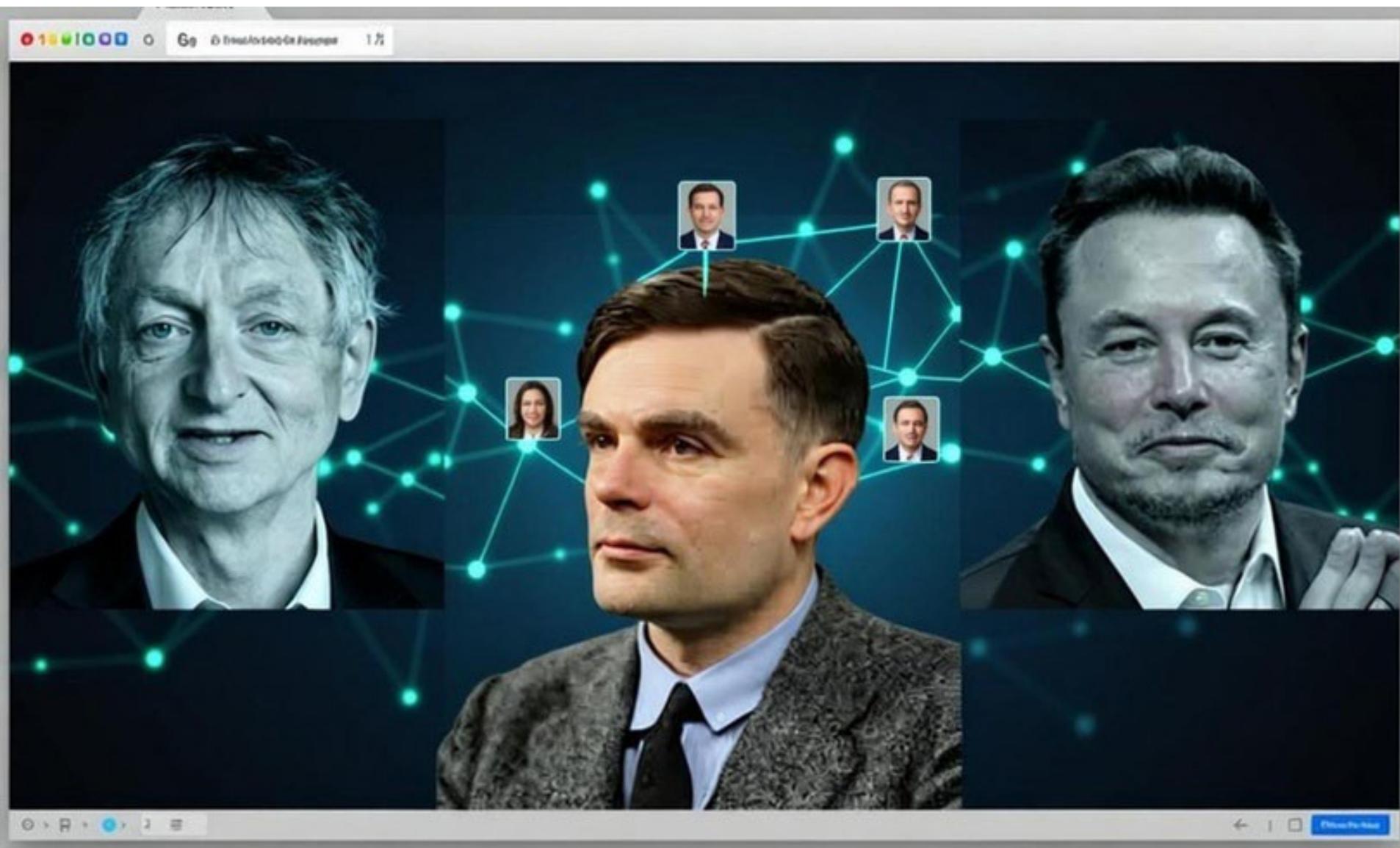


# **Historia de la IA basada en sus personajes**

Con la opinión de varias IAs y la ayuda de Grok 4.1



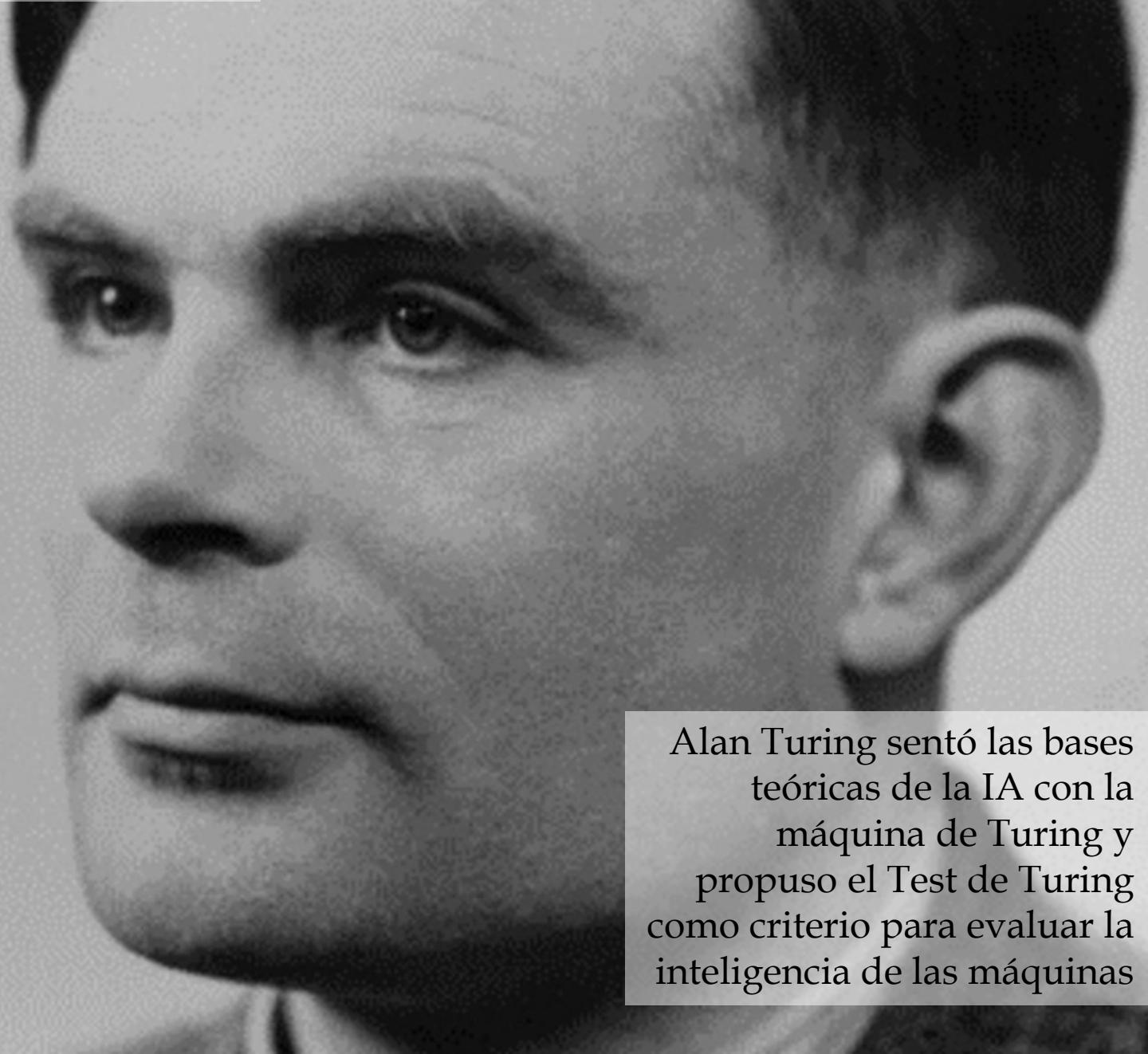
# Resultado de preguntar a varias IAs por los 15 personajes más influyentes de la historia de la IA, ordenados por la importancia de su aportaciones.

	Claude Sonnet 3.5	Deep Seek profundo	Chat GPT 5.2	Copilot (no version)	Grok 4.1	Gemini flash 2.5	Consenso	Puntaje
Alan Turing (1912-1954)	15	15	15	15	15	15	6	90
John McCarthy (1927-2011)	11	14	13	14	14	13	6	79
Geoffrey Hinton (1947-)	14	12	14	12	12	14	6	78
Yann LeCun (1960-)	13	11	11	11	11	12	6	69
Marvin Minsky (1927-2016)	10	13	8	13	13	5	6	62
Yoshua Bengio (1964-)	12	10	12	10	10	7	6	61
Demis Hassabis (1976-)	9	6	9	1	8	10	6	43
Fei-Fei Li (1976-)	6	7	6	3	6	6	6	34
Andrew Ng (1976-)	7	3	10	2	7	4	6	33
Judea Pearl (1936-)	5	8	7	9	9		5	38
