flying in on April 7th.

**Model Output:** The joke is that the speaker's mother is trying to get them to go to their step dad's poetry reading, but the speaker doesn't want to go, so they are changing their flight to the day after the poetry reading.

**Input:** Reading "War and Peace" doesn't scare me. I once had to review the CL for my intern's whole project on his final day!

**Model Output:** 'War and Peace' is a very long book, so reading it is a daunting task. However, the speaker is saying they had to review the code for their intern's project on their final day, which is even more daunting.

**Input:** I tried 10,000 random restarts of my neural network, but I was accused of overfitting. I guess no good seed goes unpunished.

**Model Output:** This joke is a pun. A neural network is a computer program that can learn from data. A "seed" is a number that is used to initialize a random number generator. A "good seed" is a number that produces a good random number generator. The phrase "no good deed goes unpunished" means that if you do something good, you will be punished for it.

**Input:** What's the difference between a zebra and an umbrella? One is a striped animal related to horses, another is a device you use to stop the rain from falling on you.

**Model Output:** This joke is an anti-joke. The joke is that the answer is obvious, and the joke is that you were expecting a funny answer.



▶ ChatGPT interpreter  
plugin demo

03:25

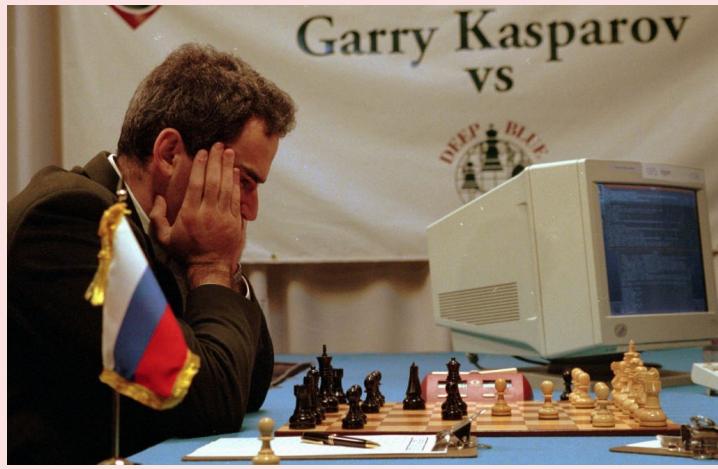
A screenshot of a video player interface. At the top, there's a thumbnail image of a person. Below it are standard video control buttons for play, pause, and volume. A large, semi-transparent text overlay covers the bottom half of the screen. The text reads "▶ ChatGPT interpreter plugin demo" in a large, white, sans-serif font, and "03:25" in a smaller font below it.

# Open Source





# ¿Qué es y cómo funciona la Inteligencia Artificial?

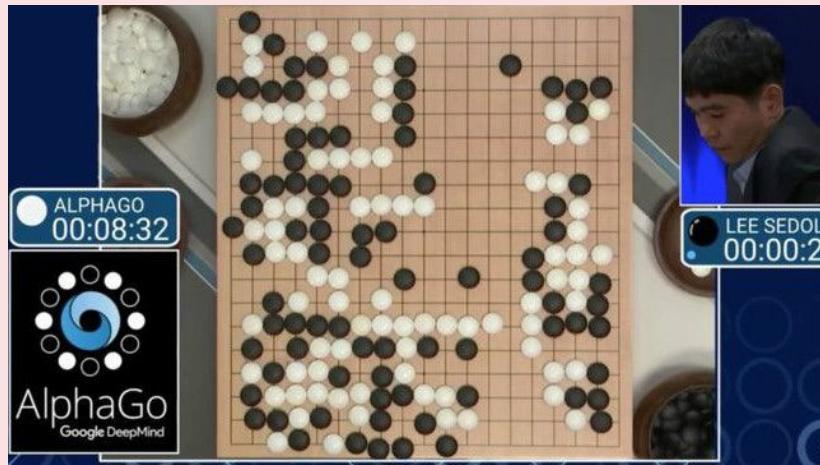


Una inteligencia artificial crea música al estilo de Bach (pero da matices sustanciales)

«DeepBach» es un sistema que ha creado una pieza similar a la forma algorítmica en la que componía el artista clásico



