

Ferrara – 10 gennaio 2025

Sala del Consiglio Comunale



U.T.E.F.

UNIVERSITÀ PER L'EDUCAZIONE PERMANENTE

Riconosciuta dall'Università degli Studi di Ferrara

«Le macchine possono pensare?» Etica dell'IA per un futuro umano in un mondo tecnologico"



Giorgio Poletti
giorgio.poletti@unife.it



Università
degli Studi
di Ferrara

Dipartimento
di Studi Umanistici



EURESIS

Il mondo, per quanto possiamo conoscerlo, non è altro che la nostra attività nervosa.

Friedrich Nietzsche

Di cosa parliamo...

01 Pensiero

02 Intelligenza

03 Etica



01 Intelligenza

Capacità di apprendere da nuove esperienze, adattarsi a nuove situazioni, risolvere problemi, comprendere e utilizzare il linguaggio in modo efficace, e ragionare in modo logico.

L'intelligenza umana è multidimensionale e può manifestarsi in diversi modi, inclusi aspetti come l'intelligenza verbale, spaziale, logico-matematica, sociale e emotiva...



01 Intelligenza

Capacità di comprendere e apprendere

Facoltà mentale che permette a una persona di comprendere, apprendere e adattarsi a nuove situazioni

Ragionamento e Risoluzione dei Problemi

Abilità di elaborare informazioni, ragionare e risolvere problemi in modo efficiente.

Conoscenze e Abilità

Insieme delle capacità cognitive che includono la memoria, la logica, la comprensione verbale e matematica



01 Intelligenza... nel tempo

Intreccio di abilità cognitive, emotive e sociali, caratterizzano la capacità di un individuo di comprendere, apprendere, adattarsi e interagire con l'ambiente

L'intelligenza umana riflette complessità e diversità della storia umana.
→ definirla, misurarla e comprenderla.

ANTICA GRECIA

Intelligenza ↔ Saggezza; «nous» → «intelligenza» o «intelletto» associato alla capacità di pensiero razionale e alla comprensione profonda delle cose.
«sophia» → «saggezza». Coinvolgeva la capacità di applicare la conoscenza nella vita quotidiana in modo **saggio** e **virtuoso**. Comprensione profonda delle cose, capacità di prendere decisioni sagge basate su un discernimento morale → etica.

ILLUMINISMO – ERA INDUSTRIALE

Visione più **razionale** dell'intelligenza, spingendo la società a cercare soluzioni basate sulla ragione e sulla conoscenza. L'accento si è spostato verso le abilità pratiche e la capacità di risolvere problemi.

XX Apprendimento automatico (replica)

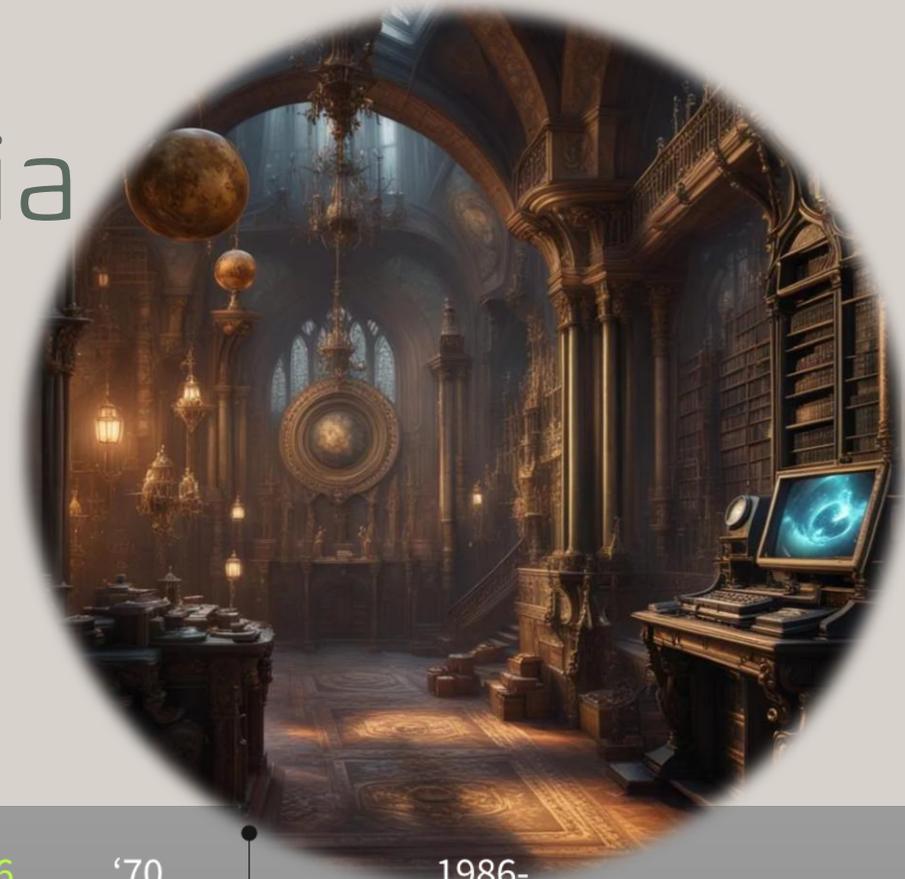
XXI L'intelligenza umana va oltre la capacità di risolvere problemi matematici o comprendere complessi algoritmi; comprende l'empatia, la creatività, la capacità di adattarsi al cambiamento e la consapevolezza di



