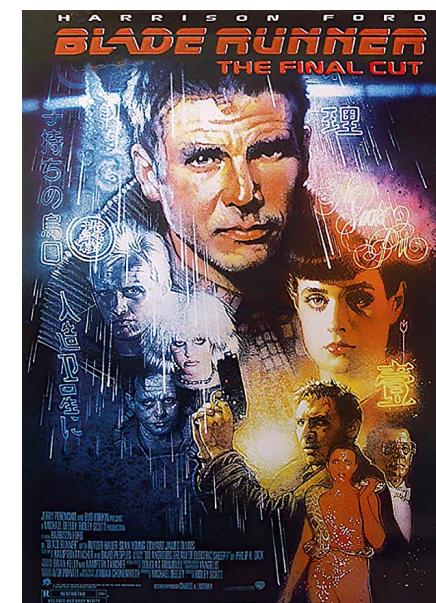
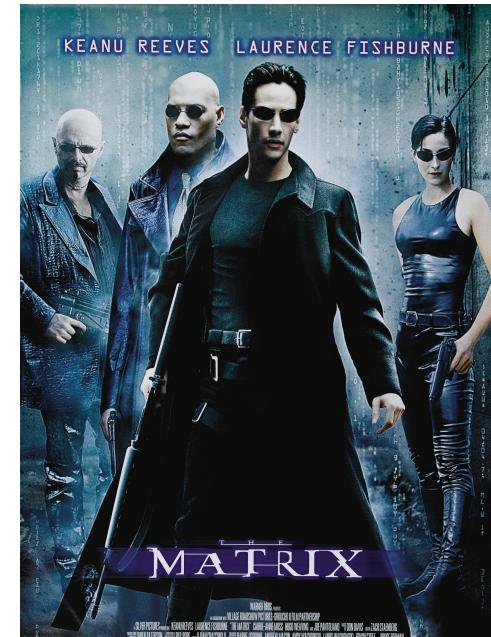
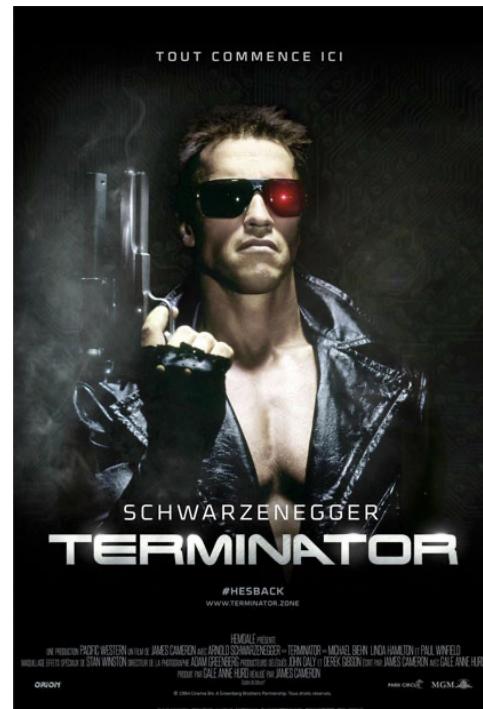
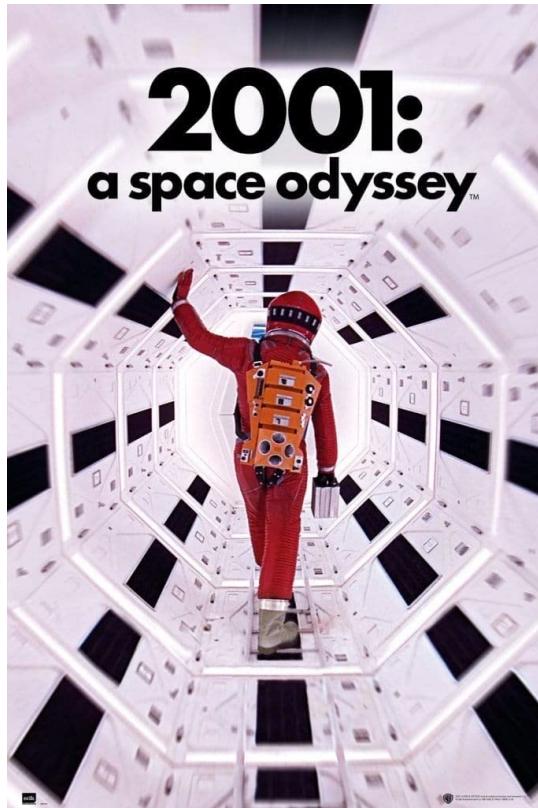


# Intelligenza Artificiale tra Passato, Presente e Futuro

Andrea Brunello

Cordovado, 7 luglio 2022





- Istante in cui la crescita tecnologica diventa **incontrollabile ed irreversibile**, con conseguenze imprevedibili per la civiltà
- Creazione di un agente intelligente, in grado di auto-migliorarsi sempre più velocemente, fino a generare una superintelligenza
- Il termine «singolarità» riflette l'idea che tale cambiamento potrebbe avvenire improvvisamente, senza essere previsto
- Concetto sostenuto da Elon Musk, Stephen Hawking, ...
- S. Hawking: «*Il successo nella creazione dell'IA sarebbe il più grande evento nella storia umana. Sfortunatamente, potrebbe anche essere l'ultimo, a meno che non impariamo come evitarne i rischi*»
- Nick Bostrom e Vincent C. Müller hanno «stimato» una probabilità del 50% che tale fenomeno possa verificarsi fra il 2040 e il 2050

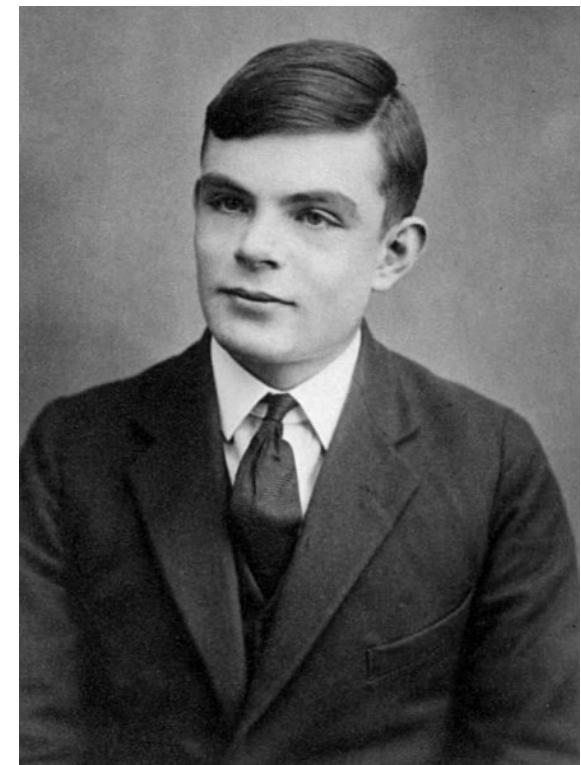
## IA Forte (Generale):

- «*Il computer, opportunamente programmato con i giusti input e output, avrebbe quindi una mente esattamente nello stesso senso in cui gli esseri umani hanno la mente*» (Searle, 1999)
- Macchina che è in grado di applicare intelligenza alla risoluzione di qualsiasi problema