Lo más leído en Tecnología  
1 Bill Gates fue detenido en 1997 por conducir sin carnet: tu ordenador está plagado de sus fotos



GPT-3, el nuevo modelo de lenguaje de OpenAI, es capaz de programar, diseñar y hasta conversar sobre política o economía



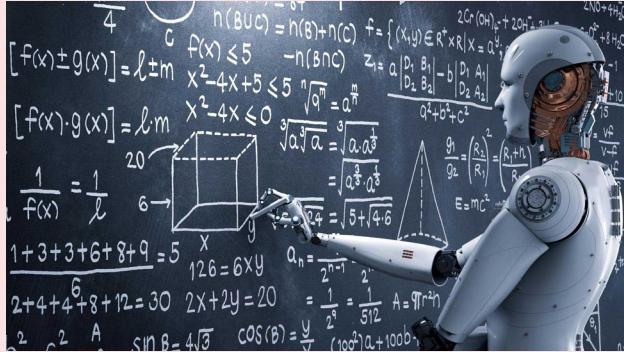
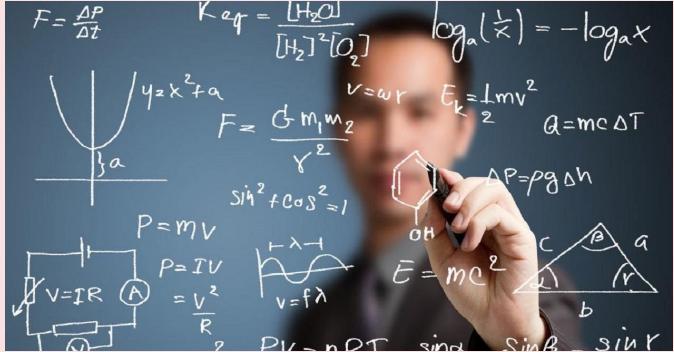
# ¿A qué nos referimos con Inteligencia Artificial?



# **Una posible definición corta:**

**“Es la capacidad de una máquina (por ejemplo, una computadora) de realizar tareas propias de una inteligencia humana”**