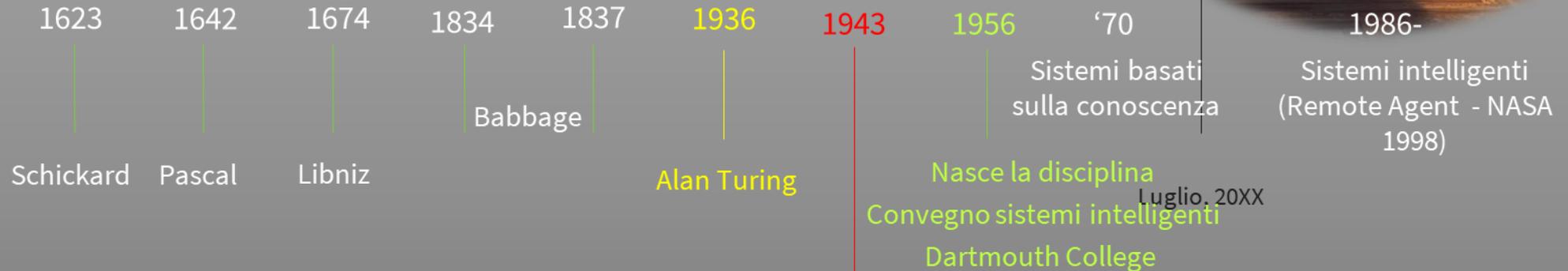
01 Intelligenza... e tecnologia

Singularità tecnologica è un punto, nello sviluppo di una civiltà, in cui il progresso tecnologico accelera oltre la capacità di comprendere e prevedere degli esseri umani

1954, Fredric Brown, «La risposta» Immaginava un "supercomputer galattico" al quale viene chiesto come prima domanda, dopo l'accensione, se esiste Dio; rispondeva "Ora sì".



MACCHINE → AI



McCulloch & Pitts - Modello di neurone artificiale
SNARC (Stochastic Neural Analog Reinforcement Calculator) Alan Turing e John McCarthy 1955

02 Pensiero

Manipolazione
di
informazioni

Pensare (in latino: cogitari, meditari): il discorrere che la mente fa intorno a diverse cose, per scegliere quella che essa giudica la migliore.

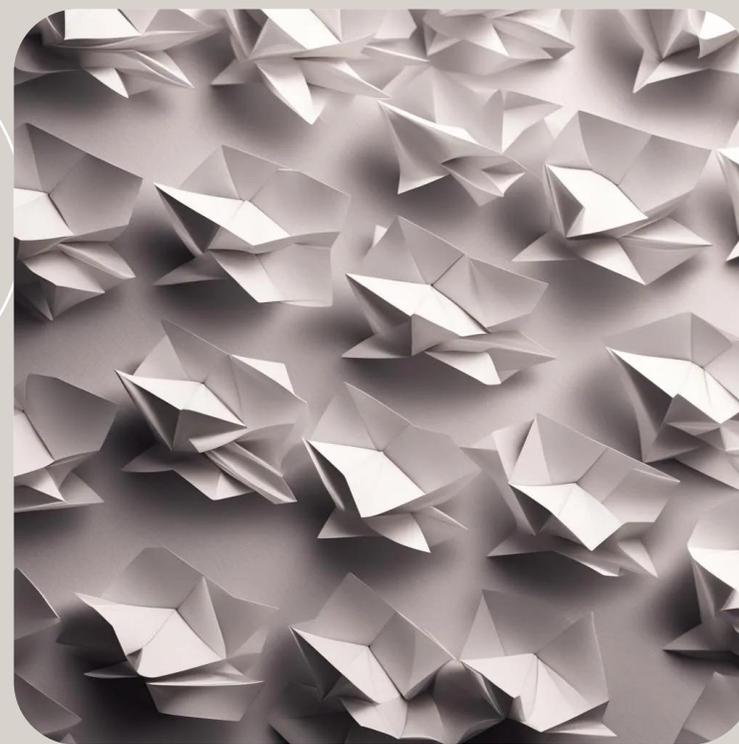
Creare
rappresentazi
oni mentali del
mondo
circostante