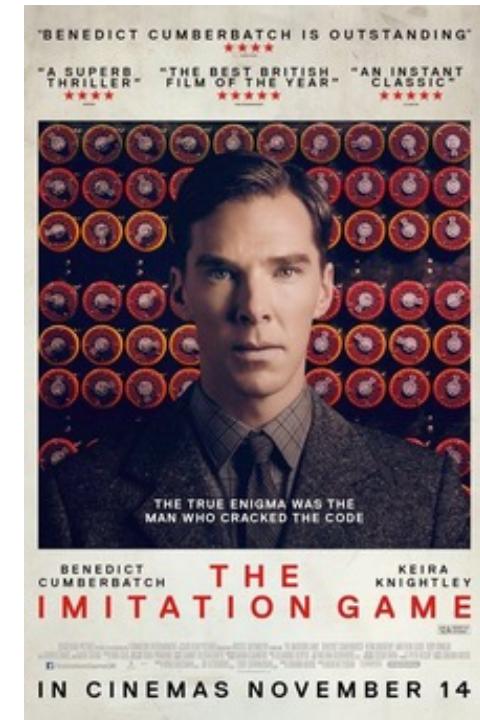
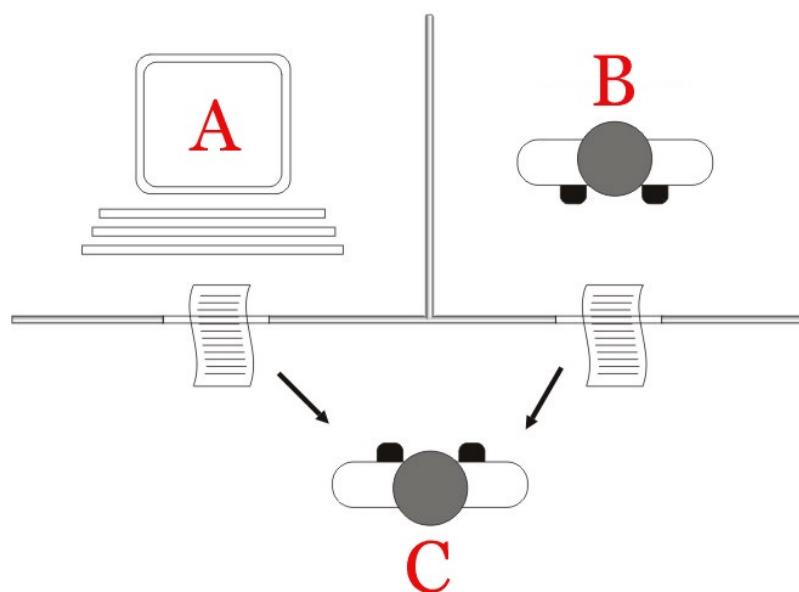
## IA Debole (Ristretta):

- Implementa una parte limitata della mente, o risolve uno specifico problema
- Ad esempio: riconoscimento di immagini, traduzione

- Matematico, logico, crittografo e filosofo britannico
- Diede un contributo fondamentale alla violazione dei cifrari tedeschi, come Enigma
- Considerato uno dei padri dell'**informatica moderna**
- Formalizzò per primo i concetti di **Algoritmo** e di **Computazione**, attraverso la Macchina di Turing



- Oltre a ciò, si interessò alla questione di definire un concetto di comportamento intelligente per le macchine
- Proposta del **Test di Turing** («*Computing Machinery and Intelligence*», *Mind*, 1950)

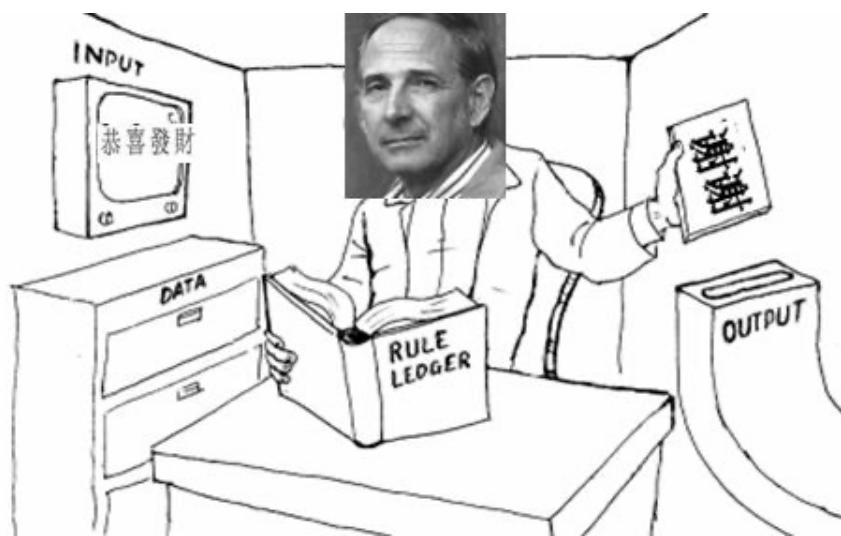


Nel corso degli anni, il Test di Turing è stato oggetto di critiche da parte di diversi scienziati e filosofi

- E' basato unicamente sul linguaggio
- E' antropocentrico
- Non tiene in alcun modo in considerazione lo **stato interno della macchina**



- Un computer digitale che esegue un programma **non può** avere una «mente», «comprensione» o «coscienza»
- Questo indipendentemente da quanto intelligente o umano il programma possa far sembrare il computer



If you see this shape,  
"什麼"  
followed by this shape,  
"帶來"  
followed by this shape,  
"快樂".

then produce this shape,  
"爲天"  
followed by this shape,  
"下式".



Storicamente, troviamo due approcci all'Intelligenza Artificiale:

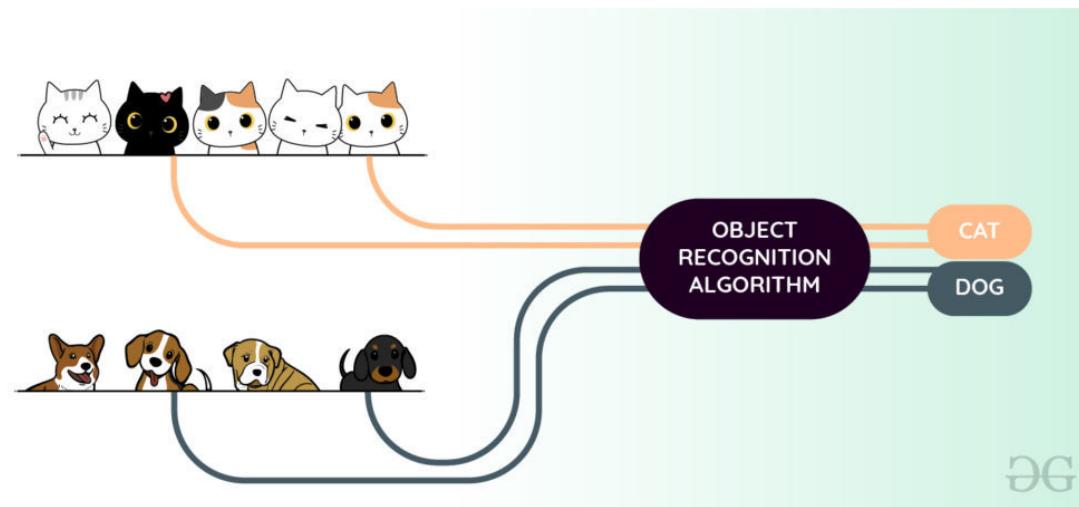
- **Simbolico** (sistemi esperti)
- **Connettivista** (reti neurali)