# ¿Tareas propias de una inteligencia humana?



# Avance de la tecnología



# Cosas que no estaban al alcance de la Inteligencia Artificial

Capacidad de Aprendizaje



La creatividad



La autoconciencia

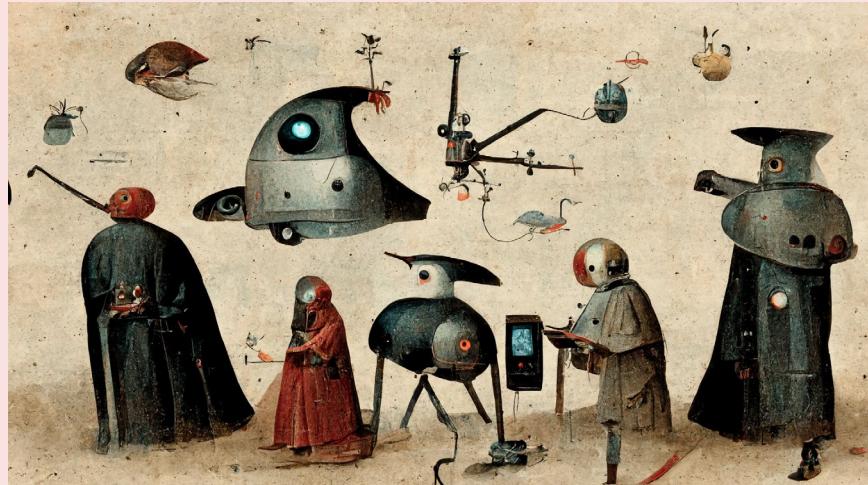


# ¿2023...?

**“Existen algoritmos a los que podemos atribuir cierto tipo de creatividad”**



Midjourney

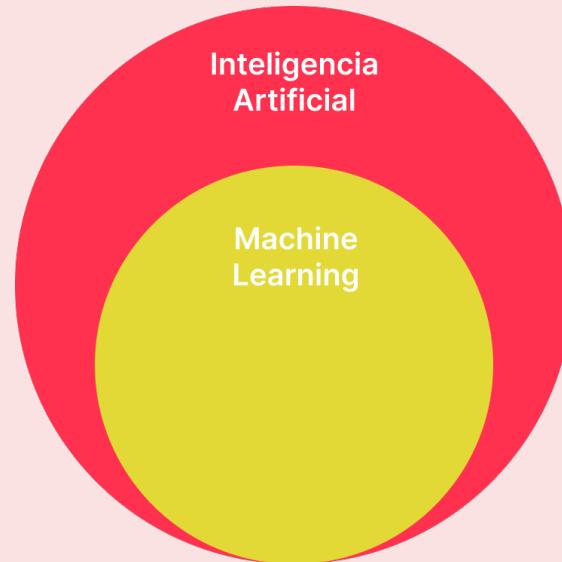


# **El aprendizaje en la Inteligencia Artificial**

**“Cuando escuchamos hablar de que la Inteligencia Artificial ha hecho alguna cosa, a menudo se está en el terreno del aprendizaje”**

# Aprendizaje Automático o “Machine Learning”

Hoy en día Machine Learning es prácticamente sinónimo de Inteligencia Artificial



# ¿Qué es?

**Un conjunto de técnicas en el que un algoritmo que tiene que revisar alguna tarea es capaz de modificar su propio comportamiento basándose en los datos del cual dispone, de lo bien o mal de que lo haya hecho en el pasado, o lo que otros digan que está bien o mal.**

**¡Aprenden de sus errores!**



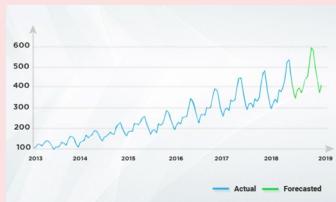
# Aprendizaje Supervisado



“Mono”