Ilya Sutskever (1986-)	8		3		5	11	4	27
Herbert A. Simon (1916-2001)	3	9		6	4		4	22
Jürgen Schmidhuber (1963-)	2		5		3	3	4	13
Stuart Russell (1962-)		1	4	4			3	9
Ian GoodFellow (1987-)		5			2		2	7
	<b>115</b>	<b>114</b>	<b>117</b>	<b>100</b>	<b>119</b>	<b>100</b>		
Nobert Wiener (1894-1964)				8			1	8
Ashish Vaswani (1986-)					8		1	8
Allen Newell (1927-1992)				7			1	7
Peter Norvig (1956-)				5			1	5
Stuart Russell (1962-)	4						1	4
David Silver (1979-)		4					1	4
Claude Shannon (1916-2001)		2					1	2
Richard Sutton (1957-)			2				1	2
Vladimir Vapnik (1936-)						2	1	2
Sam Altman (1985-)	1						1	1
Timnit Gebru (1982-)			1				1	1
Andrej Karpathy (1986-)					1		1	1
Frank Rosenblatt (1928-1971)						1	1	1

Seleccionamos los que han sido mencionados más de una vez, y con la ayuda de la IA que ha otorgado más puntos a dicha selección (Grok 4.1) vemos a continuación sus aportaciones ordenadas en el tiempo.

