



Fundamentos de la Inteligencia Artificial Generativa

MIKEL GALAR IDOATE, IRIS DOMÍNGUEZ CATENA

mikel.galar@unavarra.es , iris.dominguez@unavarra.es

upna arin

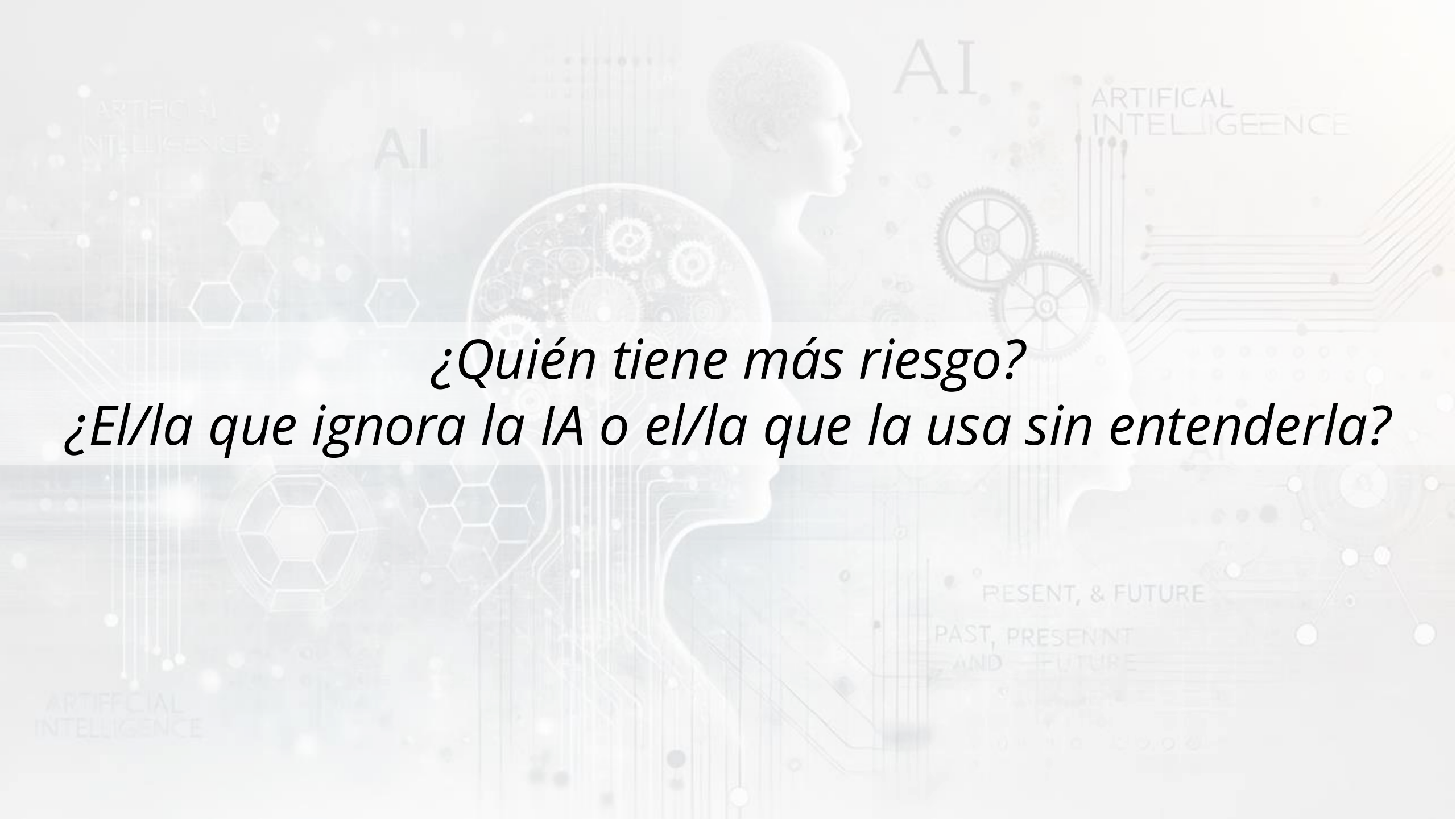
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Artificial Intelligence &
Machine Learning Research

Departamento de Estadística, Informática y Matemáticas

Institute of Smart Cities (ISC)

ARIN. Artificial Intelligence and Machine Learning Research



*¿Quién tiene más riesgo?
¿El/la que ignora la IA o el/la que la usa sin entenderla?*

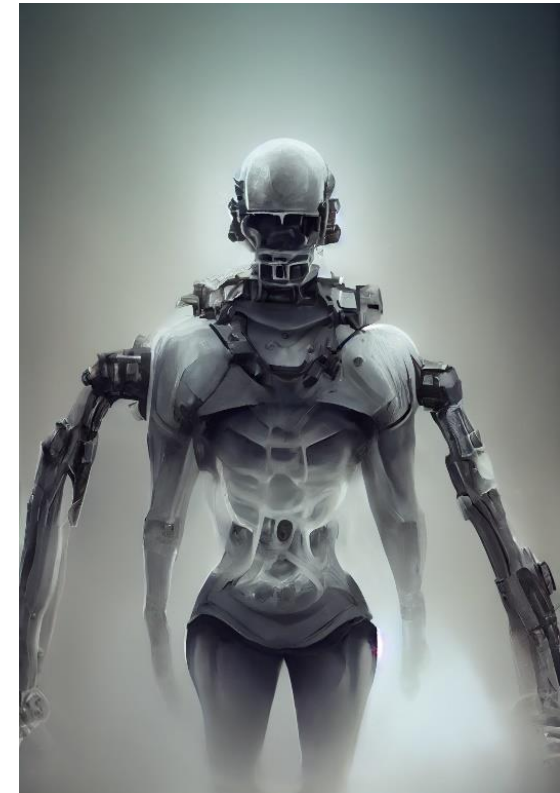
1. ¿Qué es la inteligencia artificial?
2. ¿Cómo aprenden las inteligencias artificiales?
3. ¿Qué son las inteligencias artificiales generativas?
4. Modelos del lenguaje, ¿cómo funcionan?
5. IA y medicina: guía para la asistencia sanitaria, la investigación y la docencia
6. Conclusiones

1. ¿Qué es la inteligencia artificial?
2. ¿Cómo aprenden las inteligencias artificiales?
3. ¿Qué son las inteligencias artificiales generativas?
4. Modelos del lenguaje, ¿cómo funcionan?
5. IA y medicina: guía para la asistencia sanitaria, la investigación y la docencia
6. Conclusiones

¿Qué es la inteligencia artificial?

