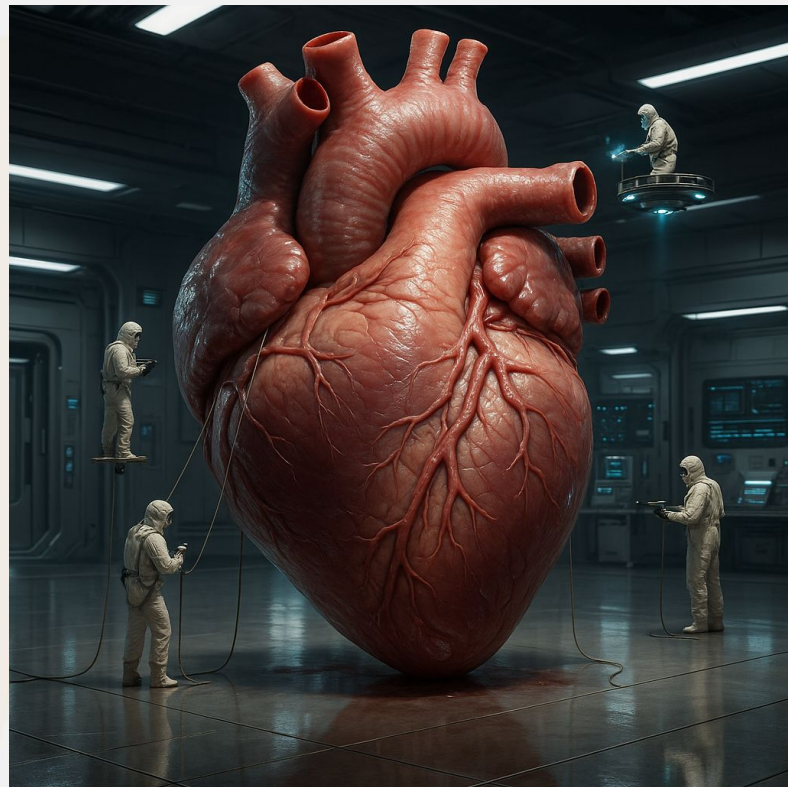


IA Generativa para la Innovación

Semana 1

Ing. Leopoldo López Gómez MPA MSc
Harvard University

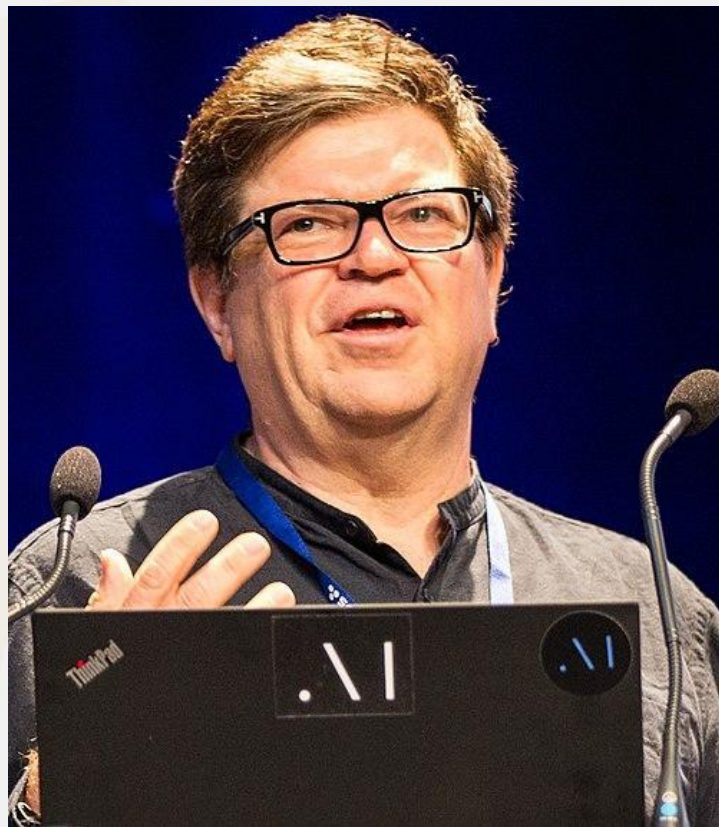


Generado con Sora, Jun 25

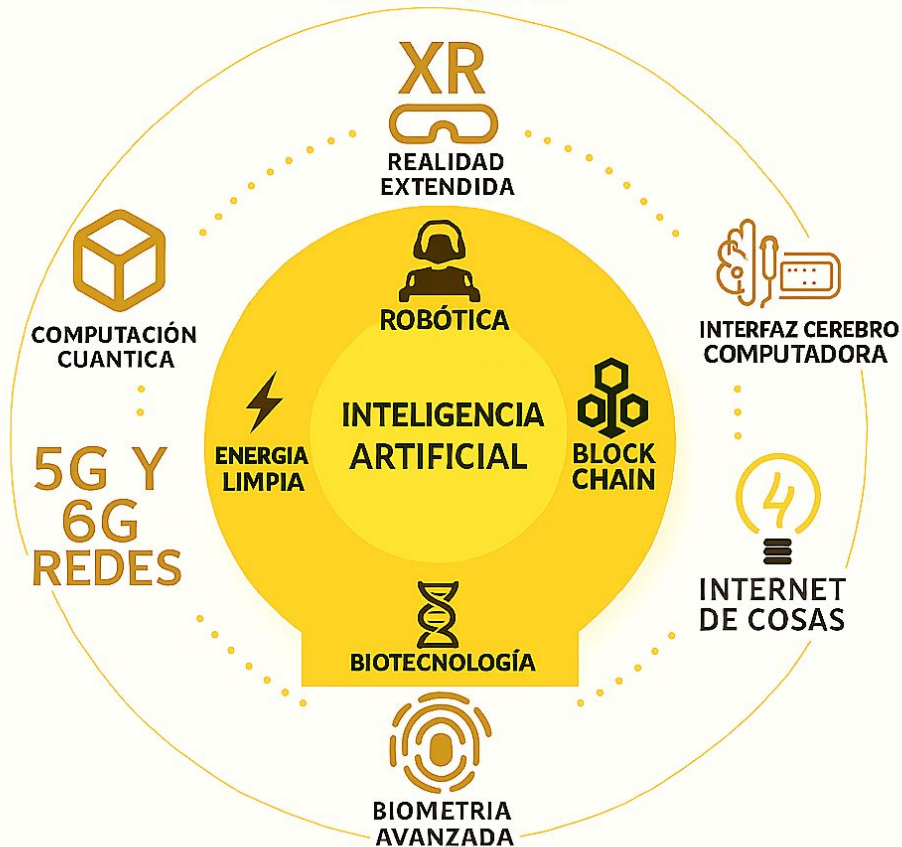
Quote of the day

“Large language models don’t understand the world – they understand the statistics of language.”

