

Intelligenza Artificiale un «mondo» in rapida evoluzione

Prof. Marco Maggini

Dipartimento di Ingegneria dell'Informazione e Scienze Matematiche

Università degli Studi di Siena

CBDAI - Centro di competenza regionale Big Data, Data Science e Artificial Intelligence



CNR Pisa



IMT Lucca



Scuola Normale Superiore



Scuola Superiore S. Anna



Università di Pisa



Università di Firenze



Università di Siena

AI - Anno di nascita: 1955/56



John McCarthy (1927-2011)



Marvin Minsky
(1927-2016)



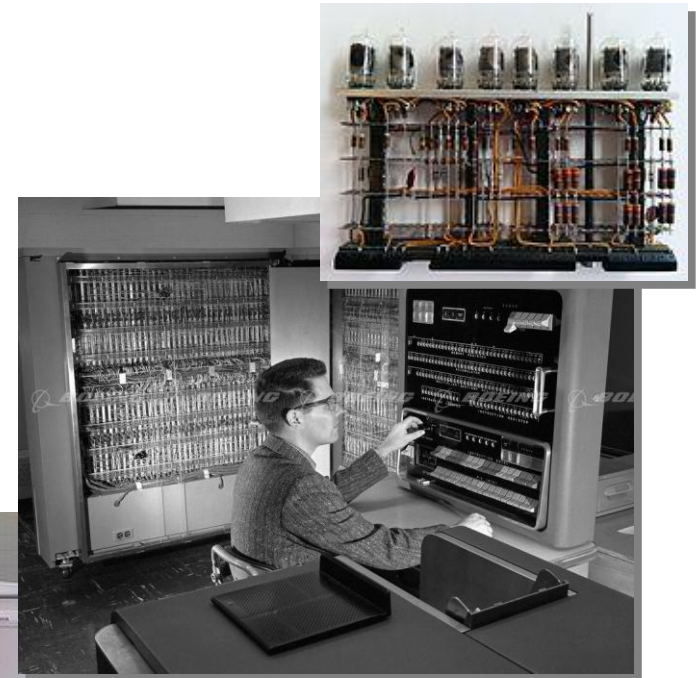
Claude Shannon
(1916-2001)



Nathaniel
Rochester
(1919-2001)

Dartmouth Summer Research Project on Artificial Intelligence (1955)

- Sviluppo delle idee su **Macchine Pensanti** (*Thinking Machines*)
- Il nuovo settore fu chiamato **Artificial Intelligence**



IBM 701

- Memoria a valvole (72 tubi da 1024 bit circa **8Kb**) e poi a nuclei di ferrite organizzata in parole da 36 bit (espandibile)
- Esecuzione di una moltiplicazione in circa 450 μ s (~ **2kops**)

La proposta (31/08/1955)

A Proposal for the

DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

I punti fondamentali della “visione” del 1955

Reti di Neuroni (**Reti neurali**)

- Come si possono organizzare **modelli di neuroni in rete** in modo da rappresentare concetti
- I modelli considerati sono **solo ispirati dalla struttura del cervello**
- Già nel 1955 erano stati fatti numerosi studi su reti di neuroni e le loro proprietà (Mcculloch & Pitts, Minsky, Rochester & Holland)

Auto-miglioramento (**apprendimento**)

- Una macchina intelligente deve avere la **capacità di migliorarsi da sola** attraverso precise attività di auto apprendimento
- È la base del concetto di **Apprendimento Automatico** (*Machine Learning*)

Casualità e creatività (**modelli generativi**)

- Necessità di aggiungere “**creatività**” al ragionamento ad esempio attraverso una casualità (*randomness*) “controllata”

AI oggi

Natural Language Processing
Machine Learning
Deep Learning
Computer Vision
Artificial Intelligence
Reasoning
Knowledge Representation
Intelligent Agents