Inteligencia Artificial (IA)

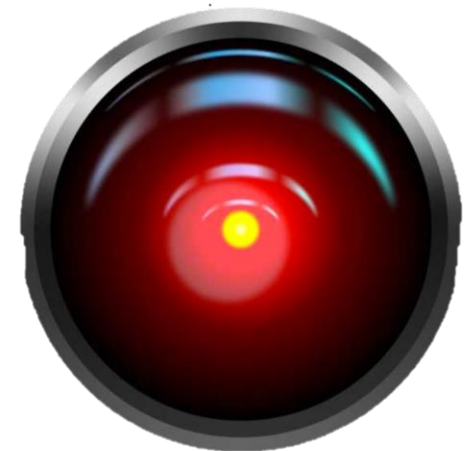
- *Es el área de informática que estudia cómo crear máquinas inteligentes*
 - Una máquina será inteligente si es capaz de realizar tareas comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico



¿Qué es la inteligencia artificial?

Diferenciamos dos tipos de IA

- **Fuerte (general):** La máquina es capaz de realizar cualquier tarea intelectual que pueda realizar el ser humano
 - Es capaz de afrontar múltiples tareas. Puede tener consciencia. Puede hacer juicios, pensar racionalmente, aprender, comunicarse en el lenguaje natural,...
 - Las IAs que aparecen en las películas
 - T-800, Wall-E, J.A.R.V.I.S., HAL



¿Qué es la inteligencia artificial?

Diferenciamos dos tipos de Inteligencia Artificial

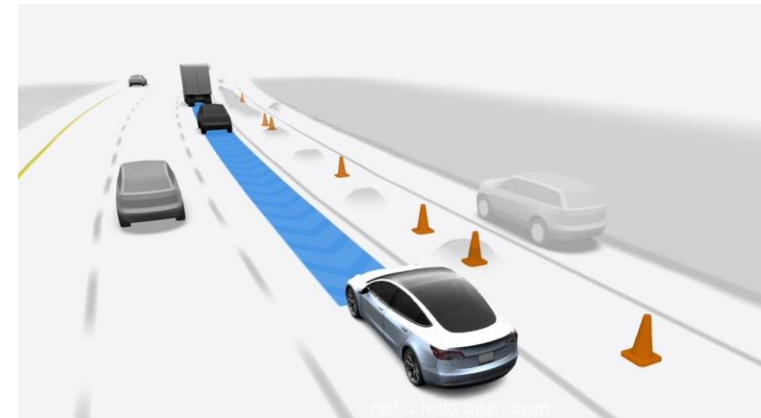
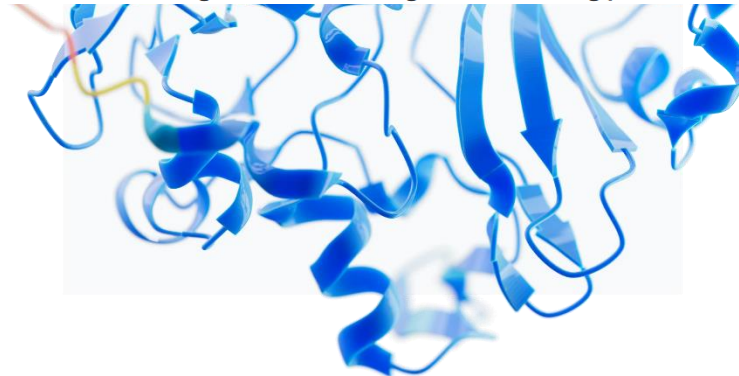
- **Débil (específica):** no tienen una capacidad cognitiva general
 - Funcionalidad limitada
 - Resuelve un conjunto concreto de problemas
 - Las IAs que conocemos
 - AlphaGo, AlphaFold, FSD Tesla, ... ¿ChatGPT, Midjourney?