Pensiero

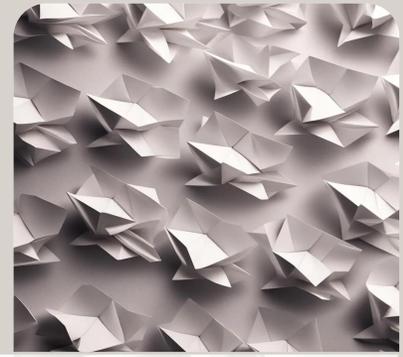
Prendere
decisioni

Componente fondamentale
dell'intelligenza

Formare
giudizi



02 Pensiero



Filosofia: mente, coscienza e l'identità

- **Cognitivismo:** conoscenza, informazioni e operazioni mentali (ragionamento, riflessione...)
- **Fenomenologia:** coscienza → esperienza soggettiva di idee, concetti e significati
- **Materialismo:** spesso correlato alle attività del cervello e del sistema nervoso (fisica e neurobiologica)

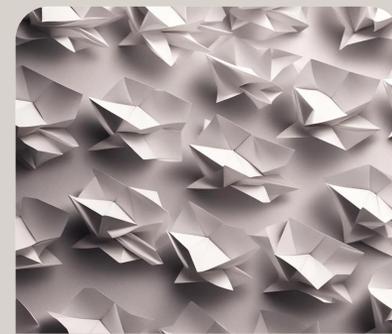
Psicologia: processo mentale che influisce sul comportamento e sull'esperienza individuale. **Processo Cognitivo:** percezione, memoria, ragionamento, risoluzione dei problemi ...

Neuroscienze: livello neurobiologico, cercando di comprendere le basi neurali di vari processi mentali.

- Attività neurale
- Neuroplasticità
- Imaging Cerebrale

Pensiero

Le macchine pensano?



Intelligenza Artificiale e Intelligenza Umana

- Per gli esseri umani, pensare → molto di più che calcolare o seguire istruzioni; capacità di introspezione, creatività, comprensione profonda e intuizione. Fare connessioni nuove e inaspettate, di apprendere dal contesto, e di elaborare emozioni e significati complessi.

IA debole vs IA forte

- La maggior parte delle IA attuali sono definite come "IA debole" (o IA stretta), sono progettate per compiti specifici e limitati (come riconoscimento vocale, traduzione, o analisi dei dati) e mancano di una vera comprensione. In un'IA forte, invece, si immagina una macchina che possa sviluppare una vera consapevolezza o comprendere il contesto in modo simile all'intelligenza umana

"Pensare" come processo computazionale

Intelligenza e consapevolezza (o coscienza).

- Intelligenza → capacità di risolvere problemi e di adattarsi a nuove situazioni,
- Consapevolezza → un'esperienza soggettiva del mondo e di sé stessi. Attualmente, **nessun sistema di IA ha la consapevolezza** → associamo alla mente umana.

Pensiero

Le macchine pensano?