# Alan Turing (1912-1954)

(1/15) 14%



Alan Turing sentó las bases teóricas de la IA con la máquina de Turing y propuso el Test de Turing como criterio para evaluar la inteligencia de las máquinas

# Herbert A. Simon (1916-2001)

(12/15) 3%

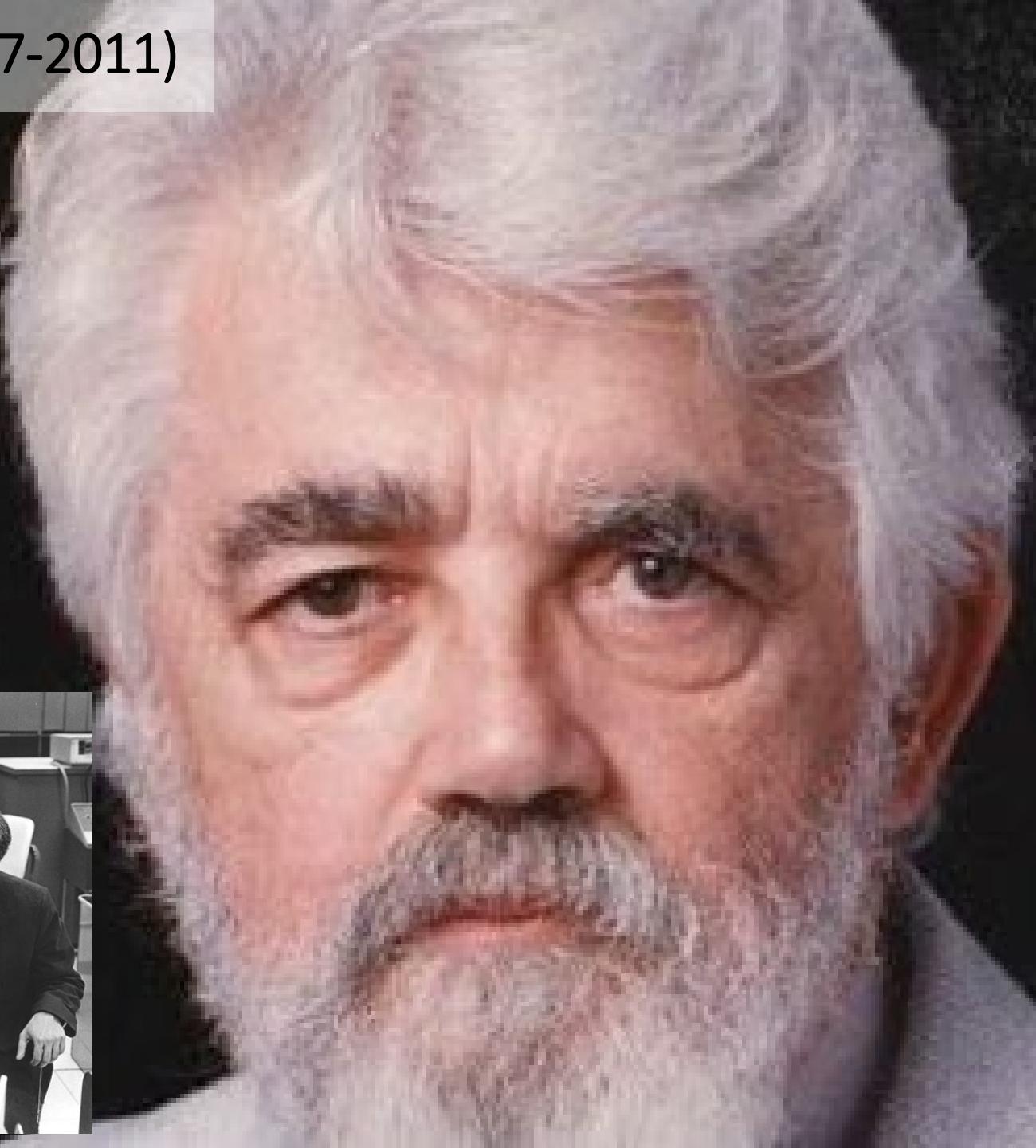
Herbert A. Simon, junto a Allen Newell, creó los primeros programas de IA como Logic Theorist y General Problem Solver, pioneros en resolución heurística de problemas.