AlphaGo

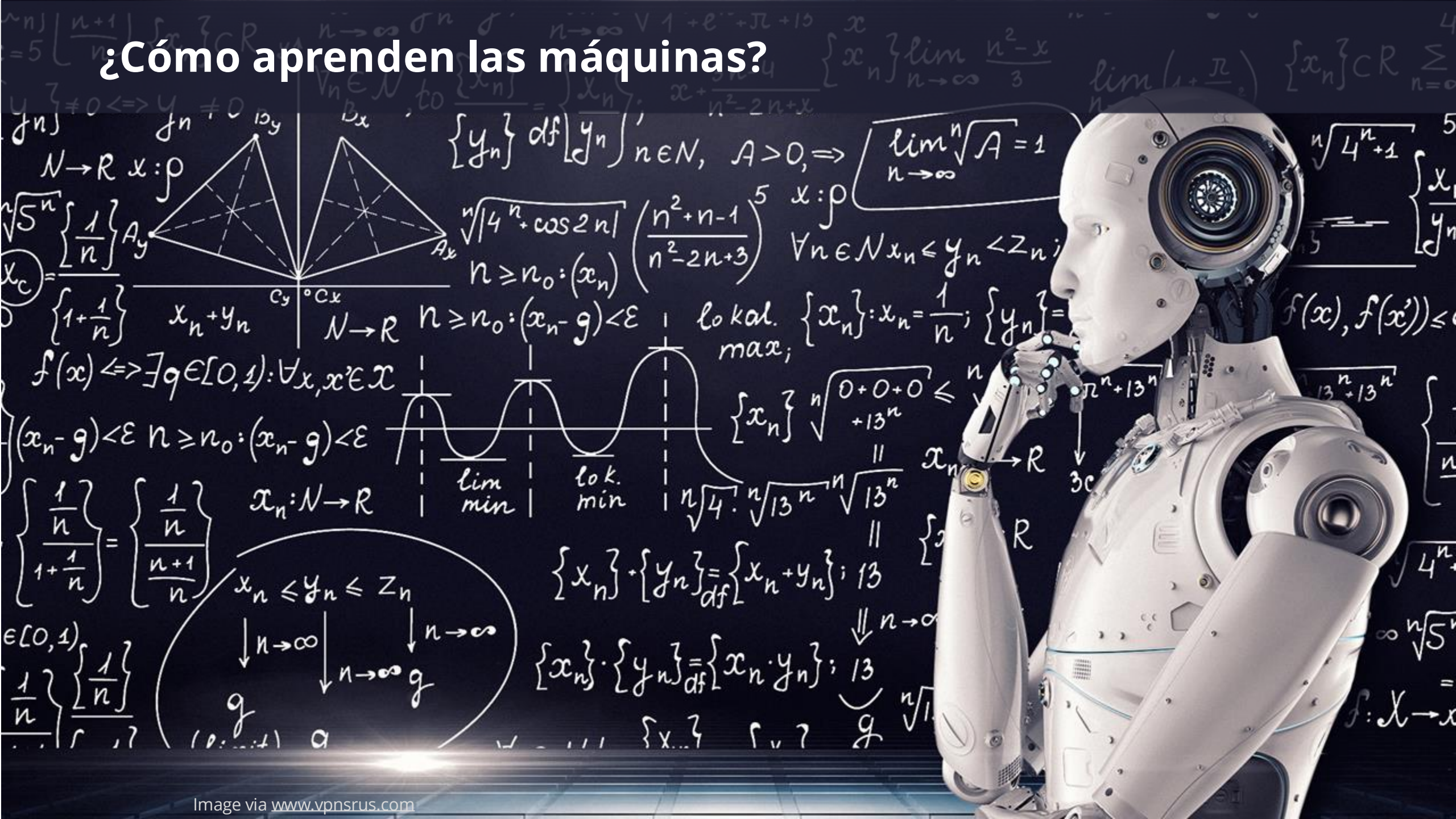
AlphaFold

Accelerating breakthroughs in biology with AI



1. ¿Qué es la inteligencia artificial?
2. ¿Cómo aprenden las inteligencias artificiales?
3. ¿Qué son las inteligencias artificiales generativas?
4. Modelos del lenguaje, ¿cómo funcionan?
5. IA y medicina: guía para la asistencia sanitaria, la investigación y la docencia
6. Conclusiones

¿Cómo aprenden las máquinas?



Aprendizaje Automático (*Machine Learning*)

- Su **objetivo** es desarrollar máquinas **capaces de aprender**
 - ¿A partir de qué?
 - De la experiencia – **¡De ejemplos!**

¿Cómo aprendemos los humanos?

Aprendizaje Automático (*Machine Learning*)

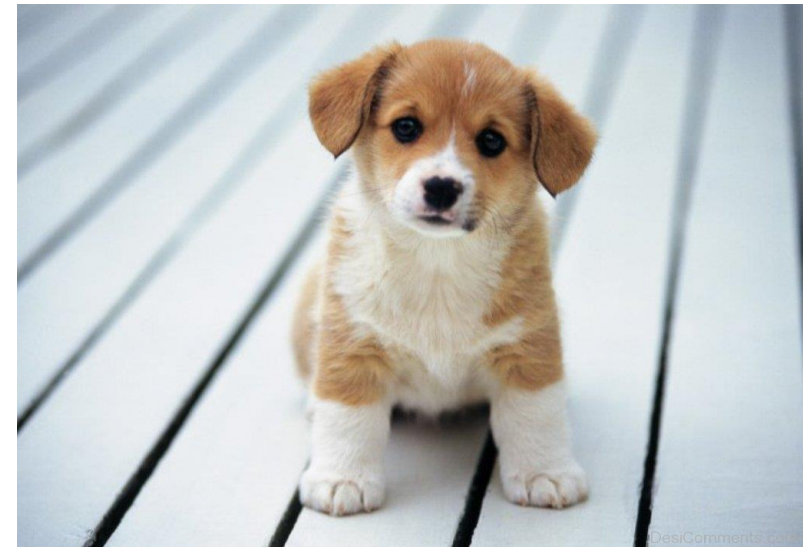
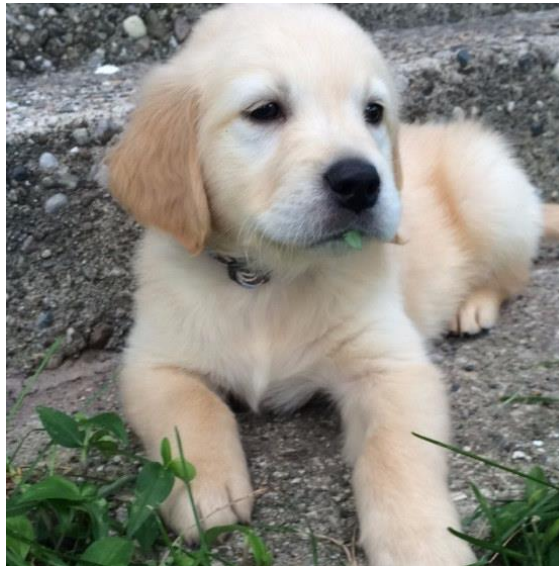
¿Cómo aprendemos los humanos?



¿Cómo aprendemos los humanos?

Aprendizaje Automático (*Machine Learning*)

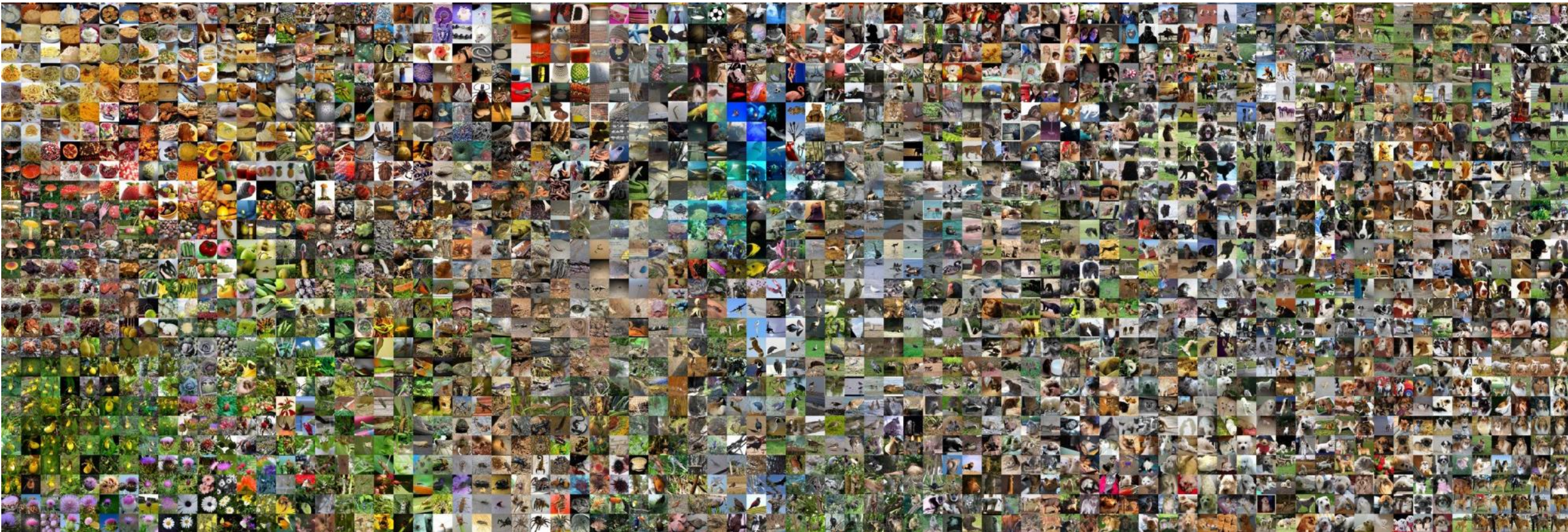
¿Cómo aprendemos los humanos?