*Figure 1. Conflict between theoretical extremes.*

- Gli esseri umani ragionano attraverso «simboli» per assegnare significato a oggetti ed eventi
- L'idea dietro all'approccio simbolico è che questi simboli sono i mattoni di qualsiasi processo cognitivo
- Sistema di IA:
  - **Base di conoscenza:** [progetto Cyc](#)
  - **Insieme di regole**
  - **Motore inferenziale**
- Interpretabile
  - Sillogismo (Aristotele):
    - **Premessa:** Tutti gli uomini sono mortali
    - **Premessa:** Tutti i greci sono uomini
    - **Conclusione:** Tutti i greci sono mortali
- Funziona bene con applicazioni che hanno regole e obiettivi ben definiti

- Non sono gli esseri umani a codificare la conoscenza in modo esplicito
- Modelli e metodi statistici che consentono al sistema di apprendere in modo autonomo dai dati
- L'idea è cercare di riprodurre le capacità intellettuali utilizzando un modello semplificato delle funzionalità cerebrali
- Richiede grandi quantità di dati di alta qualità



Il termine «Intelligenza Artificiale» viene per la prima volta utilizzato da John McCarthy nel 1956, con l'organizzazione del **Dartmouth Summer Research Project on Artificial Intelligence**

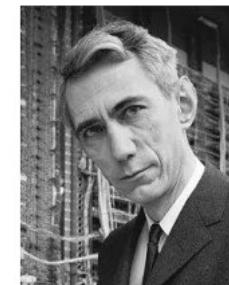
### 1956 Dartmouth Conference: The Founding Fathers of AI



John MacCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff



Alan Newell



Herbert Simon



Arthur Samuel



Oliver Selfridge



Nathaniel Rochester



Trenchard More

Da quel momento, l'Intelligenza Artificiale ha vissuto una storia fatta di continui alti e bassi

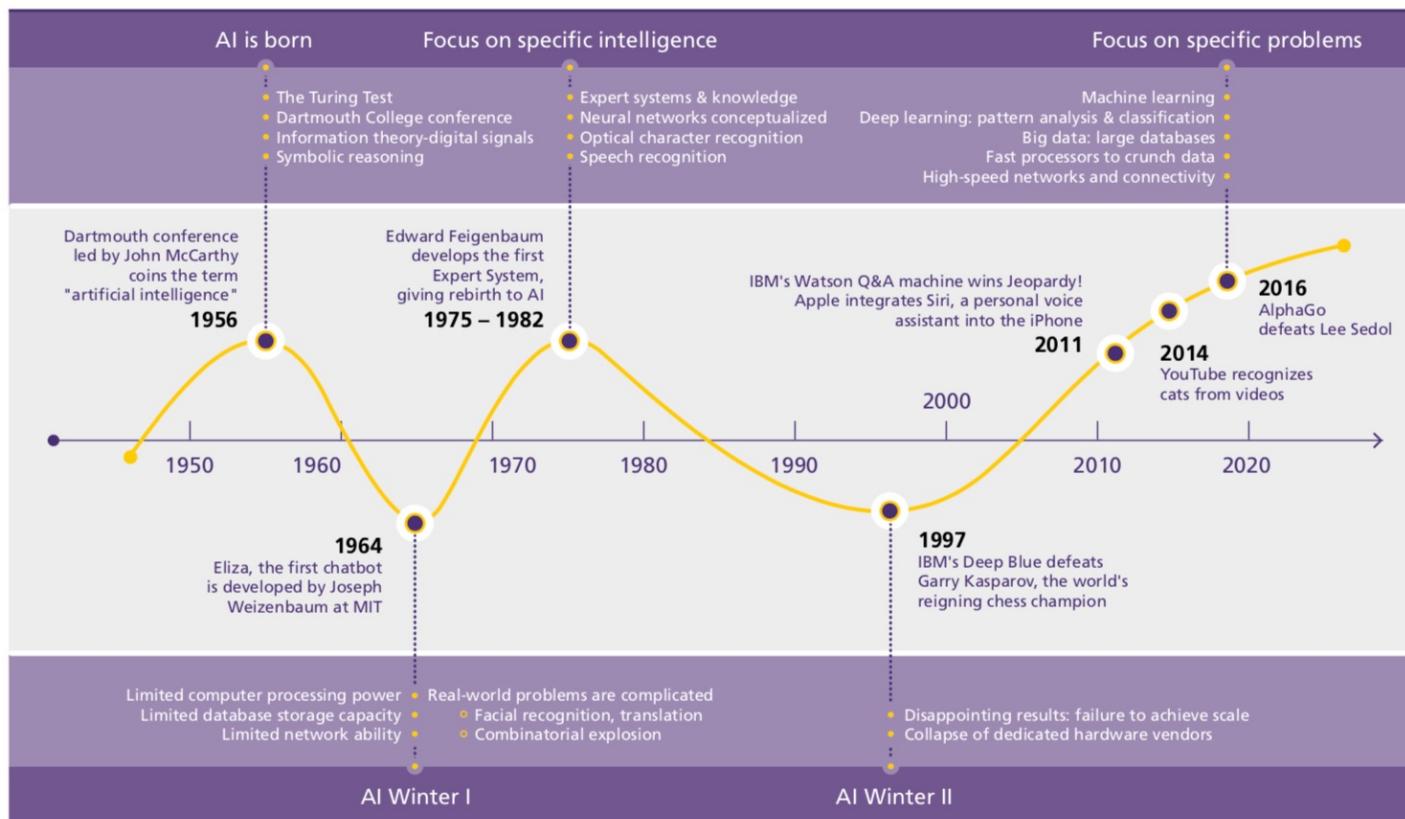


Figure 1: An AI timeline; Source: Lavenda, D./Marsden, P.

source dhl via @mikequindazzi

- I primi programmi sviluppati a seguito della conferenza apparirono, alla maggior parte delle persone, sorprendenti
- I computer iniziarono a dimostrare teoremi di geometria, a risolvere problemi matematici, e a «parlare» in inglese
- Regnava l'ottimismo:
  - **1958**, H. A. Simon e A. Newell: «*entro 10 anni un computer sarà campione del mondo di scacchi*»
  - **1965**, A. Simon: «*le macchine saranno capaci, entro 20 anni, di fare qualsiasi cosa che fanno gli esseri umani*»
  - **1965**, M. Minsky: «*entro una generazione [...] il problema di creare un'intelligenza artificiale sarà sostanzialmente risolto*»
  - **1967**, M. Minsky: «*fra 3-8 anni avremo una macchina con l'intelligenza di una persona media*»
- A fine anni '50 viene implementato il **Perceptron** (rete neurale)
- Agenzie governative come DARPA (Defense Advanced Research Project Agency) iniettarono, come risultato, quantità enormi di denaro nella ricerca, senza seguire un obiettivo specifico «*Fund people, not projects!*»