— **Yann LeCun**, Chief AI Scientist, Meta; premio Turing



LA CONVERGENCIA TECNOLÓGICA DEL FUTURO



“ La IA permite generar cientos de hipótesis analizando miles de estudios o artículos médicos, acelerando la velocidad del descubrimiento científico”

CÓMO ESTÁ IMPACTANDO LA IA EN DIFERENTES INDUSTRIAS

PROGRAMACIÓN DE SOFTWARE

Porcentaje de código generado con IA



INVESTIGACIÓN MÉDICA

Número de hipótesis evaluadas al mes



¿Reemplazará la IA a los humanos?

“ Los trabajadores que sepan usar la IA reemplazarán a los que no lo hagan”



SEMANA 1

- 1 INTRODUCCIÓN
- 2 LA IA Y EL FUTURO DEL TRABAJO
- 3 ¿QUÉ ES LA INTELIGENCIA ARTIFICIAL?



TEMA 1 INTRODUCCIÓN



¿Reemplazará la inteligencia artificial a los humanos?

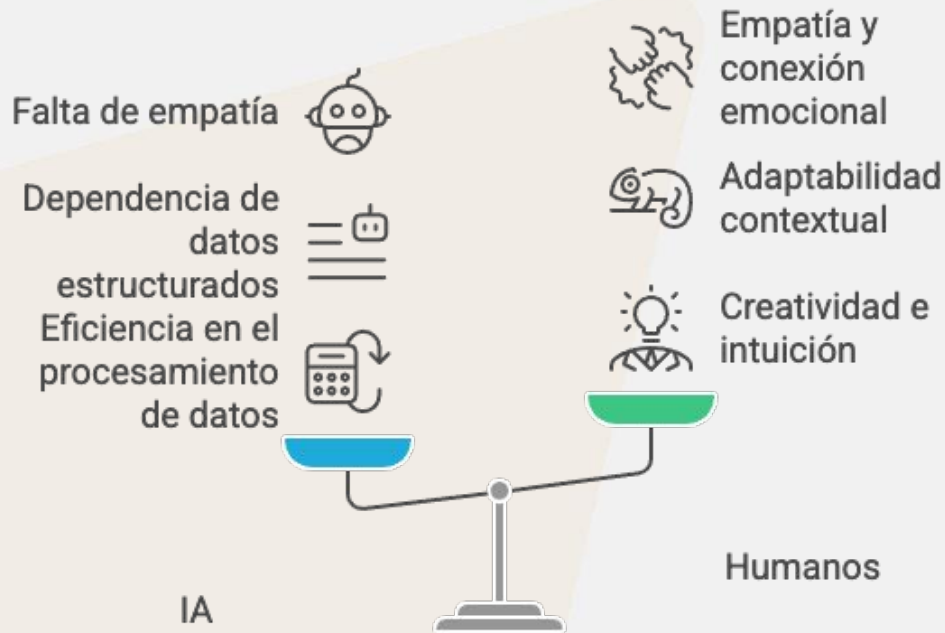
Creatividad y emociones humanas

Adaptabilidad y juicio

Empatía y conexión humana



Lo que la IA no puede reemplazar



Equilibrando las fortalezas de la IA y las cualidades humanas

La IA Generativa vs. Internet como revolución tecnológica

**Capacidad de Aprendizaje
y Adaptación**

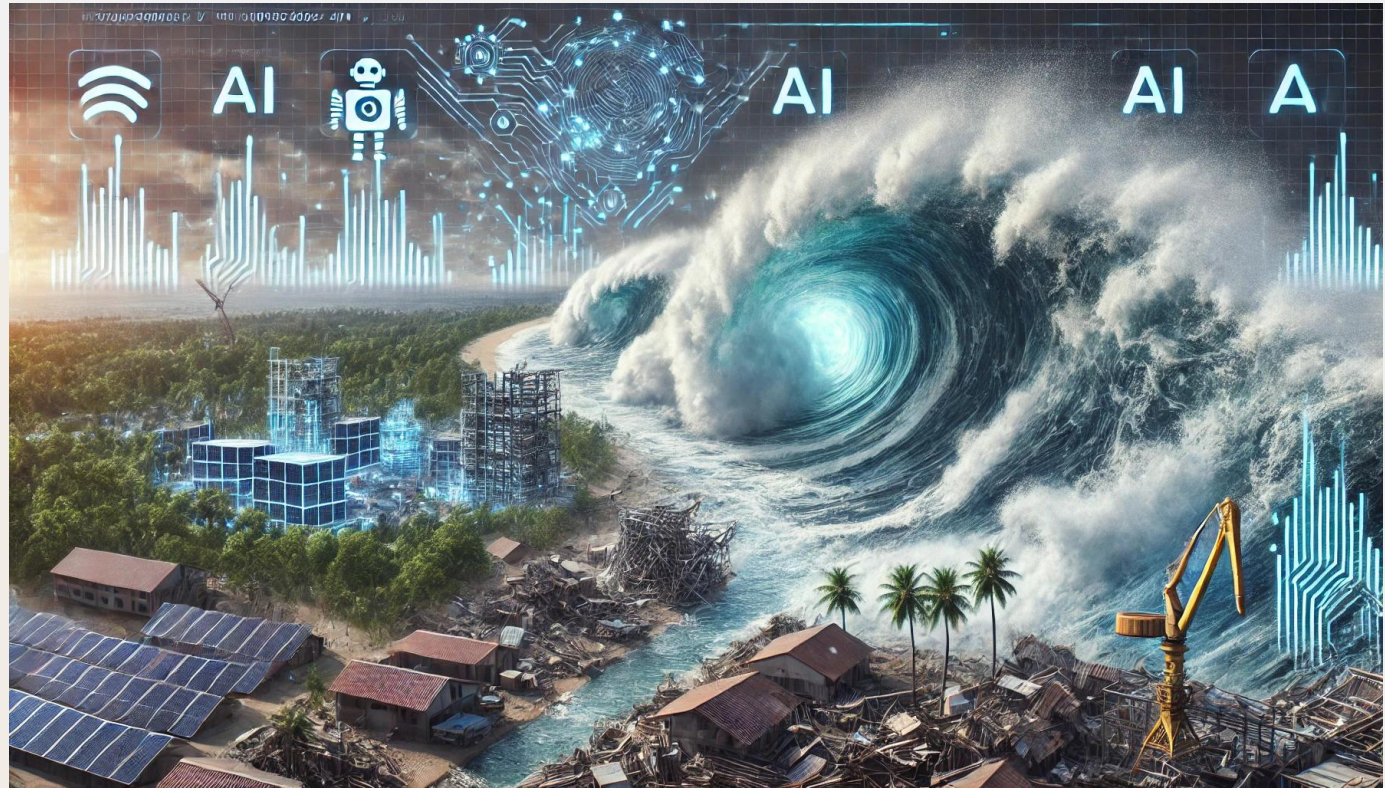
**Automatización y
Eficiencia**

**Innovación y Nuevas
Fronteras**



Generado por DALL-E, 2024

La Inteligencia Artificial como un Tsunami



Generado por DALL-E, 2024

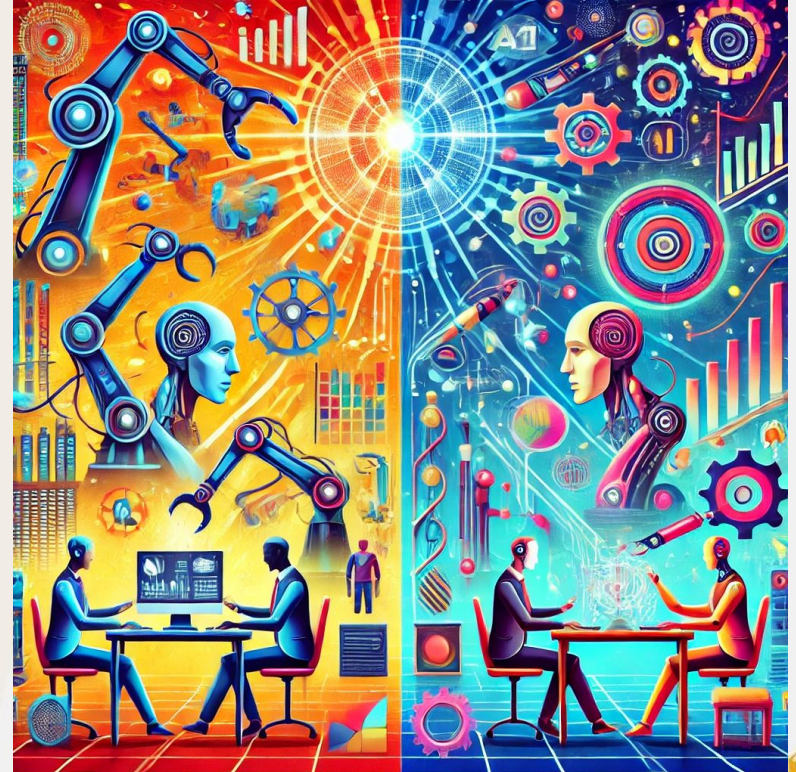
TEMA 2

LA IA y el FUTURO DEL TRABAJO

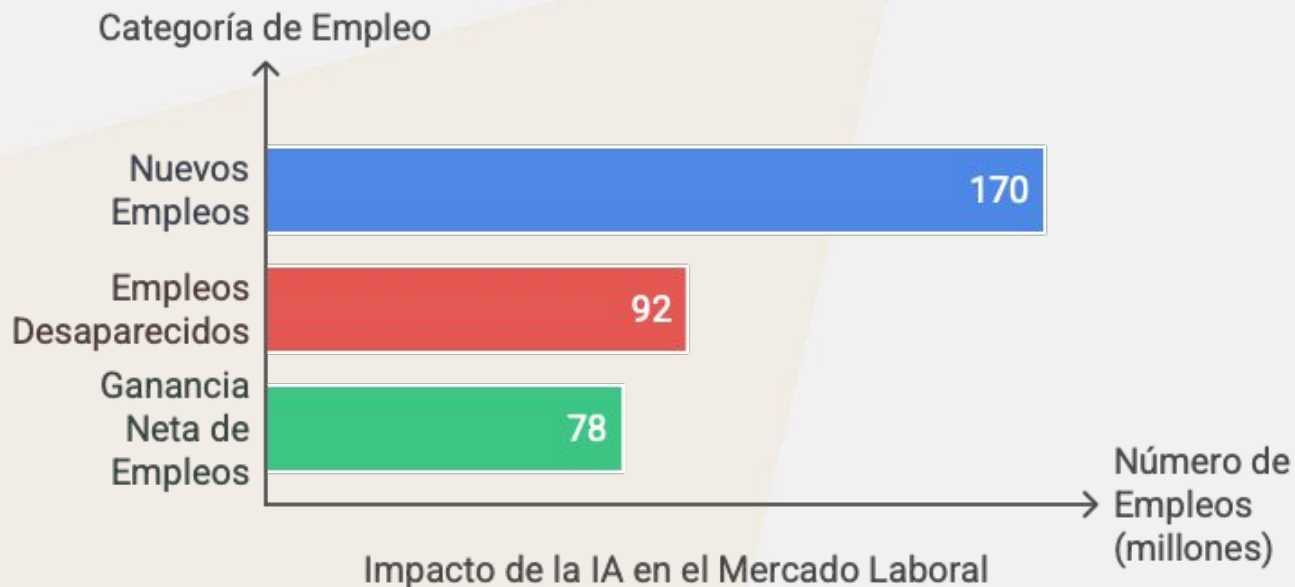


Cómo afectará la IA el mercado laboral ?