¿Cómo aprendemos los humanos?

Aprendizaje Automático (*Machine Learning*)

¿Cómo aprenden las máquinas?



Aprendizaje Automático (*Machine Learning*)

¿Por qué surge?

- Problemas que **no pueden afrontarse con programas tradicionales**
- Aplicaciones que **no pueden ser programadas a mano**
 - Procesamiento del lenguaje natural
 - Visión por computador
 - ...

Muchas definiciones...

- **Arthur Samuel** (1959)

El área de estudio que da a los ordenadores la capacidad de aprender sin ser programados explícitamente.



<https://www.xatakaciencia.com/computacion/la-emergencia-del-buen-juego-en-un-tablero-de-damas-de-1950>

1. ¿Qué es la inteligencia artificial?
2. ¿Cómo aprenden las inteligencias artificiales?
3. ¿Qué son las inteligencias artificiales generativas?
4. Modelos del lenguaje, ¿cómo funcionan?
5. IA y medicina: guía para la asistencia sanitaria, la investigación y la docencia
6. Conclusiones

Inteligencia Artificial Generativa

- Se centra en la **generación de contenido** (texto, imágenes u otras modalidades) a partir de **datos existentes**
- Trata de modelar la **distribución de datos original**
- Puede **condicionarse**

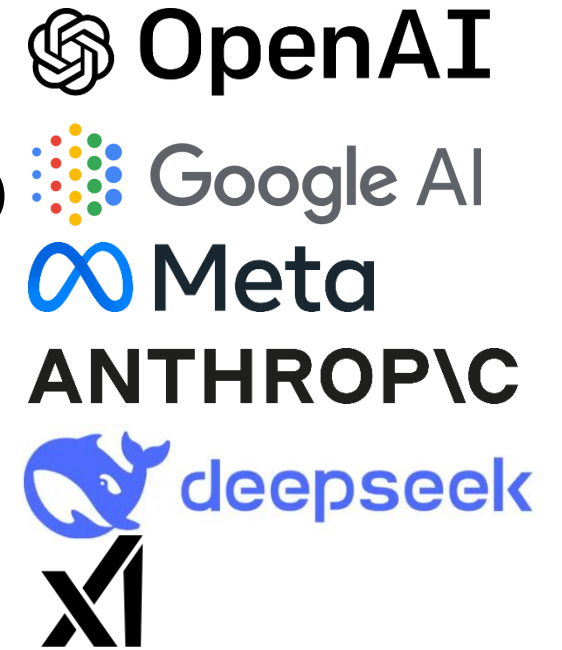
Modalidades

- **Texto**
 - Código
- **Imagen**
- Audio/Música
- Vídeo
- Moléculas
- ¿Multimodalidad?



Generadores de texto automáticos

- **OpenAI**: GPT-3 (2020), ChatGPT (2022) y GPT-4 (2023)
- **Google**: LaMDA (2021), PaLM (2021), PaLM2 (2023), Gemini (2023)
- **Meta**: LLaMA (2023), LLaMA 2 (2023), LLaMA 3 (2024)
- **Anthropic**: Claude 1 y 2 (2023), Claude 3 y 3.5 (2024)
- **DeepSeek**: V3 (2024), R1 (2025)
- **xAI**: Grok-1 (2023), Grok-2 (2024), Grok-3 (2025)



- Estos sistemas **pueden generar texto** en base a la entrada que le damos
- Pueden mantener una conversación
- Cualquiera los puede usar: [ChatGPT](#), [Gemini](#), [Llama](#)...

Generadores de texto automáticos (GPT-4)

Razonamiento avanzado

Input

Andrew is free from 11 am to 3 pm, Joanne is free from noon to 2 pm and then 3:30 pm to 5 pm. Hannah is available at noon for half an hour, and then 4 pm to 6 pm. What are some options for start times for a 30 minute meeting for Andrew, Hannah, and Joanne?

Output

Andrew: 11 am - 3 pm

Joanne: 12 pm - 2 pm, 3:30 pm - 5 pm

Hannah: 12 pm - 12:30 pm, 4 pm - 6 pm

Common availability for a 30-minute meeting: 12 pm - 12:30 pm

Creatividad

Input

Explain the plot of Cinderella in a sentence where each word has to begin with the next letter in the alphabet from A to Z, without repeating any letters.

Output

A beautiful Cinderella, dwelling eagerly, finally gains happiness; inspiring jealous kin, love magically nurtures opulent prince; quietly rescues, slipper triumphs, uniting very wondrously, xenial youth zealously.

Generadores de texto automáticos (GPT-4)

Mayor contexto

Input

[Text from: <https://en.wikipedia.org/wiki/Rihanna>]

Describe her Super Bowl performance and what made it special.

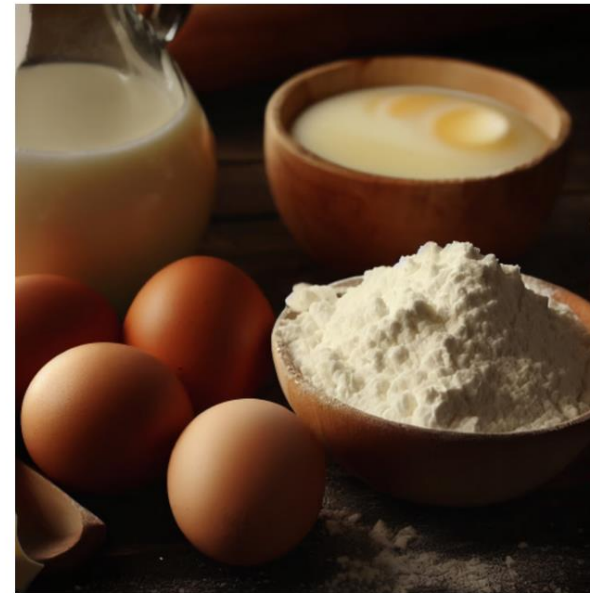
Output

Rihanna's Super Bowl LVII halftime show performance was special for a number of reasons. For one, it marked her first live performance in over five years, and it was her first time headlining the Super Bowl halftime show. Additionally, she made history by becoming the first person to perform at the Super Bowl while pregnant. The show itself was critically acclaimed, with many praising her high-energy performance and impressive vocals. It also garnered a massive viewership, becoming the second most-watched halftime show in history.