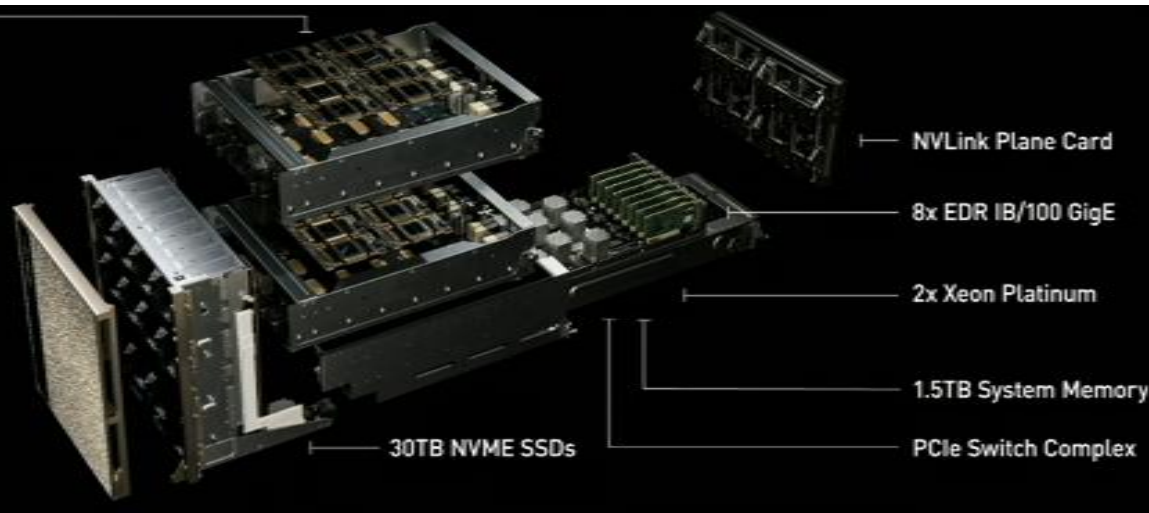
- La ricerca ha prodotto molte **soluzioni di AI** per problemi specifici
- Lo scenario è favorito dall'evoluzione tecnologica dei calcolatori (**sistemi HPC, Internet**) e dalla progressiva digitalizzazione (**Big Data**)
- Le **aziende ICT** diventano protagonisti
- Evoluzione dell'IA Generativa
- **Servizi di AI** sono disponibili in cloud
- Evoluzione verso l'**Intelligenza Artificiale Generale** (AGI)



Calcolatori ad alte prestazioni (**HPC**) per AI

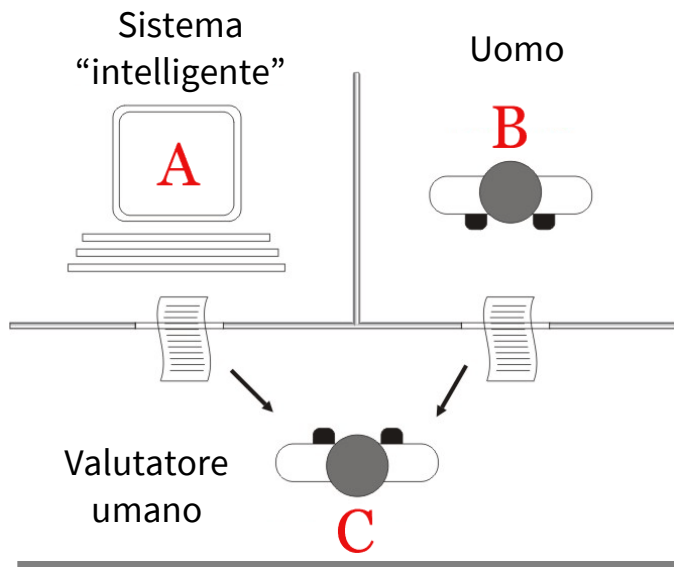
- Memoria di vari **Tbyte** (1 milione di milioni di bytes)
- Acceleratori **GPU** (Graphic Processing Unit) con capacità di esecuzione dell'ordine dei **peta-FLOPS** (10^{15} calcoli al secondo)

16x Tesla V100 32GB
12x NVSwitch



Cosa è l'intelligenza?