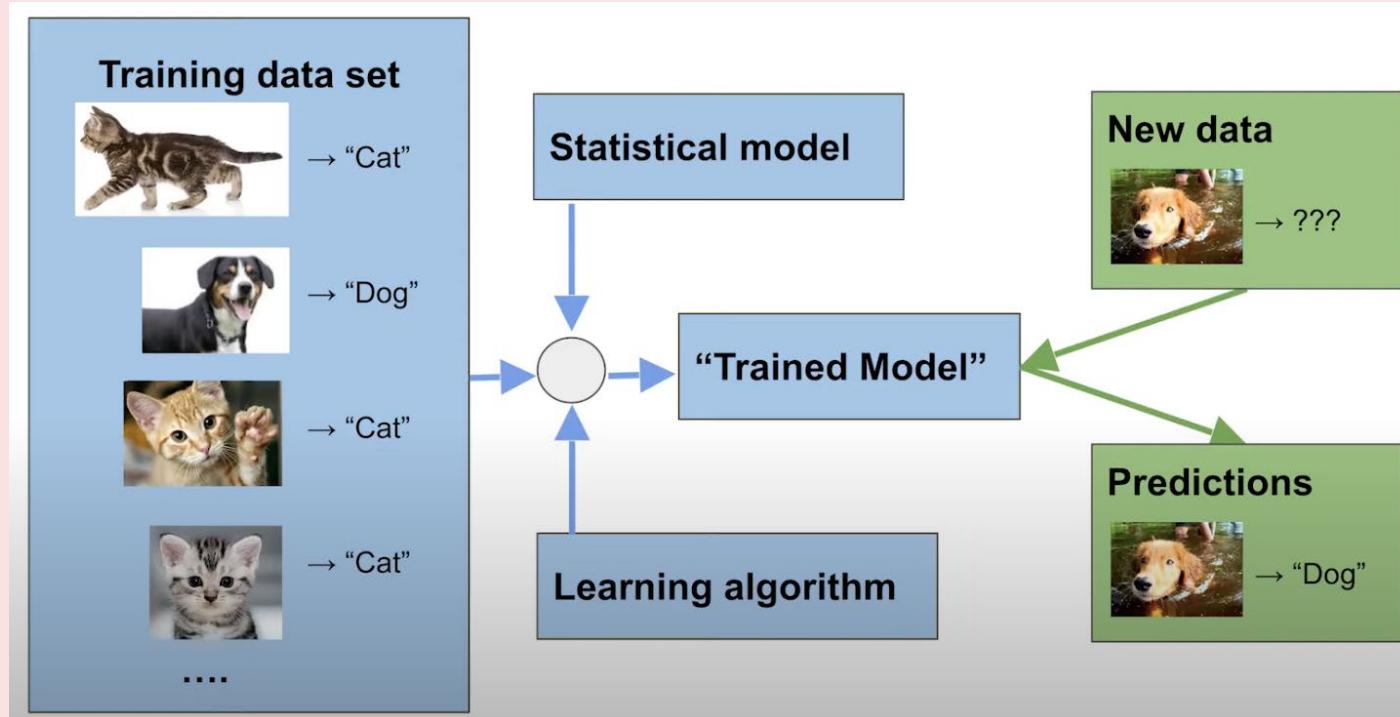


“Te felicito, qué bien actúas”



“Tendencia futura”

# Aprendizaje Supervisado



Fundamentals of Machine Learning: Types of Machine Leraning. Complexity Explorer.