Visual input

Input

What can I make with these ingredients?



Output

There are many options for what you can make with these ingredients!
Some possibilities include:

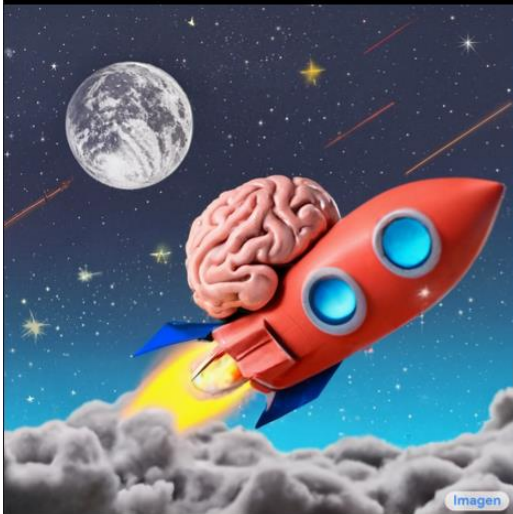
- Pancakes or waffles
- Crepes
- French toast
- Omelette or frittata
- Quiche
- Custard or pudding
- Cake or cupcakes
- Muffins or bread- Cookies or biscuits

These are just a few examples, but the possibilities are endless!



Generadores de imagen automáticos

- **OpenAI**: Dalle, Dalle-2, Dalle-3 (2021, 2022, 2023)  **OpenAI**
 - **Google**: Imagen, Imagen 2, Imagen 3 (2022, 2023, 2024)  Google AI
 - **Stability.AI**: Stable Diffusion (2022 –) – **OPEN SOURCE!**  stability.ai
 - **Midjourney** (2022 –) 
-
- Estos sistemas pueden **generar imágenes** en base nuestras órdenes
 - También pueden modificar imágenes o usarlas como punto de partida
 - Cualquiera puede utilizarlos: [Dalle-3](#), [Imagen 3](#), [Midjourney](#) o [Stable Diffusion](#)



A brain riding a rocketship heading towards the moon.



A dragon fruit wearing karate belt in the snow.



A small cactus wearing a straw hat and neon sunglasses in the Sahara desert.

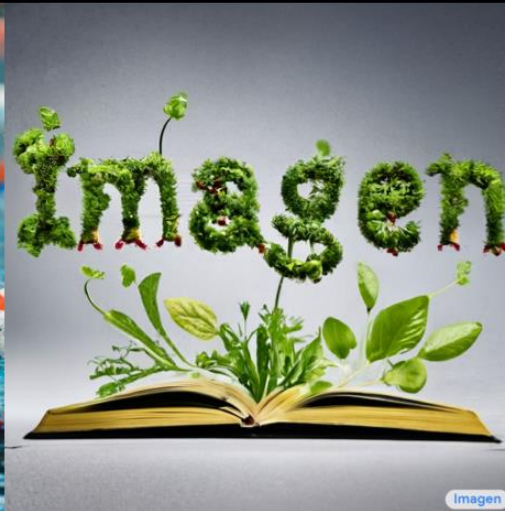


A photo of a Corgi dog riding a bike in Times Square. It is wearing sunglasses and a beach hat.

<https://imagen.research.google/>



Teddy bears swimming at the Olympics 400m Butterfly event.



Sprouts in the shape of text 'Imagen' coming out of a fairytale book.



A transparent sculpture of a duck made out of glass. The sculpture is in front of a painting of a landscape.



A single beam of light enter the room from the ceiling. The beam of light is illuminating an easel. On the easel there is a Rembrandt painting of a raccoon.

Generadores de video automáticos

- **OpenAI**: Sora (2024)  **OpenAI**
- **Google**: Veo 2, Veo 3, Flow, VideoPoet (2024, 2025)  Google AI
- **Runway**: Gen-1, Gen2, Gen 3 Alpha, Gen-4 (2023, 2024, 2025)  **runway**
- **Meta**: Movie Gen (2024)  **Meta**
- **Pika Labs** (2023, 2025, 2025) 

- Estos sistemas pueden **generar videos** en base nuestras órdenes
- También pueden modificar videos o usarlos como punto de partida
- Cualquiera puede utilizarlos: [SORA](#), [Veo](#), [Runway](#), [Pika Labs](#)



Text input summary: *A girl is running across a beach and holding a kite. She's wearing jean shorts and a yellow t-shirt. The sun is shining down.*

[Movie Gen \(Meta\)](#)



Prompt theory (Vevo 3): <https://x.com/HashemGhaili/status/1925616536791760987>

|

1. ¿Qué es la inteligencia artificial?
2. ¿Cómo aprenden las inteligencias artificiales?
3. ¿Qué son las inteligencias artificiales generativas?
4. **Modelos del lenguaje, ¿cómo funcionan?**
5. IA y medicina: guía para la asistencia sanitaria, la investigación y la docencia
6. Conclusiones

Large Language Models (LLMs)

- Modelos fundacionales: base para otras muchas aplicaciones
- Modelos muy grandes: billones de parámetros
- Entrenamiento con
 - Grandes cantidades de datos extraídos de internet – ¡Sesgos!
 - Muchas GPUs
 - Al alcance de unos pocos (Big Tech)

Large Language Models (LLMs)

- ¿Entrenan con datos etiquetados?
 - ¡No! Con **datos autoetiquetados**
 - Entrenamiento autosupervisado – **self-supervised learning**
- Ajuste por **aprendizaje por refuerzo a partir de feedback humano**
 - **Midjourney** vs. ChatGPT
- **Conocimiento limitado** a la fecha de entrenamiento
 - Conexión a internet $\neq$ conocimiento previo

Características

- Tienen un **conocimiento base muy grande**
- Tienen una **buena memoria** de trabajo (contexto actual)
- **Pero no**
 - Tienen un **monólogo interior**
 - No realizan **comprobaciones**
- Podemos verlos como un “**humano medio**” (con mucho conocimiento)
 - Interacción con internet: referencias genéricas, no libros, artículos...
 - Aparecerán **modelos específicos**: legales, administrativos, científicos

Limitaciones

- **Alucinaciones**: todas sus respuestas son asertivas a pesar de poder ser completamente falsas
- **Costes** de entrenamiento y uso
- **Dependencia** en el largo plazo
- **Sesgos**
- **Regulación**: uso, datos de entrenamiento, ¿quién será el culpable de los errores?
 - **Europa**: buena regulación, pero ¿hay riesgo de quedarnos atrás?
- **Privacidad**: ¿Qué ocurre con los datos enviados a OpenAI?