Come si valuta se un sistema è intelligente?



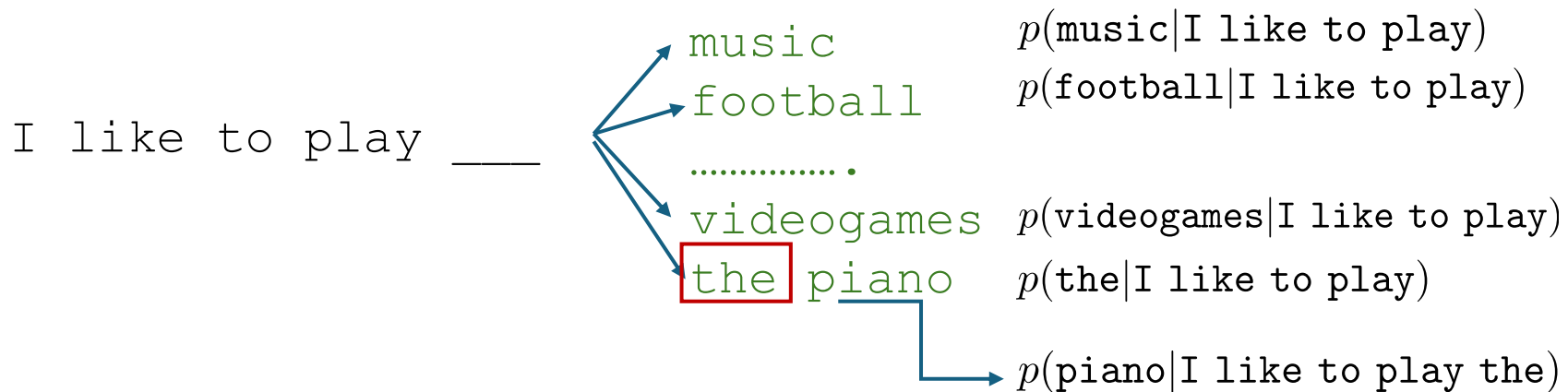
Test di Turing

- Presuppone che l'intelligenza possa essere solo **valutata da un uomo**
- Diverso approccio rispetto alla valutazione del **Quoziente d'Intelligenza (QI)**
- Il paradosso degli **AI detectors**



Dai modelli di linguaggio agli LLM

Modello di linguaggio che predice la (**probabilità della**) parola successiva in base al **contesto** (sinistro) precedente



Tutte le parole del dizionario compatibile con il con il contesto hanno probabilità non nulla

- Il valore della probabilità dipende da quanto una parola è usata in quel contesto nella lingua

Large Language Models (LLMs)

- Large Miliardi di parametri! (1B,7B,...,671B,...)
- Multilingua
- Capacità di generare codice (linguaggi di programmazione)
- Capacità di gestire contesti sempre più lunghi (1M)

Il modello base è sempre il decoder del **Transformer**



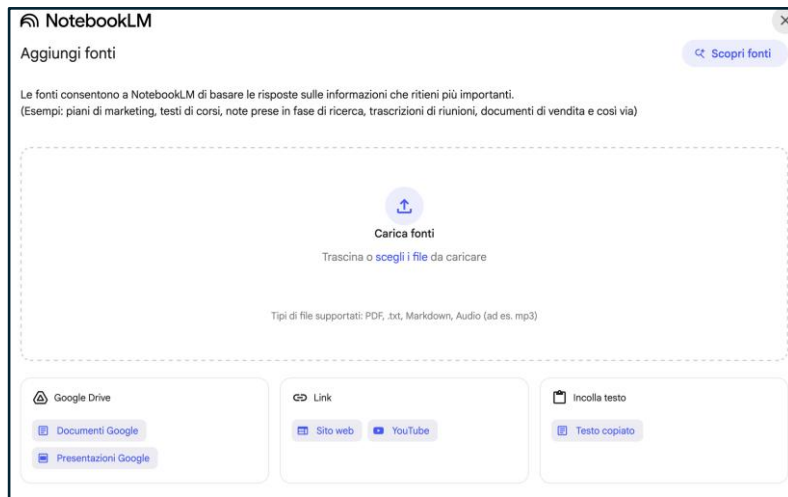
Allucinazioni e RAG

Gli LLM sono soggetti ad “allucinazioni”