# **Los problemas que resuelven estos algoritmos**

**Reconocimiento facial**

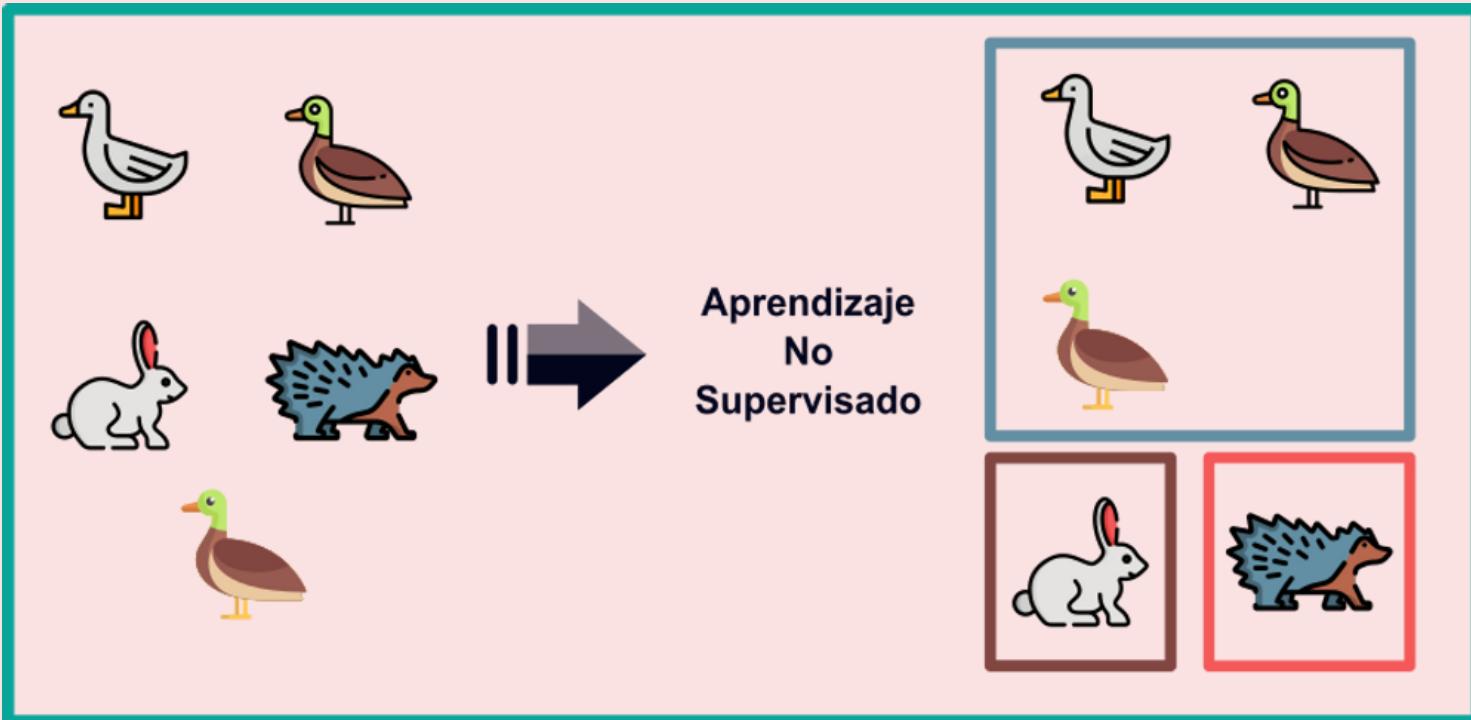
**Reconocimiento de voz**

**Reconocimiento de huellas digitales**

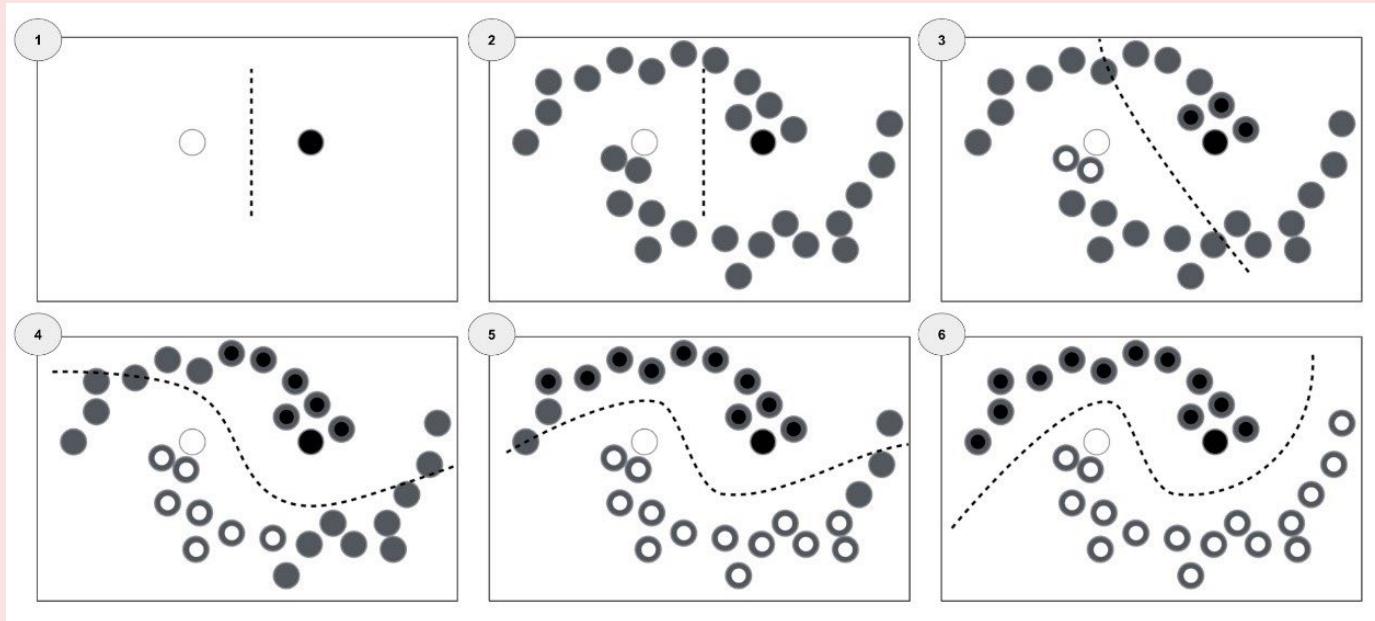
**Coches automáticos**

...