- Per quanto fossero importanti i primi risultati conseguiti dall'IA, erano in larga parte basati su una tecnica piuttosto semplice
- Per raggiungere un obiettivo, es., dimostrare un teorema con una sequenza di deduzioni logiche, procedevano passo dopo passo, un po' come si fa esplorando un labirinto
- Al raggiungimento di un vicolo cieco, tornavano indietro percorrendo una strada diversa, sino a trovare l'uscita
- I passi ammissibili, e le regole di tale «gioco», erano specificate a mano da esseri umani
- «Ragionamento come ricerca», metodo di forza bruta
- **Problema:** in certi casi, il numero di possibili percorsi attraverso il labirinto è astronomico (esplosione combinatoria)
- Approcci buoni su problemi piccoli e ben formalizzati/definiti (spoiler: il mondo reale non è fatto così)

- ELIZA (1964-1966), riconosciuto come il primo chatbot della storia
- E' in grado di simulare la conversazione con uno psicoterapeuta
- Talvolta la conversazione poteva sembrare così reale da indurre le persone a credere che stessero interagendo con un essere umano

- David Avidan, «*My Electronic Psychiatrist – Eight Authentic Talks with a Computer*»
- Tutt'oggi viene utilizzato il termine «Eliza effect»
- Implementazione di Eliza

Welcome to

```
EEEEEE LL      IIII    ZZZZZZ  AAAAAA
EE   LL      II      ZZ    AA  AA
EEEEEE LL      II      ZZZ   AAAAAAAA
EE   LL      II      ZZ    AA  AA
EEEEEE LLLLLL  IIII ZZZZZZ  AA  AA
```

Eliza is a mock Rogerian psychotherapist.

The original program was described by Joseph Weizenbaum in 1966.  
This implementation by Norbert Landsteiner 2005.

```
ELIZA: Is something troubling you ?
YOU: Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU: They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU: Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU: He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU: It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:
```

- Messi di fronte ai primi insuccessi, i ricercatori rilanciarono con frasi sempre più sensazionalistiche, nel timore di perdere i fondi
- Tuttavia, risultò chiaro ad un certo punto che li scienziati avevano grossolanamente sottovalutato le difficoltà legate all'IA:
  - Esplosione combinatoria
  - Scarse **risorse computazionali**
  - Difficoltà nella codifica del contesto e della conoscenza comune
- L'ultimo punto è estremamente rilevante nel caso della traduzione fra lingue, es., fra inglese e russo
  - «*Io spirito è forte, ma la carne è debole*»
  - «*la vodka è buona, ma la carne è marcia*»
- Sul fronte connettivista: *M. Minsky, S. Papert, Perceptrons (1969)*
- Le agenzie governative ritirarono infine i fondi, con conseguenze devastanti per la ricerca sull'Intelligenza Artificiale

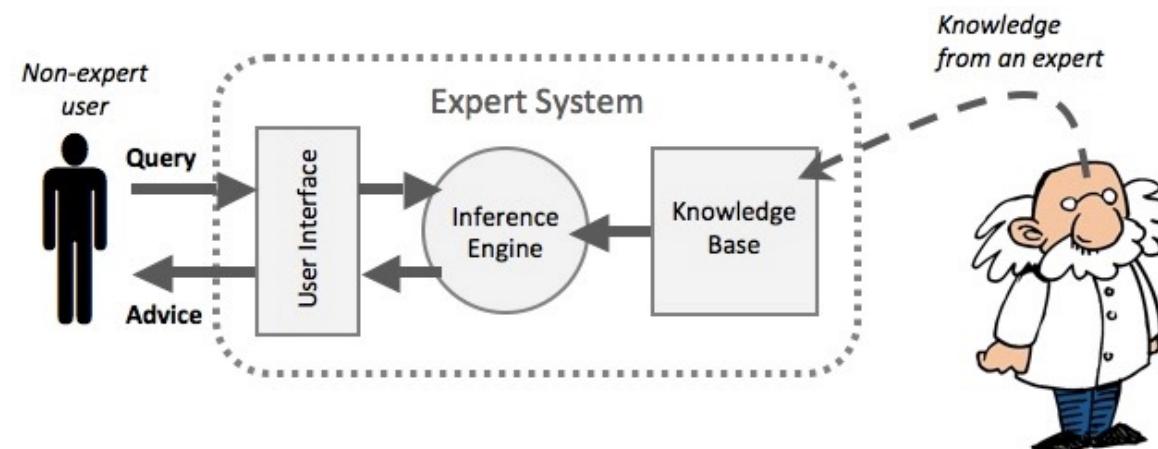
- Nel 1975 viene presentato **Cray-1**, considerato uno dei supercomputer più di successo della storia
  - CPU 80 MHz
  - Memoria 8.39 Megabyte
  - Costo di 35.000.000 \$ (del 2021)
  - Consumo di 115 kW



- 2006: Nokia n73
  - CPU 220 MHz
  - Memoria 64 Megabyte



- Gli anni '80 vedono l'esplosione di popolarità dei **sistema esperti**
- Idea: codificare la conoscenza di un insieme di esperti del dominio all'interno di una macchina, tramite fatti ed insiemi di regole
- La macchina, seguendo tali fatti e regole, sarà poi in grado di «ragionare» in modo da rispondere a domande formulate sullo stesso dominio
- Sistemi per la diagnostica, la configurazione, ... funzionano!



- Sempre più compagnie si affidano a sistemi esperti per compiti gestionali
- Nascono diverse aziende che producono hardware dedicato
- Parallelamente, si ha la rinascita dell'approccio connettivista, con la scoperta di nuove tecniche per l'addestramento delle **reti neurali**
- Nel 1981, il governo giapponese lancia il **«Fifth Generation Computer Project»**, con l'obiettivo di costruire macchine in grado di:
  - Conversare con un essere umano
  - Svolgere traduzioni fra lingue
  - Comprendere il contenuto di fotografia e immagini
  - Ragionare come gli esseri umani
- In risposta, DARPA lancia la «Strategic Computing Initiative»

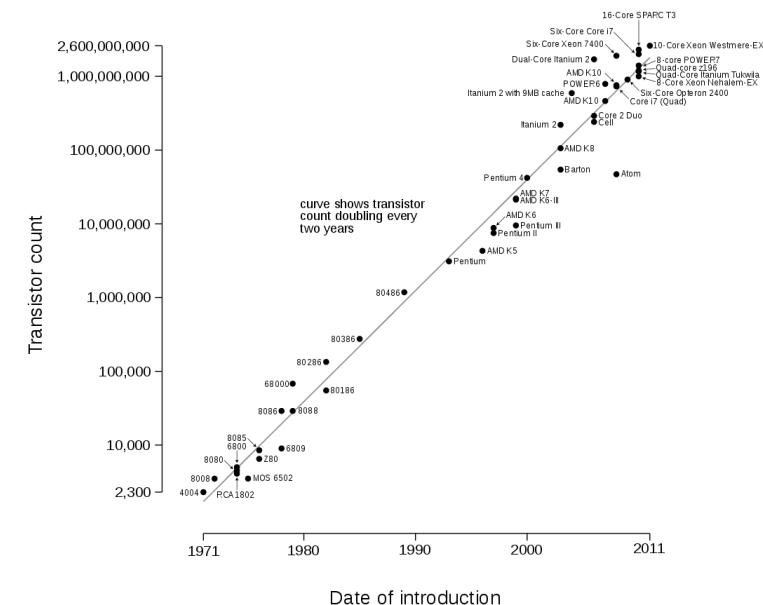
- Architetture generaliste sempre più performanti incrementano la concorrenza nei confronti dei produttori di hardware dedicato
- Contemporaneo calo della domanda di sistemi esperti, che si rivelano estremamente difficili e costosi da **gestire e mantenere** quando applicati al di fuori di contesti estremamente specifici:
  - Sistemi che non apprendono la conoscenza dai dati
  - Difficoltà di specifica: regole fragili, cablate
  - Moltissime regole: conflitti, incompletezza
- Fallimento di dozzine di compagnie produttrici di hardware
- 1992: il «Fifth Generation Computer Project» termina con un fallimento
- DARPA taglia nuovamente i fondi alla ricerca