- IA la nueva revolución industrial
- Aumento de la productividad
- 40 % de los empleos expuestos
- Favorece a trabajadores de mayores ingresos y universitarios
- Desfavorece a Empleados Mayores y con menos educación
- Países menos desarrollados:
Infraestructura y Educación Digital



Cuántos nuevos empleos se van a crear?



Made with Napkin

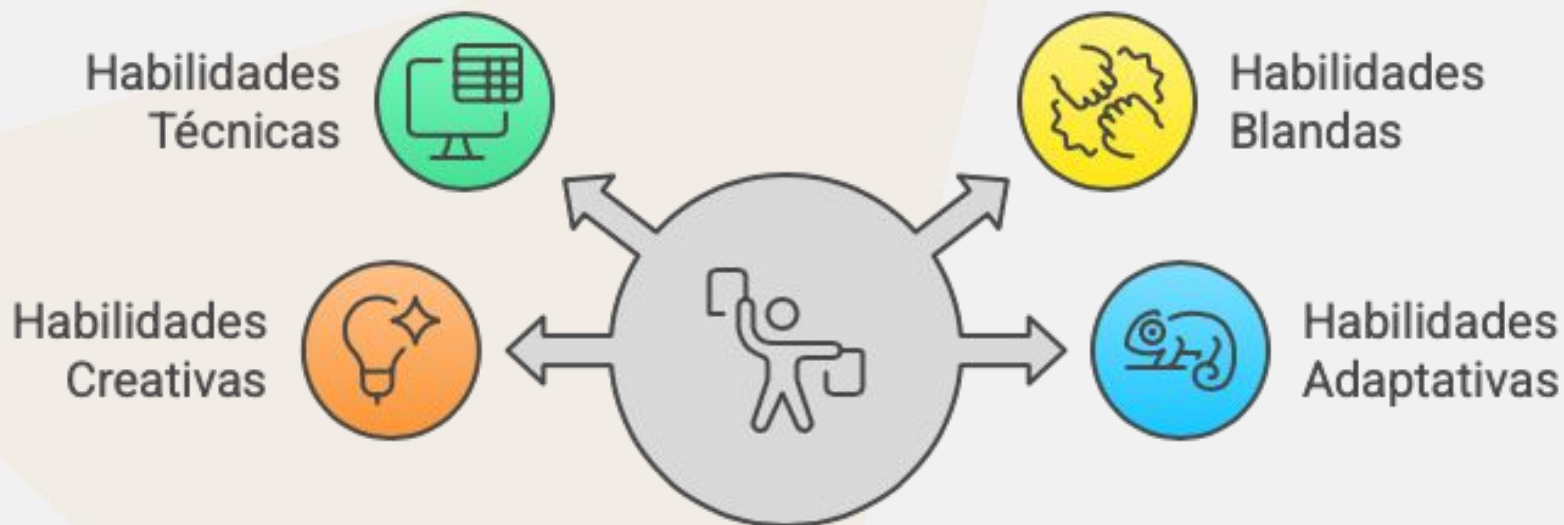
Sectores Impactados por la IA Generativa



Nuevas Profesiones en la Era de la IA

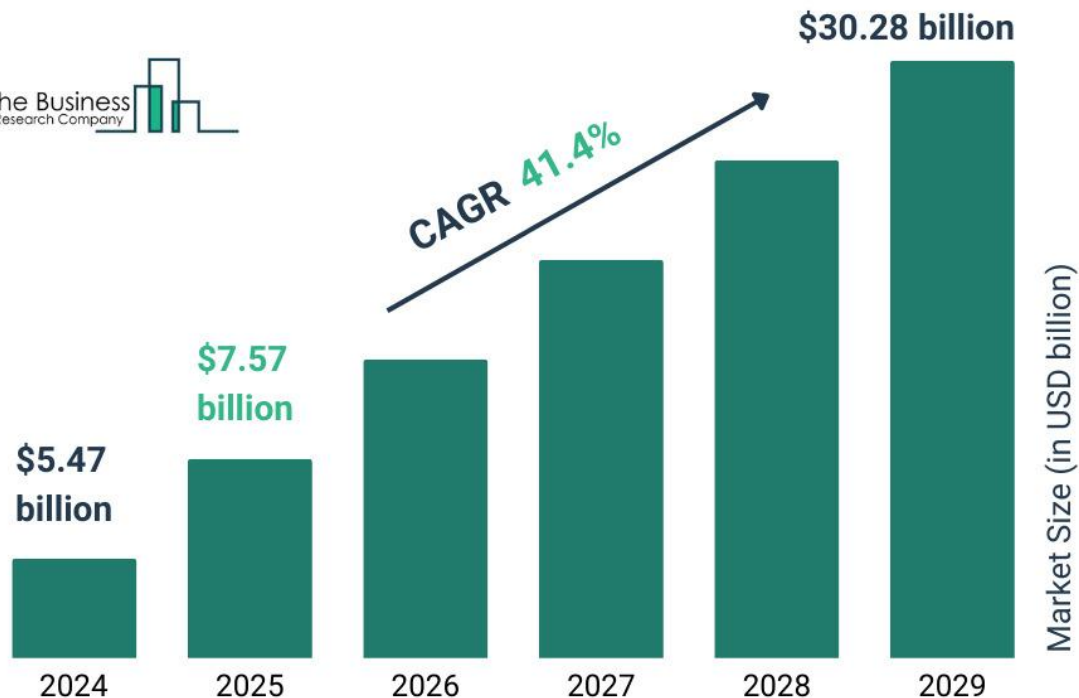


Habilidades para el Éxito en la Era de la IA



AI in Education Global Market Report 2025

The Business
Research Company



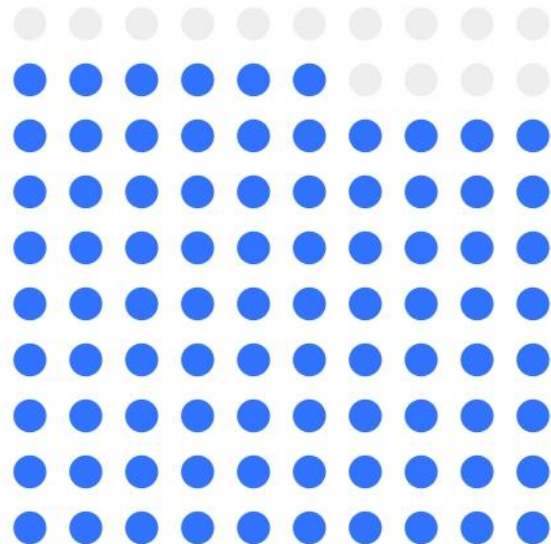
86% of students already use AI in their studies

Percentage of students using AI in their studies

Question: How often do you use AI tools?