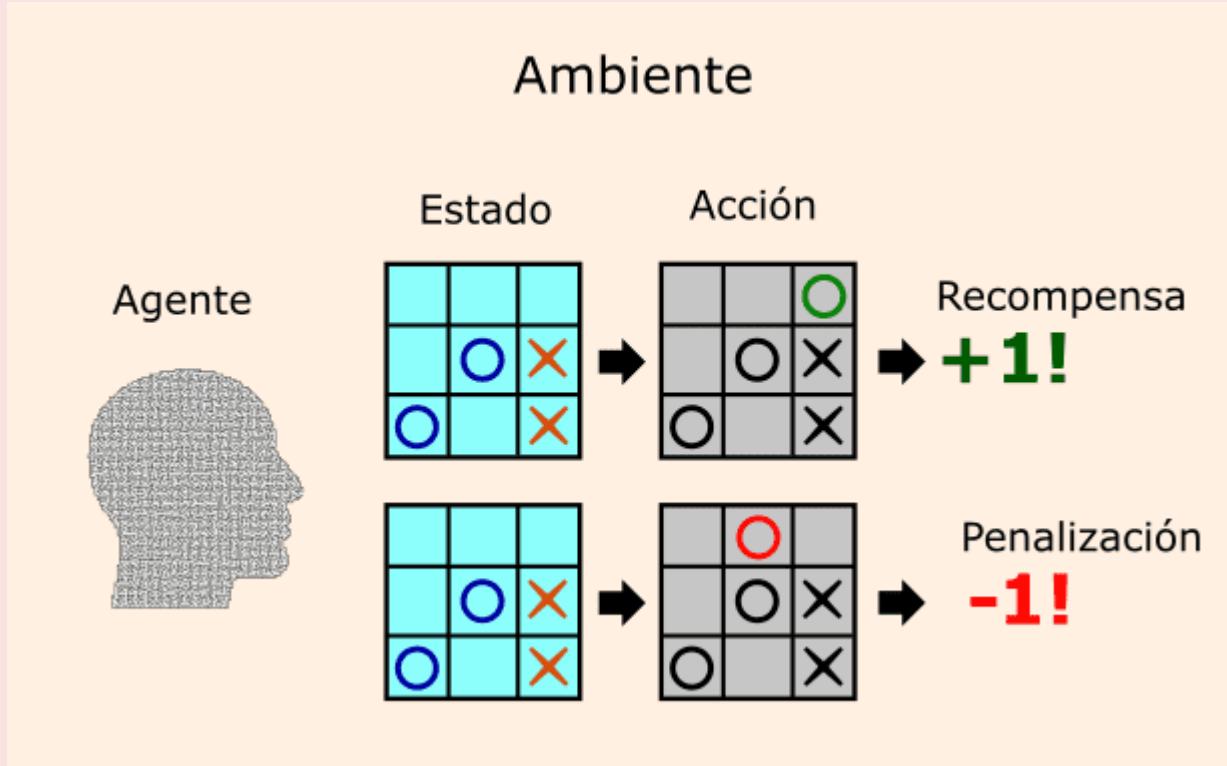
# Aprendizaje No Supervisado



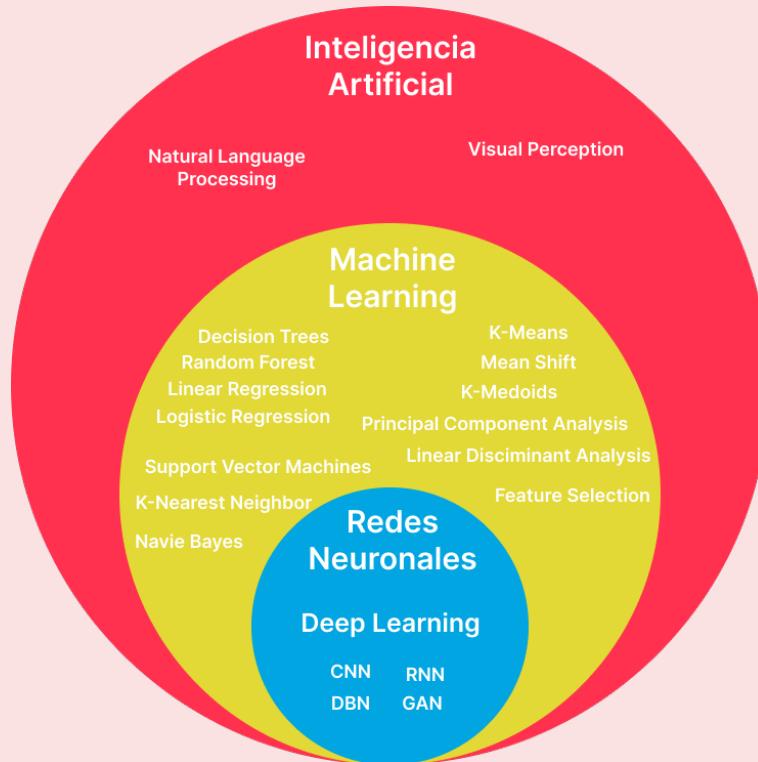
# Aprendizaje Semi-Supervisado



# Aprendizaje por refuerzo



# Para realizar estos tipos de aprendizaje hay muchas técnicas





# Revisión de Literatura impulsada por Inteligencia Artificial



# ¿Qué significa hacer una revisión de la literatura?

Primero tenemos que entender el propósito de la revisión, luego colocar el tema nuestra investigación en el contexto del trabajo existente y finalmente configurarlo para demostrar que se está haciendo una nueva contribución.

# **Para ello debemos realizar diferentes tipos de lectura**

1. Lectura exploratoria, realizada principalmente en la fase inicial
2. Lectura profunda de fuentes altamente informativas
3. Lectura amplia y específica de grandes colecciones de libros y artículos

# 1. Lectura exploratoria

Es una etapa que puede ocurrir antes o después de formular el tema o la pregunta de investigación. Nos ayuda a explorar y digerir las ideas principales en la investigación existente.

A menudo no está estructurado y está lleno de descubrimientos fortuitos. Aquí obtenemos lo que encontramos, en lugar de realizar una búsqueda específica de información particular.

Al final de esta fase debería ser capaz de esbozar un mapa aproximado de su campo de estudio

## **2. Lectura profunda de fuentes altamente informativas**

Sucede después de la fase de lectura exploratoria. Aquí se prioriza la información que lee.

Aquí nuestra atención se reducirá a un pequeño número de fuentes altamente relevantes, tal vez uno o dos libros, o un puñado de artículos, que leerá cuidadosamente, con el objetivo de comprender completamente los conceptos importantes.

A menudo va acompañado de pausas reflexivas y toma de notas significativas.

Esta fase nos ayuda a decidir qué cosas nos interesan más