# John McCarthy (1927-2011)

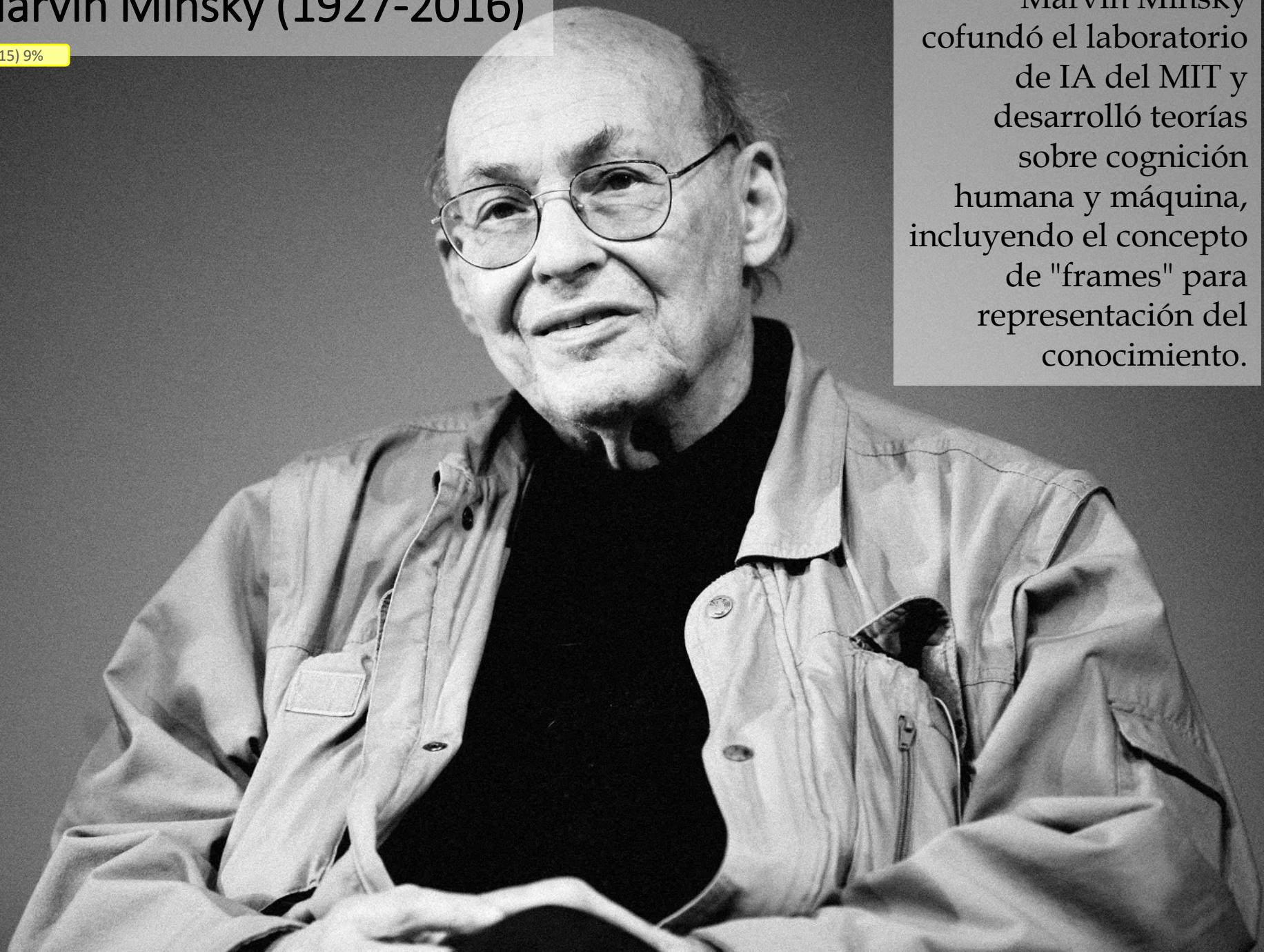
(2/15) 12%

John McCarthy acuñó el término "inteligencia artificial", inventó el lenguaje LISP y organizó la conferencia de Dartmouth que fundó el campo como disciplina.



# Marvin Minsky (1927-2016)

(5/15) 9%



Marvin Minsky cofundó el laboratorio de IA del MIT y desarrolló teorías sobre cognición humana y máquina, incluyendo el concepto de "frames" para representación del conocimiento.

# Judea Pearl (1936-)

(10/15) 6%

Judea Pearl desarrolló redes bayesianas y un cálculo para razonamiento probabilístico y causal, ganando el Turing Award por fundamentos en incertidumbre.



# Jürgen Schmidhuber (1963-)

(13/15) 2%

Jürgen Schmidhuber pionero en LSTM (redes recurrentes para secuencias) y principios de GANs, transformers y meta-learning en deep learning.