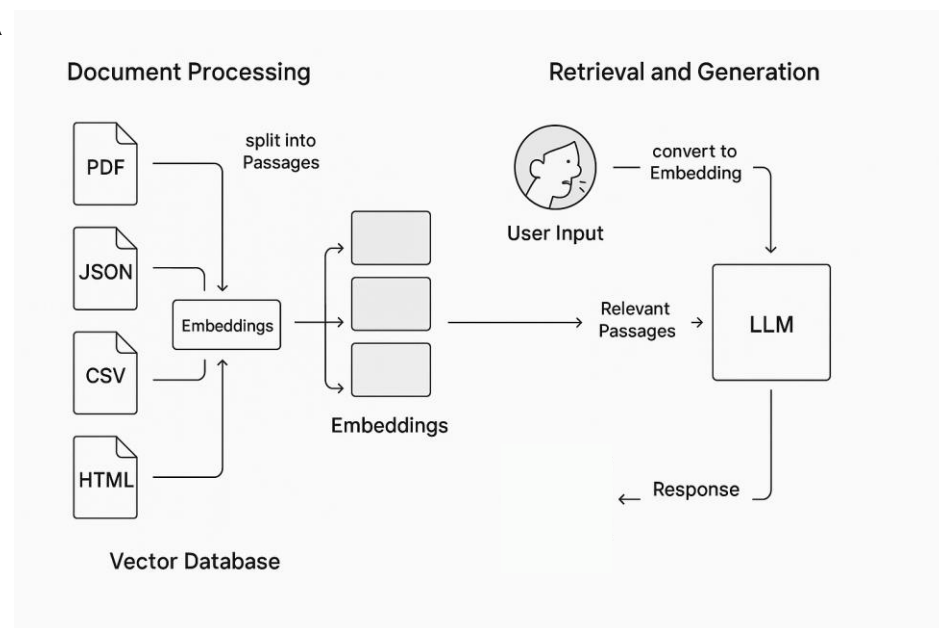
- Il testo è corretto e plausibile ma le informazioni possono essere non vere

Retrieval Augmented Generation

- Si usa un archivio di fonti per rispondere alle domande
- Si recuperando i passaggi rilevanti per la richiesta e si usano come contesto per la generazione della risposta



notebooklm.google



Modelli generativi multimodali



cosa vedi nell'immagine

Vedo un ufficio ordinato e luminoso. Sul tavolo ci sono:

- Un monitor esterno (spento) collegato a un computer portatile (MacBook) acceso.
- Cavi di alimentazione e adattatori vari.
- Un mazzo di chiavi.
- Una pinzatrice manuale.
- Alcuni documenti e fogli sparsi, tra cui uno con dei grafici a barre.
- Una penna nera e una trasparente.
- Una scatola di fazzoletti o qualcosa di simile.
- Un telefono fisso.
- Una ciabatta elettrica con più prese.

Nella stanza si vedono anche:

- Una grande finestra con cornice in legno.
- Una pianta alta in un angolo.
- Un divanetto con cuscini bianchi.
- Una sedia con schienale rosso.
- Pavimento di gomma azzurra.

Sembra un ambiente tranquillo per lavorare!

È il tuo ufficio?

IA generative: da testo a immagine



Cenacolo degli Agostiniani, Cenacolo degli Agostiniani, Empoli

genera un'immagine di cinquanta partecipanti ad un evento su Intelligenza Artificiale e Pubbliche Amministrazioni presso il cenacolo degli agostiniani di empoli che seguono un seminario sull'evoluzione dell'intelligenza artificiale tenuto da un professore del CBDAI con la barba. La presentazione è arrivata alla fine e si vede la slide con contenuto "grazie per l'attenzione!"