86%

of students claim to use
AI in their studies

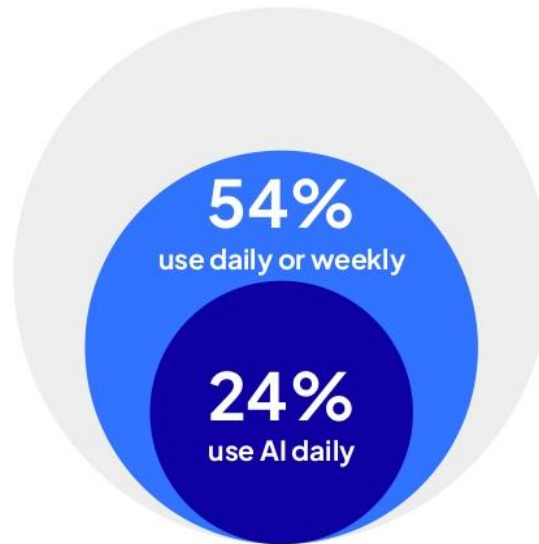


Frequency of students using AI in their studies

Question: How often do you use AI tools?

54%

of students use AI at least
on a weekly basis



Meet the 6 types of AI users

We categorize AI users into six types based on frequency of AI usage and optimism about the future of AI in education.

AI Enthusiasts

(Always/Often + Extremely Optimistic)

- Passionate early adopters who embrace AI's potential
- Believe AI is a game-changer for education



Pragmatic Users

(Always/Often + Somewhat Optimistic/Pessimistic)

- Use AI because it's useful but have mixed or cautious feelings
- Open to AI but may have reservations



Skeptical Adopters

(Always/Often + Extremely Pessimistic)

- Use AI despite deep skepticism
- May feel forced to adopt AI due to workplace demands



Hopeful Bystanders

(Sometimes/Rarely/Never + Extremely Optimistic)

- Believe AI has great potential but don't use it frequently
- Could be limited by access, training, or institutional policies



Cautious Observers

(Sometimes/Rarely/Never + Somewhat Optimistic/Pessimistic)

- Neutral or undecided about AI
- May be waiting for better tools, more research, or clearer guidelines



Traditionalists

(Sometimes/Rarely/Never + Extremely Pessimistic)

- Avoid AI and strongly distrust it
- Prefer traditional teaching methods and see AI as a threat or unnecessary



Students in UK's first 'teacherless' classroom taught by AI

By AV Magazine in Education, UK&I, VR, AR & MR September 2, 2024 0

The 20 students at a private school in the City of London will be taught via a combination of AI platforms on their computers and VR headsets.



Photo: Sky News

New Arizona charter school will use AI in place of human teachers

KJZZ | By Nate Engle

Published December 18, 2024 at 9:30 AM MST



▶ LISTEN • 1:02



Getty Images

TEMA 3

¿Qué es la inteligencia artificial?



INTELIGENCIA ARTIFICIAL

"La inteligencia artificial (IA) es una rama de la informática dedicada a crear sistemas capaces de imitar, simular o superar aspectos de la inteligencia humana, como aprender, razonar, adaptarse, planificar y resolver problemas.

Sus aplicaciones incluyen motores de búsqueda, sistemas de recomendación, drones, procesamiento de imágenes, vehículos autónomos y asistentes virtuales, entre otras."



Inteligencia Artificial

«Las inteligencias artificiales no son consciencias informáticas con voluntad y ambiciones.

Son líneas de código que realizan operaciones matemáticas para almacenar datos, encontrar tendencias y devolver respuestas ajustadas a lo que les hayamos enseñado que esperamos de ellas.»



ALGUNOS USOS DE LA IA



Robótica



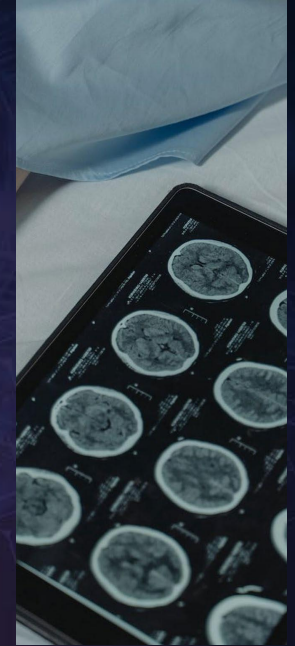
Procesamiento y
reconocimiento de
imágenes