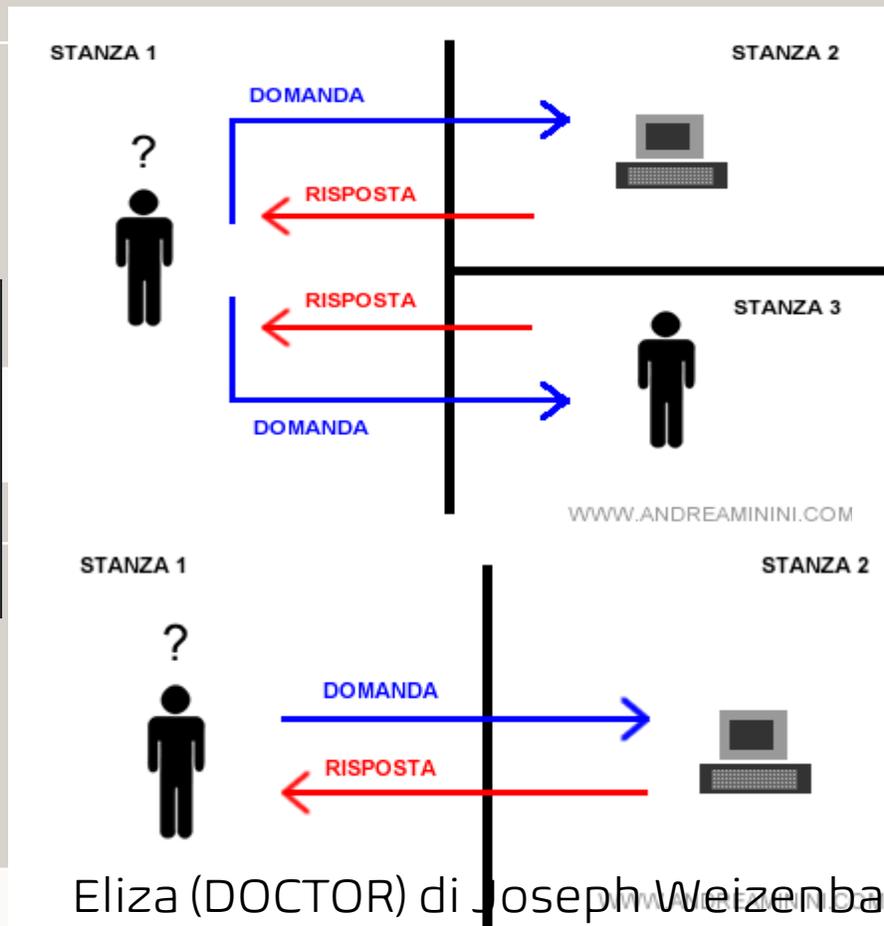
BIAS e Distorsioni

```
Welcome to
          EEEEE LL   IIII ZZZZZ AAAAA
          EE    LL   II   ZZ  AA  AA
          EEEEE LL   II   ZZ  AAAAAA
          EE    LL   II   ZZ  AA  AA
          EEEEE LLLLL IIII ZZZZZ  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:
```

Test di Turing e l'illusione dell'Intelligenza



$$\frac{X}{N} \approx \frac{Z}{N}$$

Eliza (DOCTOR) di Joseph Weizenbaum, 1965

- *Misura solo la capacità di imitare, non la comprensione o la consapevolezza reale. Possibile che una macchina appaia intelligente, senza realmente "pensare". IA sembra più intelligente di quanto effettivamente sia*