- Sistema per la pianificazione della logistica (DART, Dynamic Analysis and Replanning Tool): nel corso della Guerra del Golfo, ripaga tutti i fondi DARPA dei precedenti 30 anni
  - Le reti neurali ottengono i primi successi nel riconoscimento di caratteri (USPS)
  - Diverse applicazioni, tipicamente passate in sordina, per risolvere problemi specifici: logistica, data mining, riconoscimento del linguaggio in domini specifici
  - In generale, la ricerca si fa molto più «cauta», sistematica, ingegnerizzata
  - N. Bostrom (2006): "*Un sacco di intelligenza artificiale all'avanguardia è filtrata in applicazioni generali, spesso senza essere chiamata IA; questo perché, una volta che qualcosa diventa sufficientemente utile e abbastanza comune, non viene più etichettata come IA*"

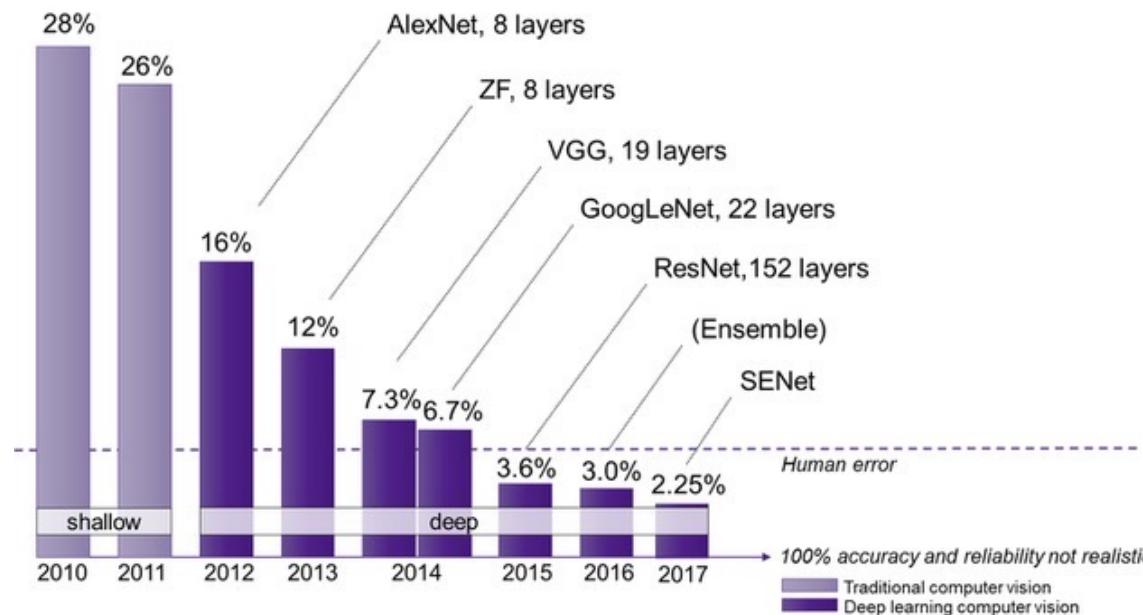
- **11 maggio 1997:** Deep Blue (IBM) diventa il primo computer a battere a scacchi un campione del mondo in carica, Garry Kasparov
- Il successo non è dovuto a nuove rivoluzionarie idee, ma ad un miglioramento delle capacità di calcolo
- Deep Blue è 10.000.000 di volte più potente delle macchine in uso nei primi anni '50: in un secondo può valutare 200.000.000 di mosse
- Legge di Moore: la velocità e le capacità di memorizzazione dei computer raddoppiano ogni due anni

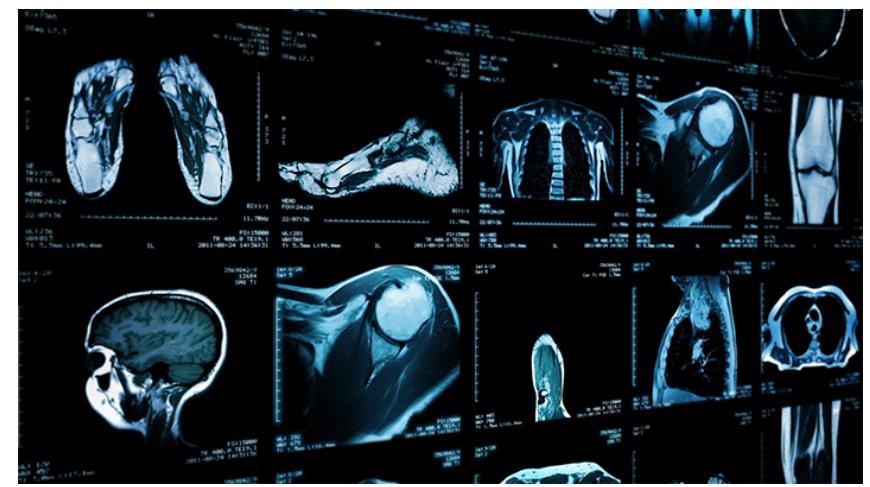


Microprocessor transistor counts 1971-2011 &amp; Moore's law



- Non si vuole costruire una IA «generale», ma la ricerca è molto più pragmatica e focalizzata su problemi specifici
- Sempre più dati a disposizione su cui addestrare soluzioni di IA (Big Data), grazie anche allo sviluppo di Internet
- Nel 2012, tramite **reti neurali**, un gruppo di ricerca improvvisamente dimezza l'errore di classificazione sul dataset ImageNet, composto da più di 14 milioni di immagini suddivise in 20.000 categorie





Shakey (1966-1972)

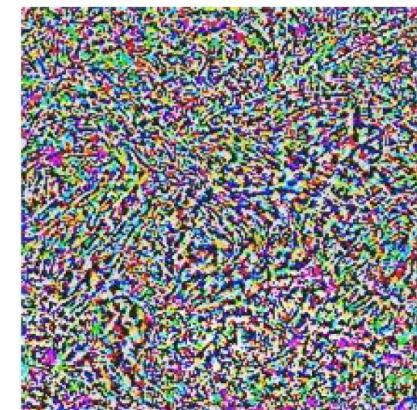
Atlas, 2021

- Ci sono ancora limiti importanti che riguardano le capacità di apprendimento di una rete neurale
- Tipicamente, sono richieste decine o centinaia di migliaia di esempi anche per imparare un semplice concetto
- I modelli addestrati possono mostrare dei comportamenti inattesi quando sottoposti ad un input mai incontrato in fase di addestramento