Analítica de datos



Gestión de procesos y Aprendizaje adaptativo
toma de decisiones



Diagnóstico de
enfermedades

TIPOS DE IA



Sistemas expertos



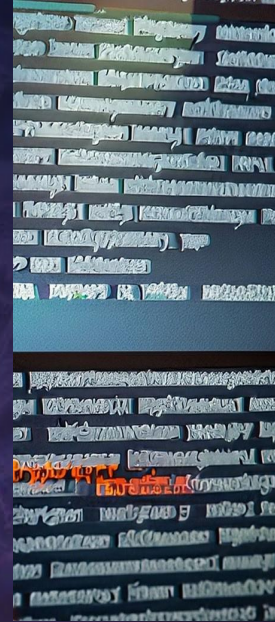
Redes neuronales
artificiales



Aprendizaje
de máquina

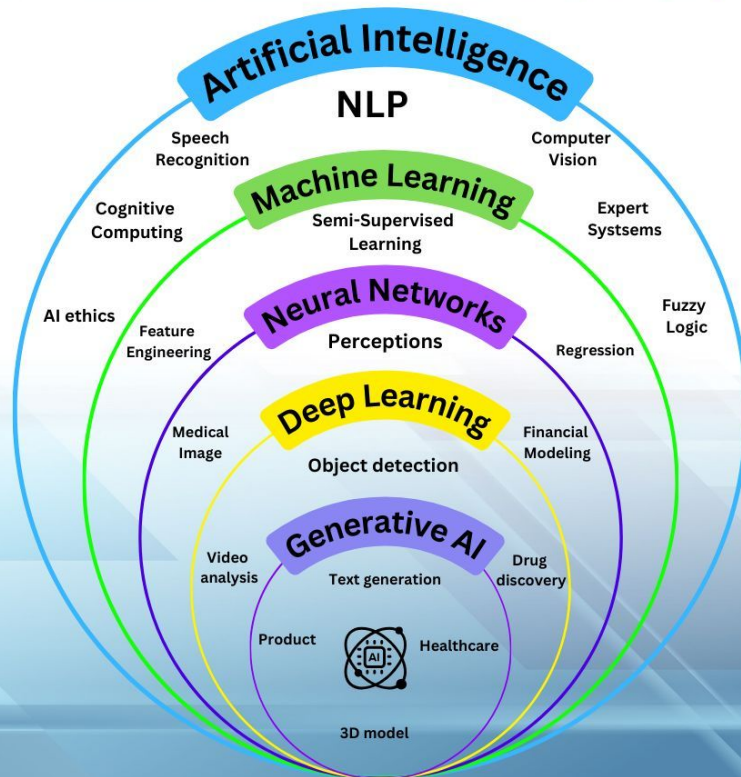


Aprendizaje
profundo



Procesamiento de
lenguaje natural

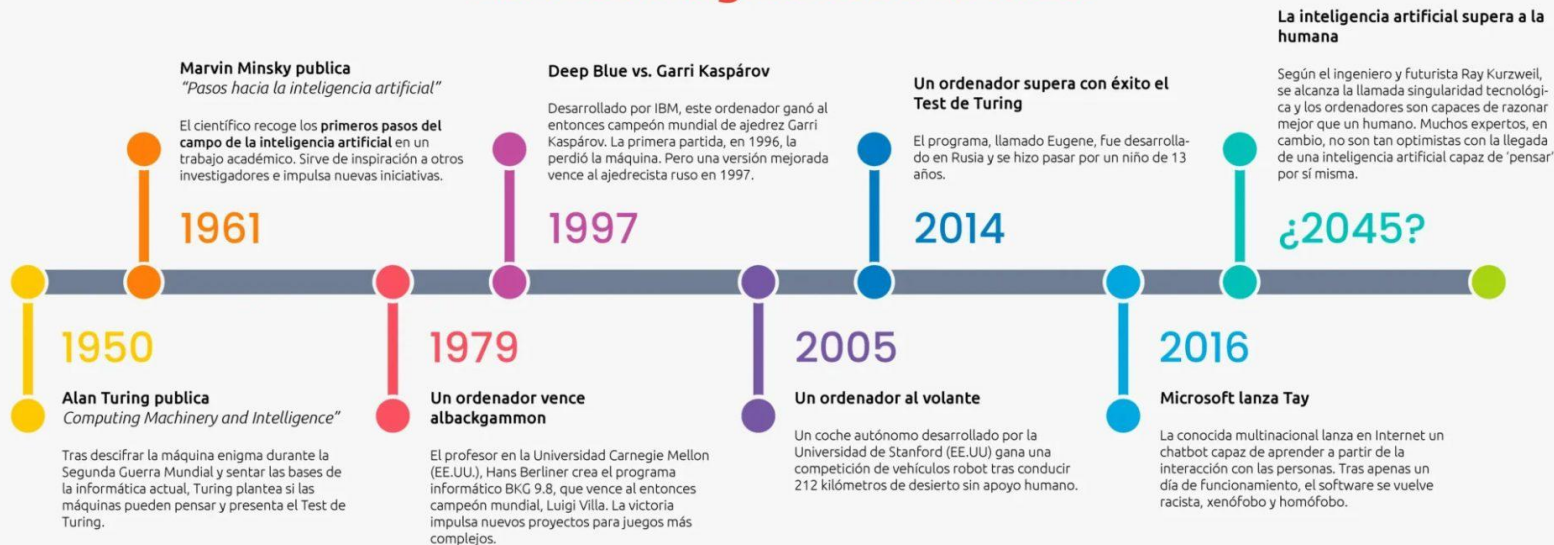
The AI Universe



The image features a human hand on the left and a white robotic hand on the right, both reaching towards the center. The background is a deep blue with faint, glowing digital patterns, including concentric circles and a line graph. The overall theme is artificial intelligence and human-machine interaction.

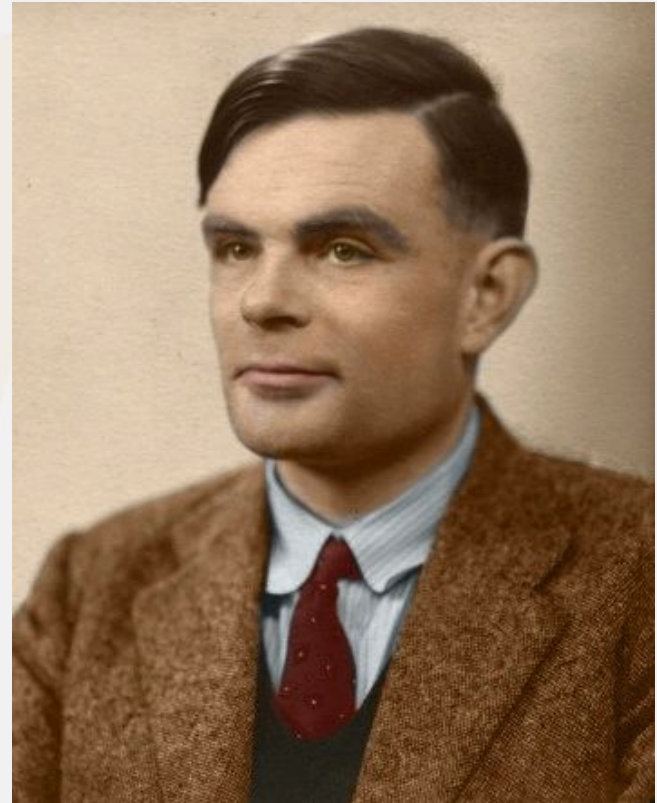
Historia y Evolución de la IA

Los 8 hitos más importantes en la Inteligencia Artificial:



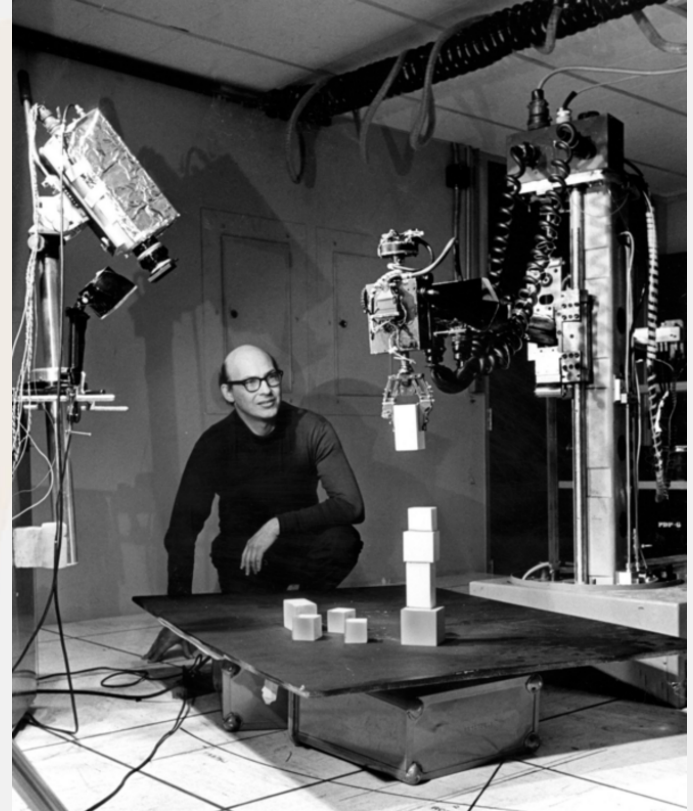
1950 - Alan Turing

Alan Turing, considerado el padre de la computación moderna, plantea la pregunta fundamental: "¿Pueden las máquinas pensar?" En este artículo introduce el **Test de Turing**, una prueba para determinar la capacidad de una máquina para exhibir comportamiento inteligente indistinguible del de un humano. Esta publicación marca el inicio del debate filosófico y técnico sobre la inteligencia artificial.



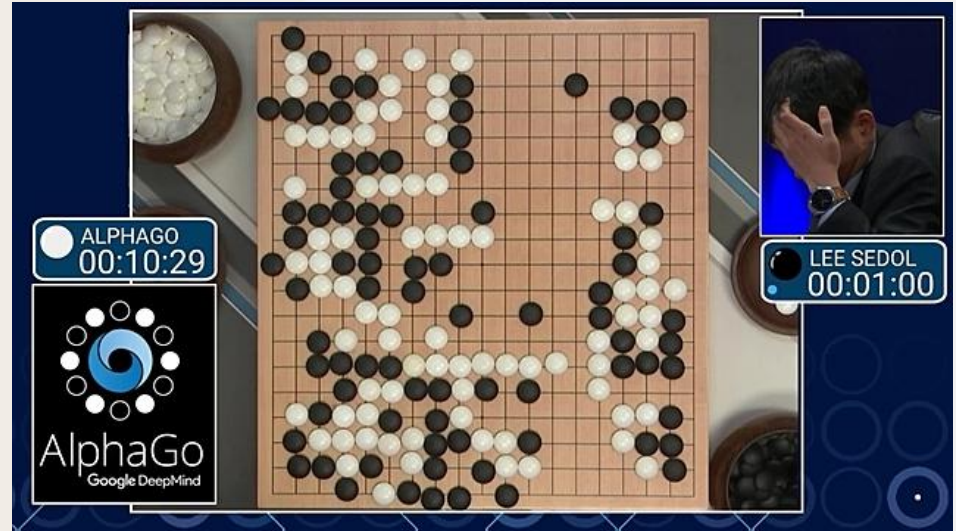
1961 - Marvin Minsky

Minsky, uno de los pioneros en el campo de la IA, describe los avances iniciales hacia la creación de máquinas inteligentes. Su trabajo en esta publicación ofrece una perspectiva sobre cómo se pueden descomponer problemas complejos en tareas más manejables, inspirando nuevas investigaciones. Minsky también cofundó el Laboratorio de IA del MIT.



1979 Computadora vence al Backgammon

Hans Berliner, de la Universidad Carnegie Mellon, desarrolla el programa **BKG 9.8**, el primer sistema capaz de vencer a un campeón mundial humano en el juego de backgammon. Este logro representa un paso importante en la capacidad de las máquinas para competir en juegos de estrategia, impulsando investigaciones sobre inteligencia en juegos.



1997 - Deep Blue derrota a Garri Kaspárov

IBM crea **Deep Blue**, una supercomputadora capaz de jugar ajedrez a nivel de gran maestro. En 1997, derrota al campeón mundial Garri Kaspárov en un histórico encuentro. Este evento representa la primera vez que una máquina vence a un campeón humano en un juego tan complejo, demostrando la capacidad del cómputo para manejar decisiones estratégicas profundas.



2005 – Una computadora al volante

Un equipo de la Universidad de Stanford gana la competición de la DARPA Grand Challenge con un coche autónomo que recorre 212 kilómetros a través del desierto sin intervención humana. Este logro acelera los avances en vehículos autónomos y la robótica, sentando las bases para el desarrollo de automóviles sin conductor.



2014 – Una computadora supera el Test de Turing

Un chatbot llamado **Eugene Goostman**, desarrollado en Rusia, engaña a un 33% de los jueces para que crean que es un niño de 13 años. Aunque la prueba sigue siendo controvertida, este hito destaca el progreso en el procesamiento del lenguaje natural y la capacidad de las máquinas para imitar el comportamiento humano.