¿Cómo funcionan?

- Tokenización
 - **Token**: subconjunto de caracteres con significado
 - El texto se tokeniza
 - Se divide en tokens
 - El número de tokens **limita el contexto del LLM**
- ChatGPT: 4k tokens
- GPT-4: 8k tokens (hasta 32k tokens)
- El límite incluye la respuesta a generar

GPT-3 Codex

Vamos a hacer una prueba para ver cómo tokeniza las palabras ChatGPT

Clear Show example

Tokens	Characters
25	68

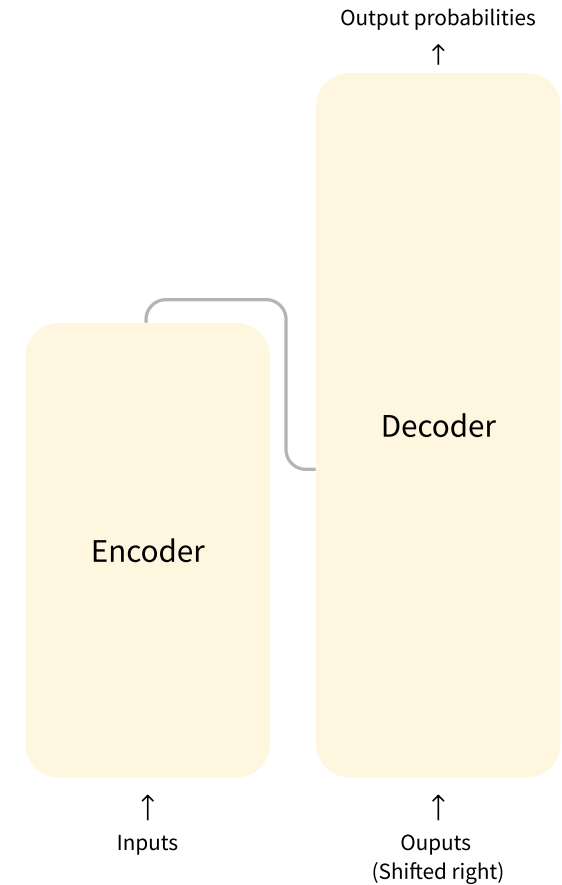
Vamos a hacer una prueba para ver cómo tokeniza las palabras ChatGPT

TEXT TOKEN IDS

A helpful rule of thumb is that one token generally corresponds to ~4 characters of text for common English text. This translates to roughly $\frac{3}{4}$ of a word (so 100 tokens $\approx$ 75 words).

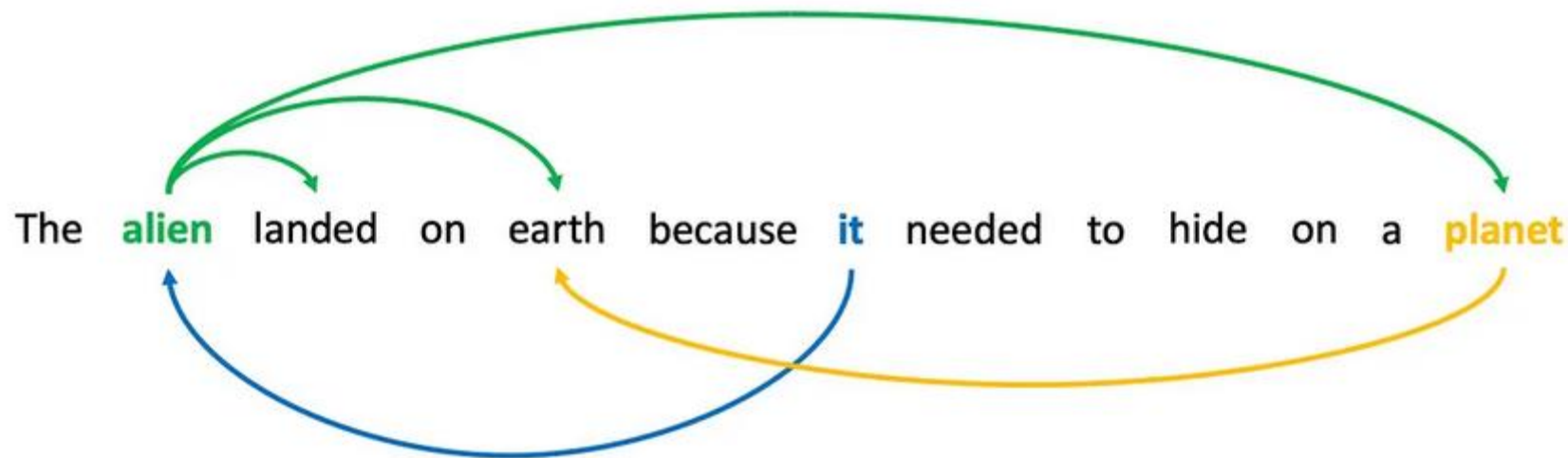
¿Cómo funcionan?

- Arquitectura: **Transformers**
- Entrenamiento
 - **Causal Language Model** vs. Masked Language Model
 - Predicción del siguiente token
 - Rellenado de huecos en una frase
- Predicción condicionada mediante *prompts*



¿Cómo funcionan?