Dog: 96.05%

(a) Original Image



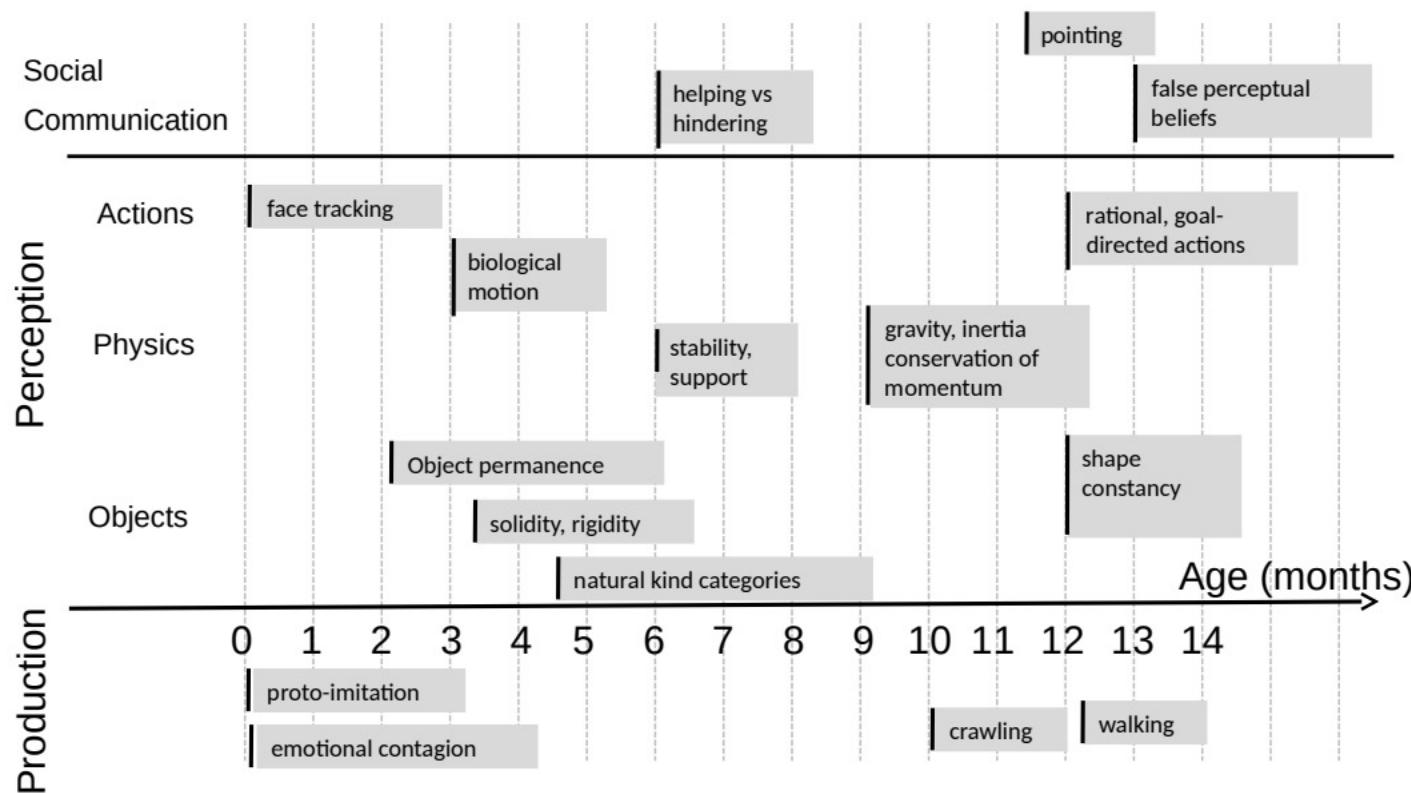
(b) Disturbance



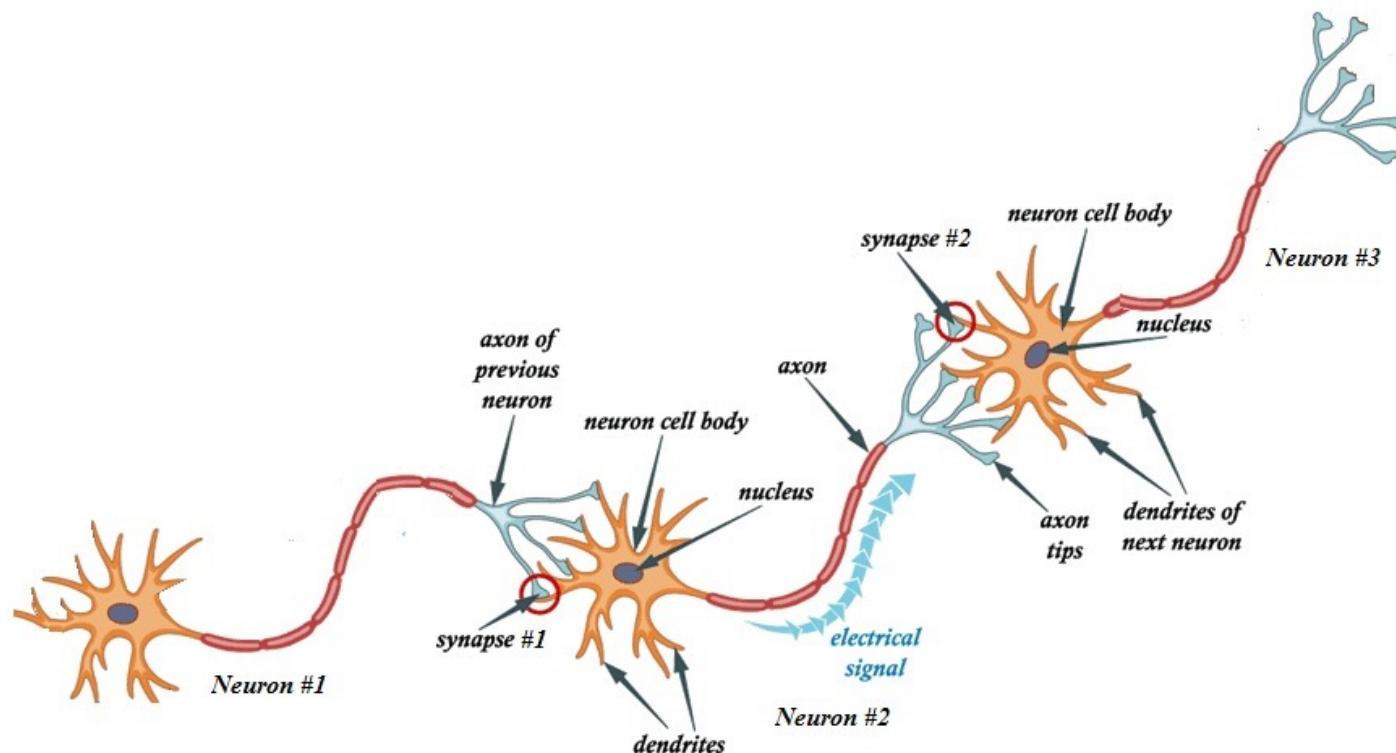
Cock: 100.00%

(c) Adversarial Examples

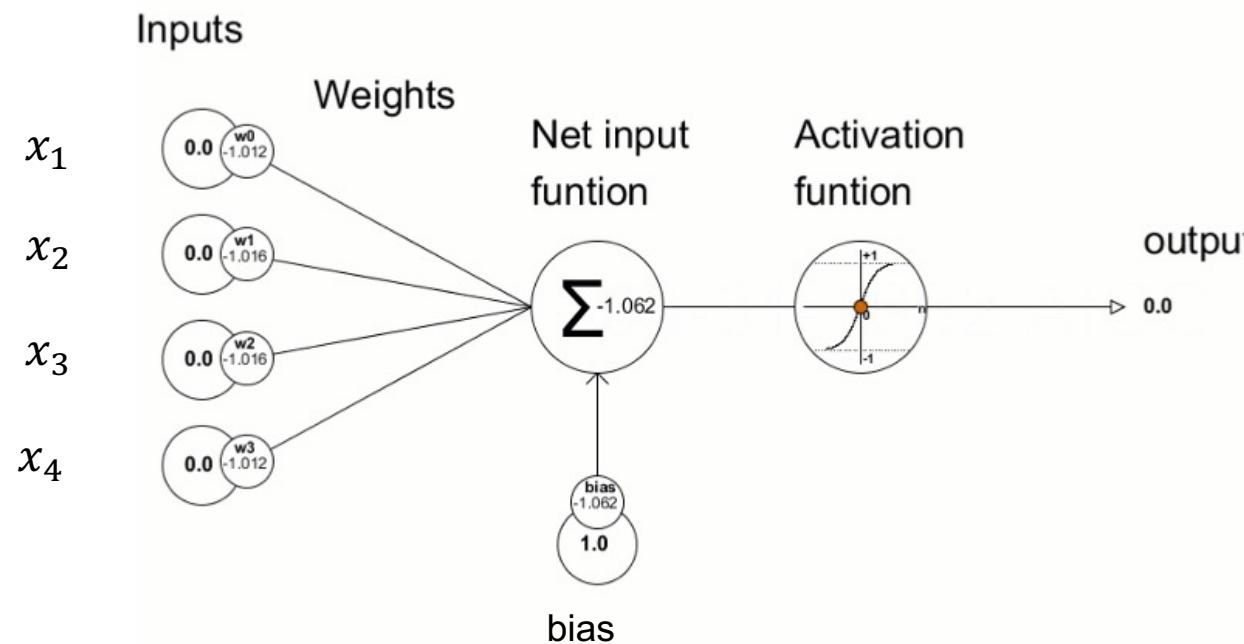
- Alle soluzioni di IA manca un **modello del mondo**, trasversale al compito specifico per il quale il sistema viene utilizzato, costruito incrementalmente
- Modello sul quale basarsi per imparare ogni nuova abilità, e da utilizzare per gestire situazioni mai osservate in precedenza, per formulare previsioni sulle conseguenze delle proprie azioni, ...



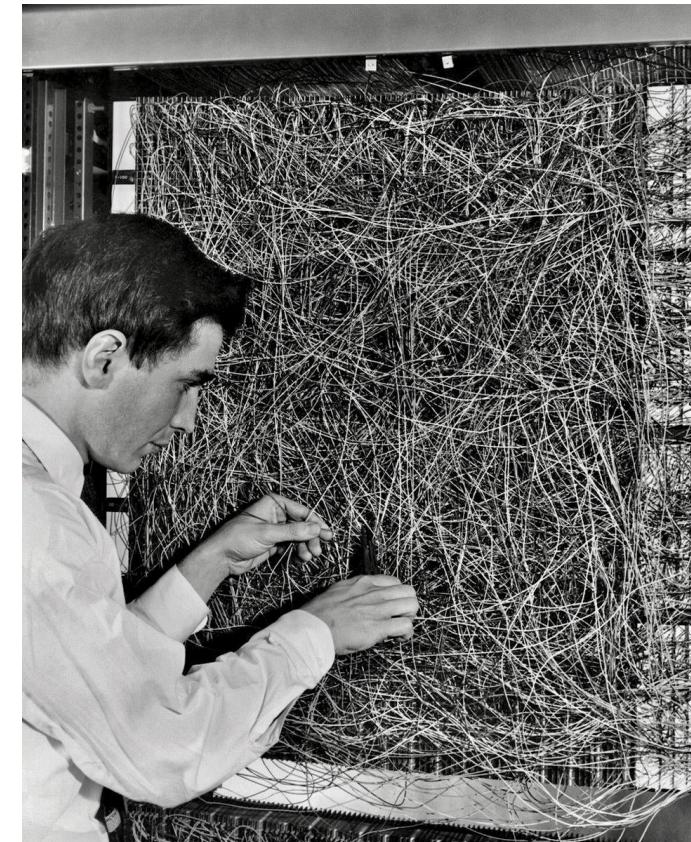
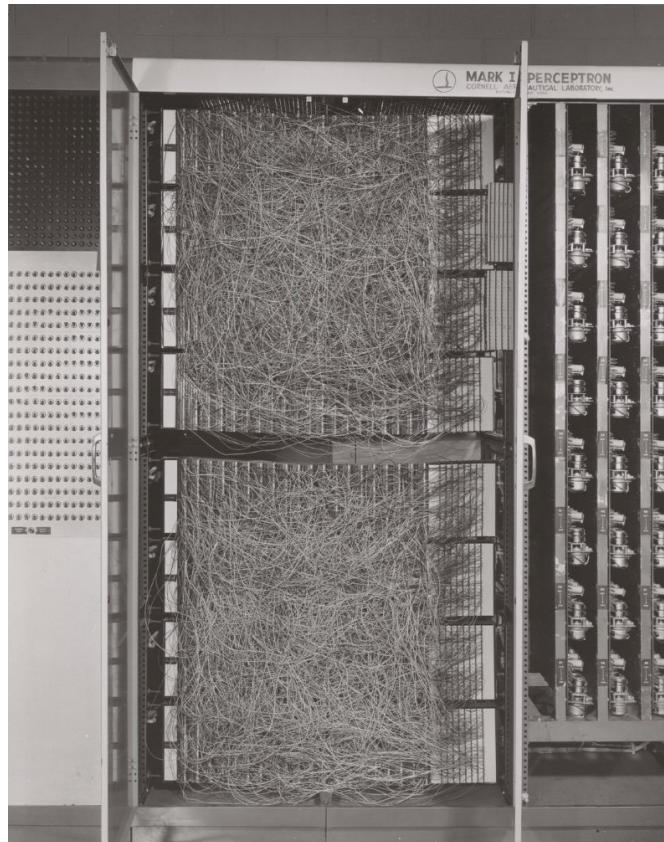
- Il cervello umano contiene all'incirca 86 miliardi di neuroni, collegati da  $10^{14}$ - $10^{15}$  sinapsi
- I dendriti ricevono i segnali dagli altri neuroni; la cellula elabora i segnali ricevuti e, se si verificano delle opportune condizioni, genera un segnale in output attraverso l'assone e le sinapsi



- Un neurone riceve in input un insieme di quantità numeriche:  
 $x_1, x_2, x_3, x_4$  nel nostro caso 1.0, 2.0, 2.1, 1.2
- Ciascuna quantità viene moltiplicata per un peso:  $w_1, w_2, w_3, w_4$
- Facciamo la somma di tali prodotti:  $x_1 * w_1 + x_2 * w_2 + x_3 * w_3 + x_4 * w_4$
- Il risultato della somma viene passato attraverso una «funzione di attivazione», che decide se e come attivare l'output di un neurone