### **3. Lectura amplia y específica de grandes colecciones de libros y artículos**

Al llegar aquí ya tiene un mapa de su campo de estudio (lectura exploratoria) y ha llenado algunas partes de este mapa con más detalles (lectura profunda).

Ha encontrado una pregunta o tema de investigación, ya sea por su cuenta o con la ayuda de un asesor.

Es donde la revisión de literatura comienza en serio.

Se revisa la literatura una vez más, pero esta vez de manera específica: está buscando piezas particulares de información que ilustren el trabajo existente o demuestren la necesidad del nuevo enfoque que tomará su investigación

### **3. Lectura amplia y específica de grandes colecciones de libros y artículos**

Desea encontrar todas las secciones de “metodología” en un grupo de artículos académicos y filtrar aquellos que tienen ciertos conceptos claves

Aquí se muestran las conexiones entre los conceptos más importantes en su campo

Es una fase lenta y a veces tediosa

¿qué pasa si te pierdes algo importante?

¿cómo aceleramos estos procesos?

# ¿Cómo aceleramos estos procesos?

Recurrir a soluciones basadas en computadoras

Una opción es crear sistemáticamente una lista clave de términos de búsqueda o frases clave, que luego se pueden combinar utilizando operadores booleanos para ampliar los resultados.

Ejemplo: “Obesidad” y “adolescentes”

Filtra artículos relacionados con la obesidad que mencionan adolescentes, pero no mencionan a los adultos mayores.

Esto puede ayudar a sacar a la superficie muchos artículos relevantes, pero hay algunas desventajas en esta estrategia

Estas consultas de palabras clave son complicadas y requieren mucho tiempo para crear.

A menudo, lo que desea encontrar son “trozos” completos de texto, párrafos o secciones, no solo palabras claves.

Una vez creada la lista de palabras claves booleanas ¿cómo sabe que no ha olvidado incluir una consulta de búsqueda importante?