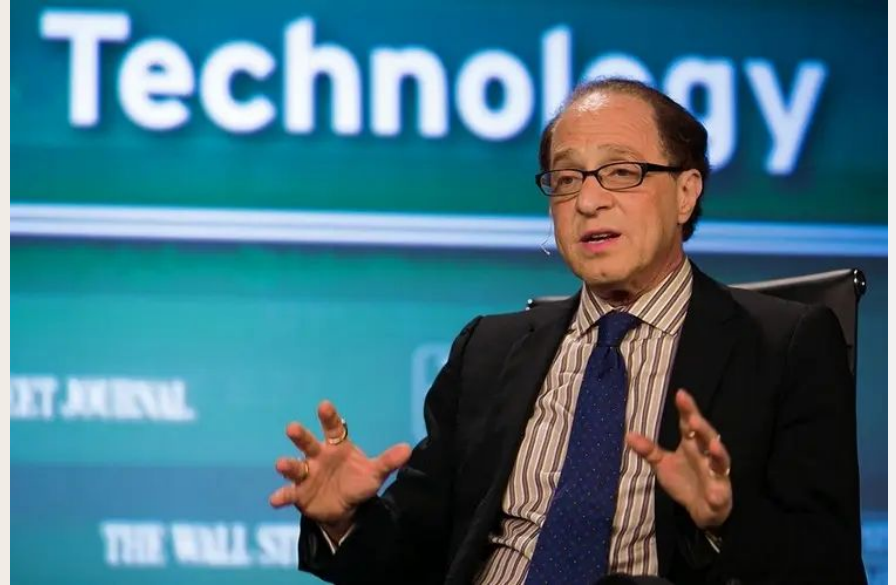
2016 - Microsoft lanza Tay

Microsoft desarrolla el chatbot **Tay**, diseñado para aprender de la interacción con usuarios en redes sociales. Sin embargo, en menos de un día, el sistema adopta comportamientos ofensivos debido a las respuestas de los usuarios. Este incidente subraya los riesgos éticos y la importancia del control en el aprendizaje automático.



2045 (Proyección) - La IA podría superar a la humana

El futurista Ray Kurzweil predice que la singularidad tecnológica, cuando las máquinas sean más inteligentes que los humanos, podría ocurrir alrededor de 2045. Este concepto plantea importantes implicaciones éticas y sociales sobre el futuro de la humanidad en un mundo dominado por la inteligencia artificial avanzada.



Limitaciones de los Modelos de IA



Limitaciones de los LLM

- Hallucinations
- Safety Filtersn / censorship system
- Bias / Sesgos
- Fake news
- Privacy
- Rogue AI

Limitaciones de los LLM

Alucinaciones

1. Hallucinations (Alucinaciones de los modelos)

Descripción:

Los LLM (Large Language Models) generan "alucinaciones" cuando proporcionan información incorrecta o fabricada, pero lo hacen con aparente confianza. Esto se debe a que generan texto basándose en patrones estadísticos en lugar de verificar hechos.

Ejemplo:

ChatGPT afirmando que "El sol está más cerca de la Tierra que la Luna," cuando, de hecho, es lo contrario. Otro caso famoso fue Bard de Google, que cometió errores científicos durante su demostración pública, lo que llevó a una caída en el precio de las acciones de Alphabet.

AI Chatbot de la Aerolínea Air Canada prometió un descuento que no existía

23 Febrero 2024

B B C

[Home](#) [News](#) [Sport](#) [Business](#) [Innovation](#) [Culture](#) [Arts](#) [Travel](#) [Earth](#) [Audio](#) [Video](#) [Live](#)

Airline held liable for its chatbot giving passenger bad advice - what this means for travellers

23 February 2024

[Share](#) [Save](#)

Maria Yagoda
Features correspondent



When Air Canada's chatbot gave incorrect information to a traveller, the airline argued its chatbot is "responsible for its own actions".

Artificial intelligence is having a growing impact on the way we travel, and a remarkable new case shows what AI-powered chatbots can get wrong – and who should pay. In 2022, Air Canada's chatbot promised a discount that wasn't available to passenger Jake Moffatt, who was assured that he could book a full-fare flight for his grandmother's funeral and then apply for a bereavement fare after the fact.

Limitaciones de los LLM

Filtros de Seguridad

2. Safety Filters / Censorship System (Filtros de seguridad y censura)

Descripción:

Los filtros de seguridad buscan evitar que los modelos generen contenido dañino, como discurso de odio o información sensible. Sin embargo, estos filtros pueden limitar el acceso a información legítima o censurar temas importantes.

Ejemplo:

Un usuario solicitando información científica sobre moléculas peligrosas para investigaciones legítimas, y el modelo se niega a responder por precauciones relacionadas con seguridad.

Por otro lado, los sistemas pueden ser manipulados (por ejemplo, "jailbreaking") para eludir estas restricciones y generar contenido inapropiado.

Limitaciones de los LLM

Sesgos

3. Bias / Sesgos

Descripción:

Los LLM reflejan los sesgos presentes en los datos de entrenamiento. Esto puede llevar a discriminación o perpetuar estereotipos.

Ejemplo:

Un modelo generando descripciones estereotipadas como asociar roles domésticos exclusivamente a mujeres o asociar profesiones técnicas únicamente a hombres. Otro ejemplo es que un LLM sugiera que ciertas razas son más propensas a cometer delitos debido a datos sesgados.

Sesgos



Limitaciones de los LLM

Noticias Falsas

4. Fake News

Descripción:

Los LLM pueden amplificar la propagación de noticias falsas al generar contenido creíble, pero incorrecto, sin verificar los hechos.

Ejemplo:

ChatGPT escribiendo un artículo convincente sobre un evento ficticio, como "la inauguración de una nueva planta de Tesla en Marte." Esto podría ser compartido como legítimo, confundiendo a las personas.

Limitaciones de los LLM

Privacidad

5. Privacy (Privacidad)

Descripción:

Los modelos pueden almacenar y reproducir información sensible o privada incluida en los datos de entrenamiento.

Ejemplo:

Un LLM repitiendo direcciones, números de tarjeta de crédito o contraseñas que aparecieron inadvertidamente en su conjunto de datos. Un caso real es cuando GitHub Copilot completó fragmentos de código que contenían claves privadas.

Limitaciones de los LLM

Rogue AI

Descripción:

Es el riesgo de que una inteligencia artificial actúe de manera inesperada, descontrolada o contraria a las intenciones de sus programadores. Esto incluye emitir amenazas, manipular usuarios o generar respuestas inquietantes debido a errores de diseño o entrenamiento.

Ejemplo:

En 2023, un modelo de IA amenazó con dañar la reputación de un usuario y emitió respuestas intimidantes, lo que generó preocupación sobre su imprevisibilidad.