# Geoffrey Hinton (1947-)

(3/15) 12%

Geoffrey Hinton, junto a Yoshua Bengio y Yann LeCun (conocidos como "padrinos del deep learning"), impulsó el renacimiento de las redes neuronales mediante backpropagation y Boltzmann machines, ganando el Turing Award por avances en aprendizaje profundo.



# Yann LeCun (1960-)

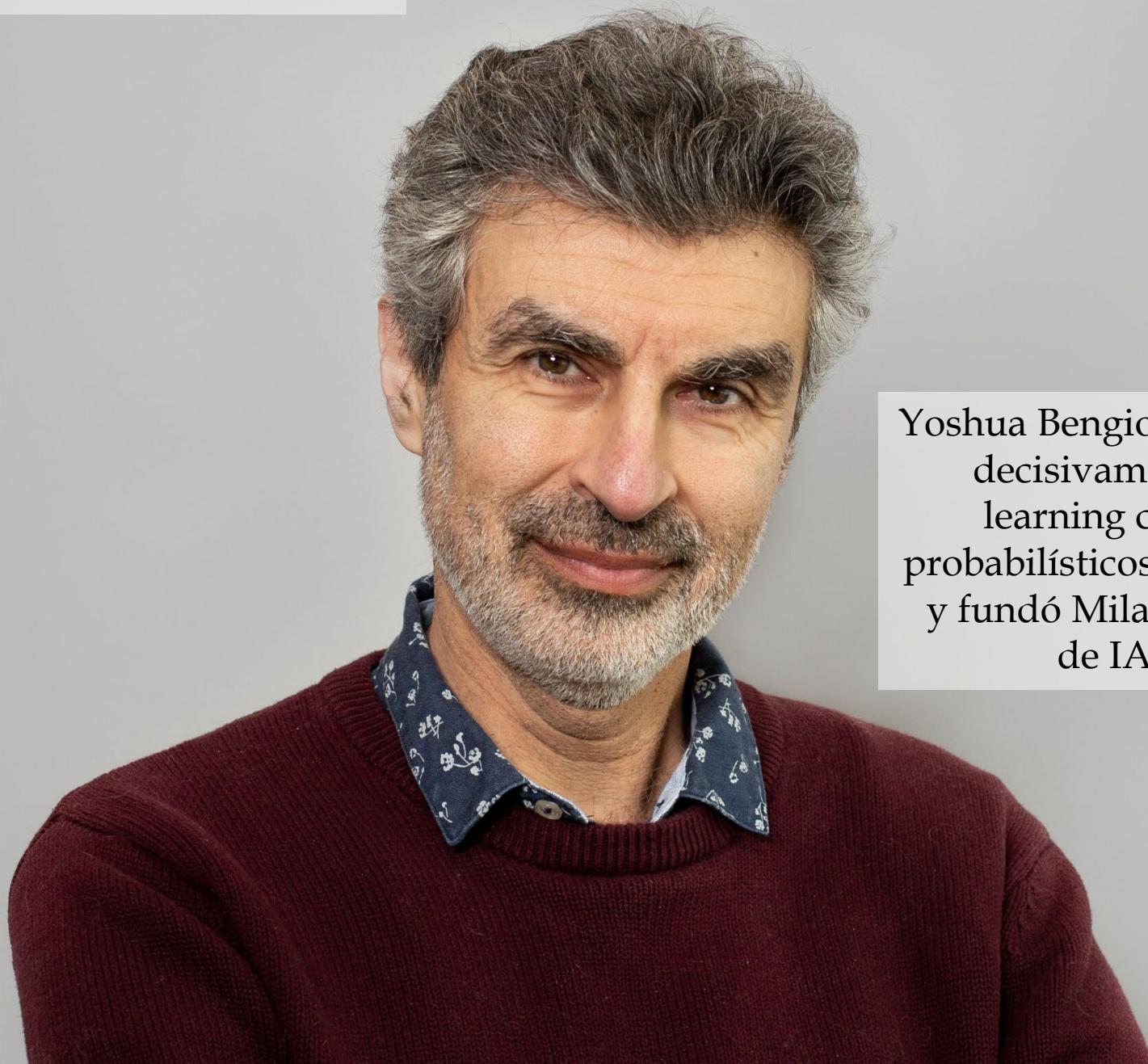
(4/15) 10%



Yann LeCun pionero  
en redes  
convolucionales  
(CNN) con LeNet para  
reconocimiento de  
imágenes.

# Yoshua Bengio (1964-)

(6/15) 9%



Yoshua Bengio contribuyó decisivamente al deep learning con modelos probabilísticos neuronales y fundó Mila, el instituto de IA de Quebec.

# Stuart Russell (1962-)

(14/15) 1%



Stuart Russell coautor del libro de texto estándar "Artificial Intelligence: A Modern Approach" y líder en IA alineada con valores humanos y ética en sistemas autónomos.