- Mecanismo de atención
 - Para predecir una nueva palabra, se analiza todo el **contexto previo**
 - El **contexto es el texto ya generado** (se empieza por un token de inicio)



<https://mark-riedl.medium.com/a-very-gentle-introduction-to-large-language-models-without-the-hype-5f67941fa59e>

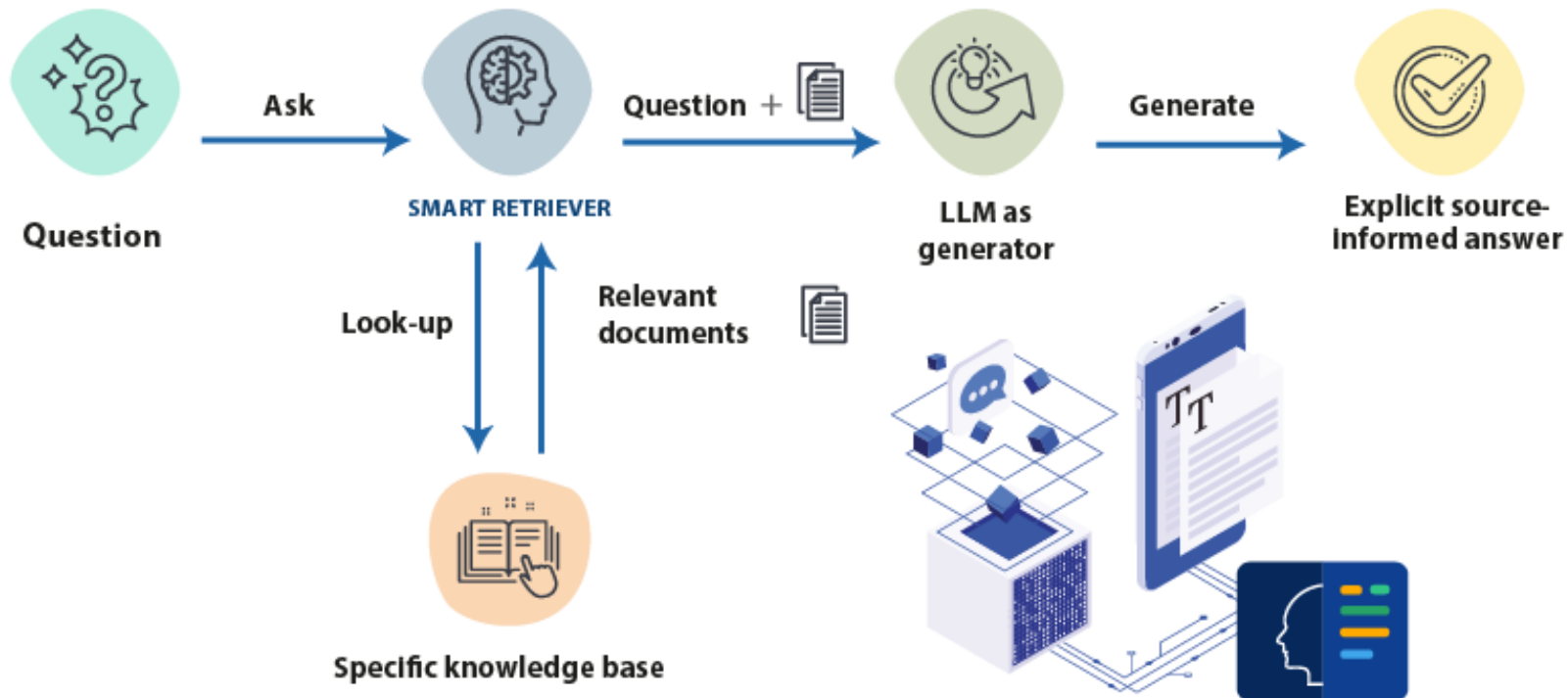
RAG: Retrieval Augmented Generation

LLMs para consultar documentos

There are 3 phases:

1 RETRIEVAL 2 AUGMENTED 3 GENERATION

The process is as follows:



Aplicaciones – nuestro “Copilot”

1. Corrección de textos
2. Traducción de textos
3. Resumen de textos
4. Correos y copywriting
5. Ayuda preguntas informáticas
6. Programación: escribir un programa
7. Turismo: guía turística
8. Psicología: asesorar sobre algo
9. Escritura: escribir un microrrelato, una propuesta...
10. Entrenador personal: generar un plan de trabajo
11. Educación: generar exámenes y soluciones, tutor personal
12. Generador de imágenes y vídeos

Modelos de razonamiento – ya son una realidad

Modelos que piensan (generan tokens de pensamiento)

Deep research (investigación profunda) – emergiendo (desde hacer 6-8 meses)

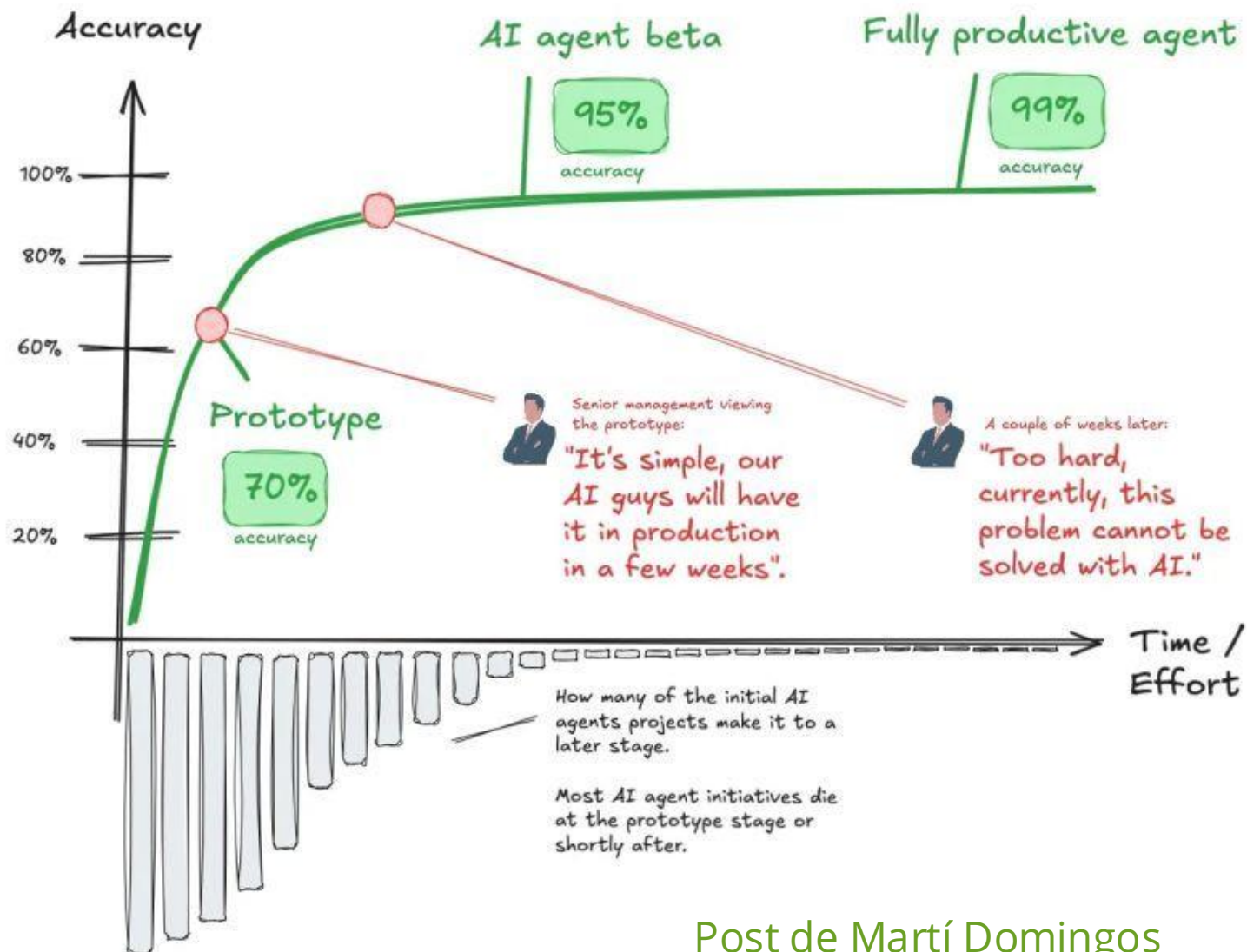
Búsquedas profundas con informes completos