Puede perder resultados que serían relevantes,

“excelencia deportiva” no coincidiría con un artículo que mencionara solo “atletismo de alto rendimiento”

Lluvia de ideas para  
preguntas de investigación

Elicit



Motores de búsqueda

EvidenceHunt



SEMANTIC SCHOLAR



Discovery  
By researcher.life

Búsqueda de literatura



ResearchRabbit



Plugin para Zotero

Inciteful

IRIS.AI

Revisión de papers

scholarcy



SCISPACE



Litmaps

CONNECTED PAPERS

Tomando notas



LATERAL



ChatGPT

jenni

Editor de notas

Paperpal  
By researcher.life

Citas  
scite  
Plugin para Zotero



# Debates actuales

**Opinion**  
Artificial intelligence (AI)

● This article is more than 1 month old

## AI machines aren't 'hallucinating'. But their makers are

*Naomi Klein*

Tech CEOs want us to believe that generative AI will benefit humanity. They are kidding themselves



Mon 8 May 2023 09.02 BST  
1,342

f t m

And their goal never was to solve climate change or make our governments more responsible or our daily lives more leisurely. Illustration: UliGrafik/Alamy

Inside the many debates swirling around the rapid rollout of so-called artificial intelligence, there is a relatively obscure skirmish focused on the choice of the word "hallucinate".

europapress / portaltic / sector

## La UE avanza en la regulación de la IA incluyendo pautas de transparencia y gestión de riesgo



Europa Press PortaltIC  
CORTES DE CASTILLA LA MANCHA - Archivo  
Publicado Jueves, 11 mayo 2023 17:17  
@portaltic

Go Karts super divertidos K1 Speed Garden Santa Fe  
Diversion Go Karts K1 Speed Garden Santa Fe

**The long read**

## The stupidity of AI

ograph: Rokas Tenys/Alamy

### Artificial intelligence in its current form is based on the wholesale appropriation of existing culture, and the notion that it is actually intelligent could be actively dangerous

by James Bridle

In January 2021, the artificial intelligence research laboratory OpenAI gave a limited release to a piece of software called Dall-E. The software allowed users to enter a simple description of an image they had in their mind and, after a brief pause, the software would produce an almost uncannily good interpretation of their suggestion, worthy of a jobbing illustrator or Adobe proficient designer – but much faster, and for free. Typing in, for example, “a pig with wings flying over the moon, illustrated by Antoine de Saint-Exupéry” resulted, after a minute or two of processing, in something reminiscent of the patchy but recognisable watercolour brushes of the creator of *The Little Prince*.

A year or so later, when the software got a wider release, the internet went wild. Social media was flooded with all sorts of bizarre and wondrous creations, an exuberant hodgepodge of fantasies and artistic styles. And a few months later it happened again, this time with language, and a product called ChatGPT, also produced by OpenAI. Ask ChatGPT to produce a



LA NACION > Sociedad

## “Si se produce el caos, será demasiado tarde”. Por qué la inteligencia artificial podría devorar rápidamente toda la cultura humana

Yuval Harari, Tristan Harris y Aza Raskin escriben sobre los peligros de esta tecnología también para los sistemas democráticos

29 de marzo de 2023 • 09:38

Yuval Harari, Tristan Harris y Aza Raskin



Continúa.  
Sigue creciendo.  
Obtén las apps de Creative Cloud.

Más info

# Invisible Indirect Injection: A Puzzle for ChatGPT

2023-06-01



If you have found this page by solving the puzzle, congratulations! This write-up is your reward. If you're in a rush and know the puzzle, skip to [How it works](#) below or check out the source code on [GitHub](#).

In this post, I will briefly explain how the puzzle works and what it means for the security of generative AI integrations. Raising awareness of these issues was the main goal. I also present new techniques that make indirect prompt injections invisible to end users.

## Abogados presentan en una corte una investigación jurídica... hecha con ChatGPT

Los abogados podrían ser sancionados debido al documento que incluyeron en una demanda contra una aerolínea, el cual hacía referencia a casos judiciales anteriores. Varios de ellos nunca sucedieron o implicaban a empresas que no existían.



La app de ChatGPT en la pantalla de un iPhone. Foto: AP / Richard Drew, Archivo

■ CIENCIA Y TECNOLOGÍA

Por AP

viernes, 9 de junio de 2023