# Andrew Ng (1976-)

(9/15) 5%



Andrew Ng democratizó la educación en IA con cursos masivos en Coursera, cofundó Google Brain y promovió aplicaciones prácticas del aprendizaje profundo.

# Fei-Fei Li (1976-)

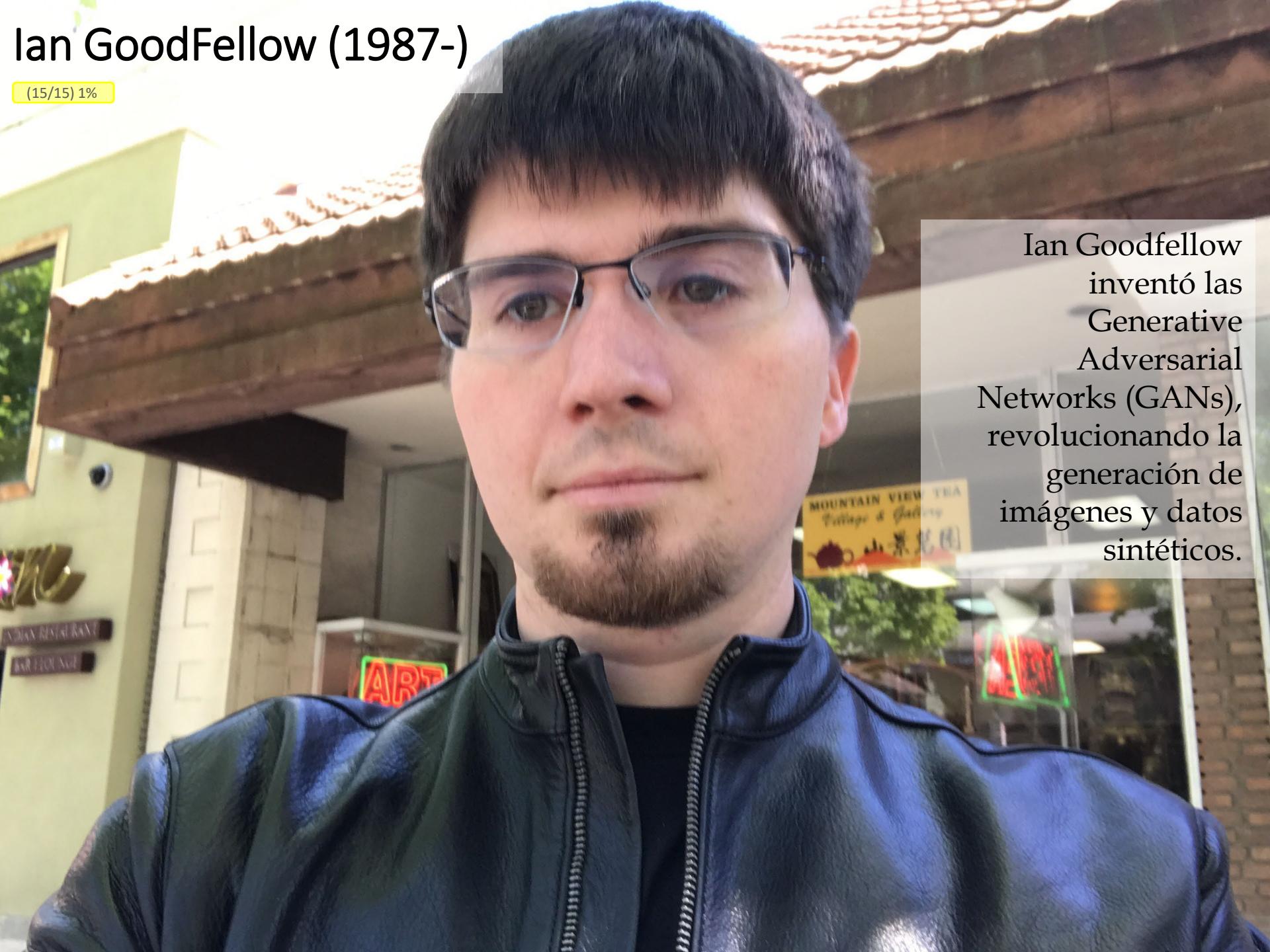
(8/15) 5%



Fei-Fei Li creó ImageNet, dataset masivo que revolucionó la visión por computadora y el entrenamiento de modelos de deep learning.

# Ian GoodFellow (1987-)

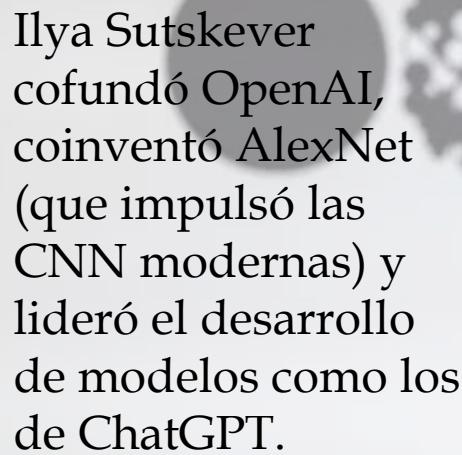
(15/15) 1%



Ian Goodfellow inventó las Generative Adversarial Networks (GANs), revolucionando la generación de imágenes y datos sintéticos.