Agentes de IA – emergiendo (desde hacer 10-16 meses)

Cuando damos a los LLMs la capacidad de ser autónomos

AGI (Artificial General Intelligence) – nos quedan entre 6 meses y 100 años

The AI Agent Paradox



1. ¿Qué es la inteligencia artificial?
2. ¿Cómo aprenden las inteligencias artificiales?
3. ¿Qué son las inteligencias artificiales generativas?
4. Modelos del lenguaje, ¿cómo funcionan?
5. IA y medicina: guía para la asistencia sanitaria, la investigación y la docencia
6. Conclusiones

La IA ya está aquí

- La IA ha pasado de la promesa a la práctica:
 - Diagnósticos por imagen, notas clínicas, predicción de ocupación hospitalaria
- También acelera la investigación: revisión bibliográfica, redacción, análisis de datos
- Pero su uso sin reflexión puede amplificar sesgos, causar errores diagnósticos o vulnerar derechos

Los datos que alimentan la IA

- Ninguna IA es mejor que los datos que la alimentan
- Importa la representatividad (¿entrena con población similar a la tuya?)
- **Riesgo**: perpetuar inequidades si el modelo no refleja diversidad demográfica o técnica (ej. retinopatía en poblaciones no representadas)
- **Pilares esenciales**: Gobernanza, privacidad y trazabilidad de datos

Principios éticos fundamentales (UNESCO)

- Equidad y no discriminación
- Transparencia y explicabilidad
- Supervisión humana y rendición de cuentas
- Participación y confianza del paciente
 - El profesional debe ser mediador cualificado entre el algoritmo y el paciente

IA en la práctica clínica: entorno de alto riesgo

- Los sistemas de IA médica están regulados por:
 - MDR (productos sanitarios con IA, marcado CE)
 - RIA (Reglamento Europeo de IA, 2024)

Marcado CE

- El marcado CE no basta: garantiza la validación en un contexto concreto, no la seguridad universal
- **Responsabilidad compartida**: fabricante, institución y profesional

El profesional debe poder:

- Entender capacidades y limitaciones del sistema
- Interpretar y, si es necesario, anular sus salidas
- Evitar el sesgo de automatización (aceptar sin cuestionar)
- Formación → obligación legal y ética (RIA, art. 14 y 26)

Artículo 14 del RIA describe cómo debe realizarse dicha **supervisión humana**:

El usuario debería tener la capacidad de: (a) "comprender adecuadamente las capacidades y limitaciones pertinentes del sistema de IA de alto riesgo", (b) "ser conscientes de la posible tendencia a confiar automáticamente o en exceso en los resultados", (c) "interpretar correctamente los resultados", (d) "hacer caso omiso, anular o invertir el resultado", así como (e) "intervenir en el funcionamiento del sistema (...) o interrumpirlo mediante un botón de parada".

Riesgos en el uso de LLMs clínicos

- Omisión crítica: dejar fuera información vital
- Alucinación plausible: generar contenido falso pero convincente
- Pérdida de contexto: interpretar mal datos aislados
- Riesgo de confidencialidad: datos clínicos en plataformas públicas

LLMs en investigación: potencial y precaución

- **Útiles para:** Revisiones bibliográficas, generación de código, redacción científica, revisión virtual por pares
- **Buenas prácticas:**
 1. No introducir datos personales
 2. Tú eres el autor
 3. Verifica todo
 4. Detecta sesgos
 5. Documenta el uso
 6. Domina el prompting
 7. Itera y contrasta
 8. Protege la propiedad intelectual
 9. Conoce sus límites temporales
 10. Potencia, no sustituyas, tu juicio experto

Propuesta curricular en IA sanitaria

- Grado → Fundamentos, ética, regulación
- Residencia → Aplicación práctica con IA y LLMs
- Formación continuada → Actualización normativa y tecnológica

1. ¿Qué es la inteligencia artificial?
2. ¿Cómo aprenden las inteligencias artificiales?
3. ¿Qué son las inteligencias artificiales generativas?
4. Modelos del lenguaje, ¿cómo funcionan?
5. IA y medicina: guía para la asistencia sanitaria, la investigación y la docencia
6. Conclusiones

Principios que deben guiar la formación sanitaria en IA

1. Alfabetización en IA transversal (práctica clínica e investigación)
2. Supervisión humana competente
3. Integridad científica
4. Responsabilidad indelegable del profesional
5. Pensamiento crítico

“El valor de la IA no está en la tecnología, sino en el profesional que la usa con juicio, ética y humanidad.”

Hay esperanza: Permitirá más avances científicos y poder centrar la atención donde aportamos valor. **Aprovechemos su potencial y formémonos**



Fundamentos de la Inteligencia Artificial Generativa

MIKEL GALAR IDOATE, IRIS DOMÍNGUEZ CATENA

mikel.galar@unavarra.es , iris.dominguez@unavarra.es

upna arin

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Artificial Intelligence &
Machine Learning Research

Departamento de Estadística, Informática y Matemáticas

Institute of Smart Cities (ISC)

ARIN. Artificial Intelligence and Machine Learning Research