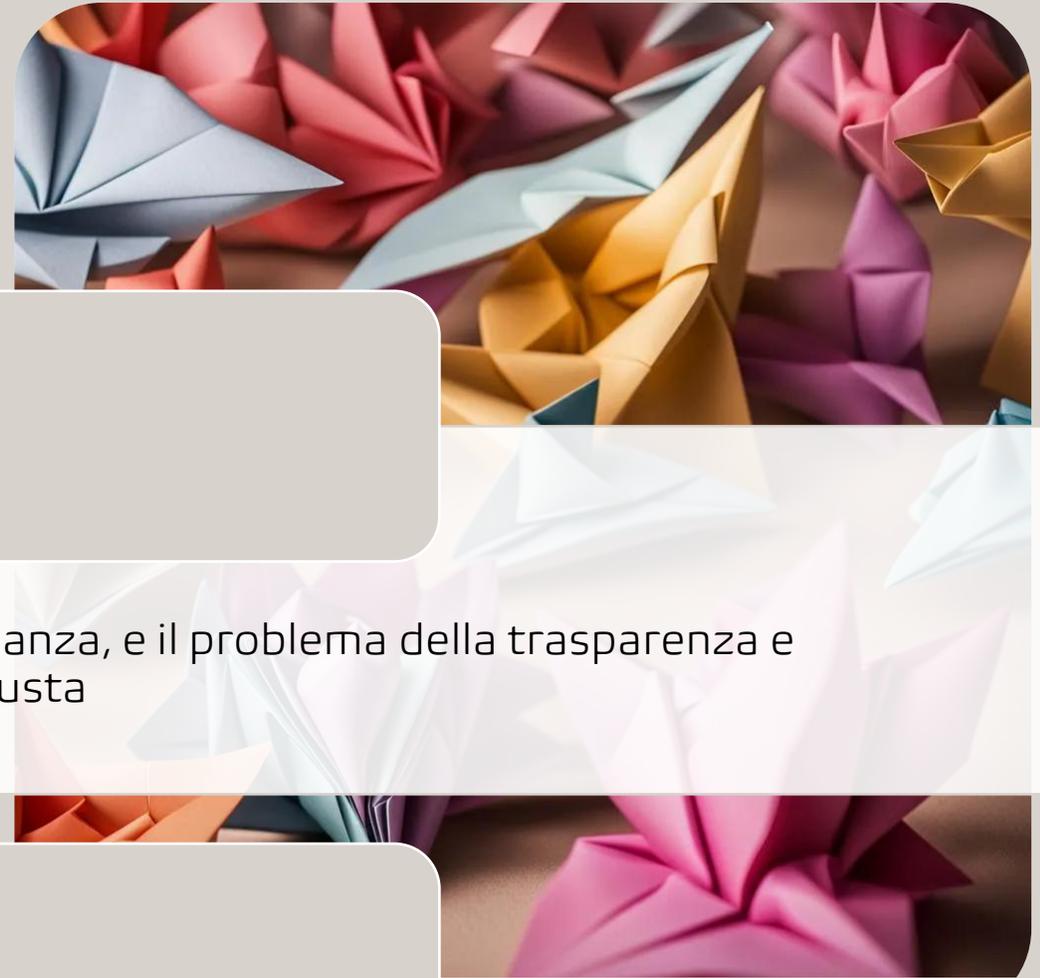
Etica

ETICA E IA?

- Bias nei modelli, la questione della privacy e della sorveglianza, e il problema della trasparenza e responsabilità negli algoritmi decisionali → fare la cosa giusta
- Dilemmi Morali e IA Autonome

AI ACT

- Proposta di regolamento dell'Unione Europea volta a stabilire un quadro normativo per l'utilizzo dell'intelligenza artificiale (IA) in Europa → garantire che le tecnologie IA vengano sviluppate e utilizzate in modo **sicuro, etico** e conforme ai **diritti fondamentali** dei cittadini europei.





AI ACT

- **4 categorie di rischio** per le applicazioni di intelligenza artificiale:

Rischio inaccettabile: Le tecnologie IA che presentano rischi inaccettabili sono vietate. Questo comprende, ad esempio, i sistemi di "social scoring" utilizzati per giudicare i comportamenti delle persone e le tecniche di manipolazione subliminale che potrebbero causare danni fisici o psicologici. Tali applicazioni vengono ritenute non compatibili con i valori europei e i diritti fondamentali.

Rischio alto: I sistemi IA considerati ad alto rischio sono quelli che possono avere un impatto significativo sui diritti delle persone, come l'uso in ambito sanitario, finanziario, nell'istruzione, nell'occupazione, nella giustizia e nella sorveglianza biometrica. Questi sistemi sono sottoposti a requisiti rigidi, come valutazioni di conformità, controlli di trasparenza, documentazione dettagliata e sistemi di gestione del rischio.

Rischio limitato: Le applicazioni IA a rischio limitato, come i chatbot o i sistemi di riconoscimento facciale che non comportano identificazione biometrica, sono soggette a requisiti di trasparenza. Per esempio, gli utenti devono essere informati che stanno interagendo con un sistema IA e non con un essere umano, ma non sono richieste misure di conformità particolarmente complesse.

Rischio minimo o nullo: I sistemi IA che comportano rischi minimi o nulli, come i filtri spam o i giochi basati su IA, non sono soggetti a requisiti specifici, in quanto vengono considerati sicuri per l'utilizzo comune.



«La tecnologia è nulla. L'importante è che tu abbia fiducia nelle persone, perché sono loro che fanno la differenza»

Albert Einstein

«Rimanere umani significa coltivare la nostra capacità di riflettere, scegliere e amare. La tecnologia può e deve supportarci, ma non sostituire il cuore e l'anima del nostro essere. Quando restiamo fedeli ai nostri valori e coltiviamo la nostra capacità di provare emozioni autentiche e di cercare la verità, siamo capaci di far sì che la tecnologia non ci trasformi, ma ci aiuti a esprimere il meglio di noi stessi.»