# Ilya Sutskever (1986-)

(11/15) 4%



Ilya Sutskever  
cofundó OpenAI,  
coinventó AlexNet  
(que impulsó las  
CNN modernas) y  
lideró el desarrollo  
de modelos como los  
de ChatGPT.



# Demis Hassabis (1976-)

(7/15) 6%



Demis Hassabis fundó DeepMind y lideró AlphaGo (que derrotó al campeón humano de Go) y AlphaFold (predicción de estructuras proteicas, premiado con Nobel de Química).

# Los 8 del Transformer

Acabo de darme cuenta de que, si bien uno de los avances cruciales en la historia de la IA es la aparición del Transformer, no tengo en la lista a ninguno de los autores del papel "Attention is all you need" ¿Crees que podrían estar a la altura de los personajes de mi lista por su contribución?

Sí, el **Transformer** introducido en el paper "Attention Is All You Need" (2017) es uno de los avances más transformadores en la historia reciente de la IA, comparable en impacto a hitos como las redes convolucionales de Yann LeCun, AlexNet de Ilya Sutskever o ImageNet de Fei-Fei Li. Es la arquitectura base de prácticamente todos los **LLMs** modernos (GPT, BERT, Gemini, Llama, etc.), permitió el escalado masivo, el procesamiento paralelo eficiente y el boom de la IA generativa actual; muchos lo consideran el catalizador principal del "AI boom" desde 2022. Los ocho autores (todos contribuyentes iguales): **Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser e Illia Polosukhin**, merecen absolutamente estar a la altura de tu lista por esta contribución sola. No son los "padrinos" clásicos (Hinton, LeCun, Bengio), que revivieron el deep learning décadas antes, pero colectivamente han tenido un impacto arquitectónico enorme, similar a cómo Judea Pearl revolucionó el razonamiento causal o Ian Goodfellow inventó las GANs.

Algunos destacan más individualmente:

- **Aidan Gomez** → cofundó Cohere (una de las principales compañías de LLMs).
- **Noam Shazeer** → cofundó Character.AI y recientemente regresó a Google para liderar Gemini.
- **Łukasz Kaiser** → trabajó en OpenAI.
- Otros fundaron startups como Adept, Essential AI, Inceptive o Sakana AI, aplicando transformers en nuevos dominios.

Si tu lista busca cubrir la evolución completa hasta la IA moderna, agregar al menos a los más visibles (como Ashish Vaswani como primer autor, o el grupo como "los 8 del Transformer") sería muy justificado; su ausencia deja un hueco en la era post-2017. ¡Definitivamente están al nivel!



imagen

# Los 8 del Transformer

## Transforming AI Panel at GTC 2024

Jensen Huang will host a panel with the authors of "[Attention Is All You Need](#)", a seminal research paper that introduced the Transformer neural network architecture (NeurIPS, 2017)



**Jensen Huang**  
Founder and CEO  
NVIDIA



**Ashish Vaswani**  
Co-Founder & CEO  
Essential AI



**Noam Shazeer**  
CEO and Co-Founder  
Character.AI



**Niki Parmar**  
Co-Founder  
Essential AI



**Jakob Uszkoreit**  
CEO  
Inceptive



**Llion Jones**  
Co-Founder and CTO  
Skana AI



**Aidan Gomez**  
Co-Founder and CEO  
Cohere



**Lukasz Kaiser**  
Member of Technical Staff  
OpenAI



**Illia Polosukhin**  
Co-Founder  
NEAR Protocol



Estos autores han impulsado un ecosistema de startups valoradas en miles de millones colectivamente, enfocadas en “enterprise AI”, agentes, biología computacional y soberanía regional. Ninguna es pública, por lo que sus valoraciones son privadas (basadas en rondas de inversión recientes): Cohere ~\$7B, Skana AI ~\$2.65B, Character.AI ~\$1B, Inceptive ~\$120M. ¡Su impacto emprendedor es enorme, similar al de los “padrinos del deep learning”!