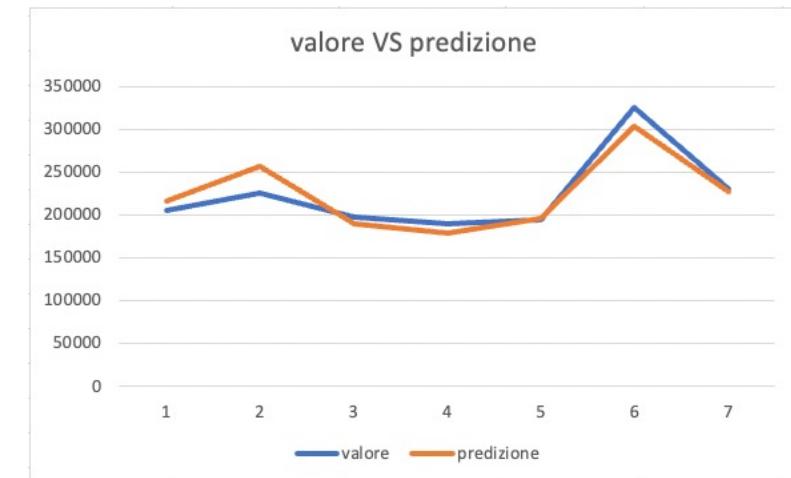
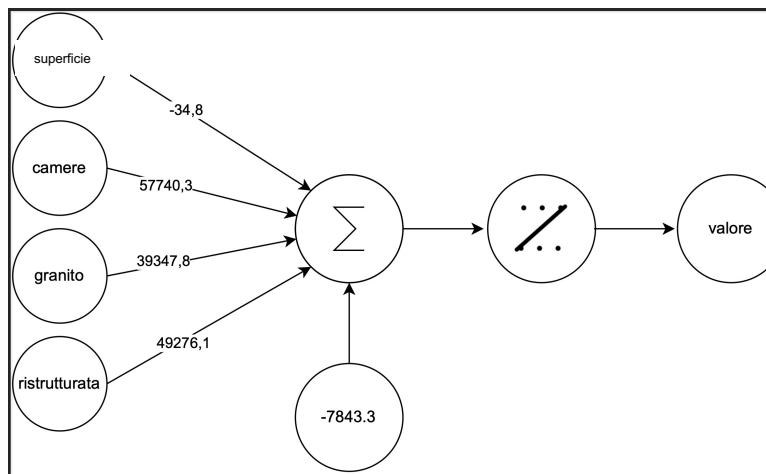


- Le prime implementazioni dei neuroni artificiali, negli anni '50 e '60, sono hardware
- Collegamenti costruiti tramite **cavi**, numeri codificati attraverso un **segnale elettrico**, pesi implementati con **resistenze e potenziometri**
- Oggi, le reti neurali vengono simulate all'interno del calcolatore



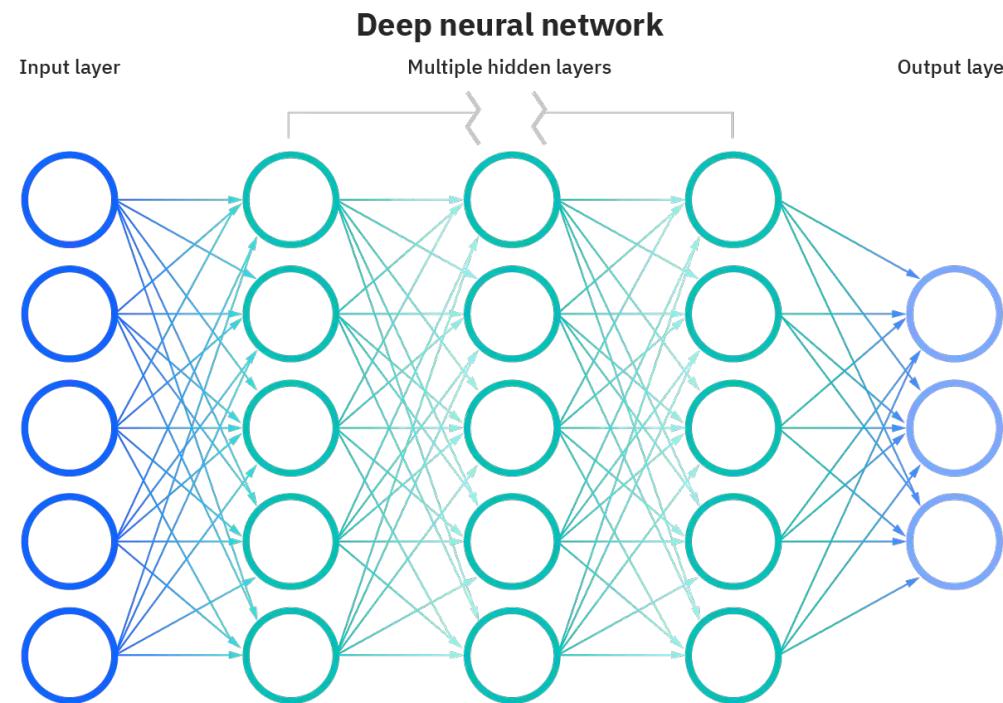
- Alla rete neurale vengono mostrati una serie di esempi relativi al concetto che si vuole apprendere
- Ciascun esempio è costituito da degli input, associati con un output
- Si vogliono imparare dei pesi, nella rete neurale, che consentano di costruire una mappatura fra input e output
- La rete neurale apprende dunque una **funzione matematica** fra input e output numerici

	<b>superficie</b>	<b>camere</b>	<b>granito</b>	<b>ristrutturata</b>	<b>valore</b>
1	3529	6	0	0	205000
2	3247	5	1	1	224900
3	4032	5	0	1	197900
4	2397	4	1	0	189900
5	2200	4	0	1	195000
6	3536	6	1	1	325000
7	2983	5	0	1	230000



$$\begin{aligned}
 \text{valore} = & - 34.8 * \text{superficie} \\
 & + 57740.3 * \text{camere} \\
 & + 39347.8 * \text{granito} \\
 & + 49276.1 * \text{ristrutturata} \\
 & - 7843.3
 \end{aligned}$$

- Ci sono dei limiti a quello che è possibile fare con un singolo neurone
- Più neuroni, organizzati su più livelli (deep neural network)
- L'architettura è stabilita a priori, solo i pesi vengono appresi



[Esempi interattivi](#)

[Neural network zoo](#)

- [Video: Prof. Nick Bostrom - Artificial Intelligence Will be The Greatest Revolution in History, 2017](#)
- [Video: Creating AI Could Be the Biggest & Last Event in Human History | Stephen Hawking |Google Zeitgeist, 2015](#)
- [Video: Storia e Futuro dell'Intelligenza Artificiale, - con Chris Bishop, 2017](#)
- [Video: Consciousness in Artificial Intelligence | John Searle | Talks at Google, 2015](#)
- [A Brief History of AI: How to Prevent Another Winter \(A Critical Review\), 2021](#)
- [The History of Artificial Intelligence, University of Washington, 2006](#)