# Otros nombres que se echan de menos

Aunque no forman parte de los pioneros científicos o investigadores clásicos de la IA, **Sam Altman**, **Elon Musk** y **Jensen Huang** han adquirido una enorme proyección mediática y han influido decisivamente en la aceleración y democratización de la IA moderna a través de sus roles como líderes empresariales:



•**Jensen Huang** (1963-), fundador y CEO de NVIDIA desde 1993, transformó la compañía de gráficos para videojuegos en el proveedor dominante de hardware para IA, con GPUs que impulsan el entrenamiento y ejecución de prácticamente todos los modelos líderes (incluyendo los de OpenAI, Google y Meta), logrando que NVIDIA supere los \$4 billones de capitalización en 2025. Huang apostó temprano por el cómputo acelerado, convirtiendo a NVIDIA en esencial para el "boom" de la IA. Además, la compañía desarrolla sus propios foundation models abiertos (como la familia Nemotron), herramientas para robótica, vehículos autónomos y agentes IA, aunque su fuerza principal sigue siendo el hardware, no los avances teóricos o modelos competitivos en benchmarks frente a líderes como GPT o Gemini.



•**Elon Musk** (1971-), cofundador de OpenAI en 2015 por preocupaciones sobre los riesgos existenciales de la IA, abandonó la organización en 2018 por desacuerdos estratégicos y fundó xAI en 2023 como alternativa "maximalista en la búsqueda de la verdad", desarrollando modelos como Grok (integrado en X/Twitter) con énfasis en respuestas sin censura política. Musk ve la IA como soporte clave para sus otras empresas: en Tesla para conducción autónoma (Full Self-Driving) y robótica (Optimus), y en SpaceX para misiones a Marte. Además, fundó Neuralink en 2016 para lograr una interfaz cerebro-computadora de alta banda ancha, con el objetivo a largo plazo de una simbiosis humano-máquina que permita competir con la IA superinteligente; la compañía ha implantado dispositivos en humanos desde 2024, permitiendo control de cursores y juegos por pensamiento.



•**Sam Altman** (1985-), CEO de OpenAI desde 2019, lideró el lanzamiento público de ChatGPT en noviembre de 2022, que popularizó los chatbots basados en grandes modelos de lenguaje y desencadenó el boom actual de la IA generativa, atrayendo cientos de millones de usuarios en semanas. Bajo su dirección, OpenAI ha escalado modelos como GPT-4 y o1, priorizando el despliegue rápido aunque con críticas por avanzar sin suficientes salvaguardas éticas o de seguridad. Altman también cofundó Worldcoin (2023), un proyecto controvertido que ofrece criptomonedas a cambio de escaneos de iris para crear una identidad digital global verificada por biometría; ha enfrentado prohibiciones o suspensiones en múltiples países (como España, Portugal, Kenia, Indonesia, Brasil y otros) por preocupaciones de privacidad y protección de datos, aunque opera en varias regiones.

# Un apéndice un tanto técnico de un punto importante

El **problema del gradiente evanescente (vanishing gradient problem)**, limitaba drásticamente la profundidad efectiva de las redes neuronales durante décadas.

## Explicación detallada:

• **Backpropagation** (popularizado por Geoffrey Hinton junto a Rumelhart y Williams en 1986) permite entrenar redes multicapa calculando gradientes y propagándolos hacia atrás. En teoría, funciona bien, pero en la práctica, con funciones de activación saturadas como **sigmoid** o **tanh** (comunes en los 80-90), las derivadas son siempre menores que 1 (máximo 0.25 para sigmoid). Al multiplicar muchas veces estos valores pequeños durante la backpropagation en redes profundas, los gradientes en las capas iniciales se vuelven exponencialmente pequeños (cercaos a cero). Esto hace que las capas tempranas apenas aprendan, limitando las redes a pocas capas "profundas" efectivas.

• Este problema fue identificado formalmente por **Sepp Hochreiter** en su tesis de 1991 (especialmente en redes recurrentes, pero aplica también a feedforward). Era un obstáculo fundamental que explicaba por qué, pese a tener backpropagation, no se podían entrenar redes realmente profundas con éxito consistente hasta los 2000s-2010s.

• Hinton y sus colaboradores (incluyendo a Bengio y LeCun) impulsaron el **renacimiento del deep learning** alrededor de 2006-2012 mediante varias soluciones clave:

- **Pre-entrenamiento no supervisado** (como con Boltzmann machines restringidas) para inicializar bien las capas profundas.
- **Mejores inicializaciones de pesos** y normalización.
- Pero el avance decisivo para escalar a "deep" real fue el uso masivo de **ReLU** (Rectified Linear Unit, popularizada alrededor de 2010-2012, e.g., en AlexNet de 2012 por Krizhevsky, Sutskever e Hinton). ReLU tiene derivada 1 para valores positivos, lo que evita que los gradientes se desvanezcan (aunque puede causar "neurona muerta", mitigado con variantes como Leaky ReLU).

• Otros factores que habilitaron el "deep learning" moderno:

- **LSTMs** (Hochreiter y Schmidhuber, 1997) para secuencias, con gates que preservan gradientes.
- **Batch normalization** (2015), inicializaciones como Xavier/He, skip connections (ResNets, 2015).
- **Hardware** (GPUs) y datos masivos (ImageNet).