

Intelligenza artificiale: rischi e opportunità

Angelo Montanari

Università degli Studi di Udine

«L'IA è probabilmente la cosa migliore o la cosa peggiore

mai accaduta all'umanità» Stephen Hawking

Pordenone, 16 febbraio, 2026

Reale o
artificiale?

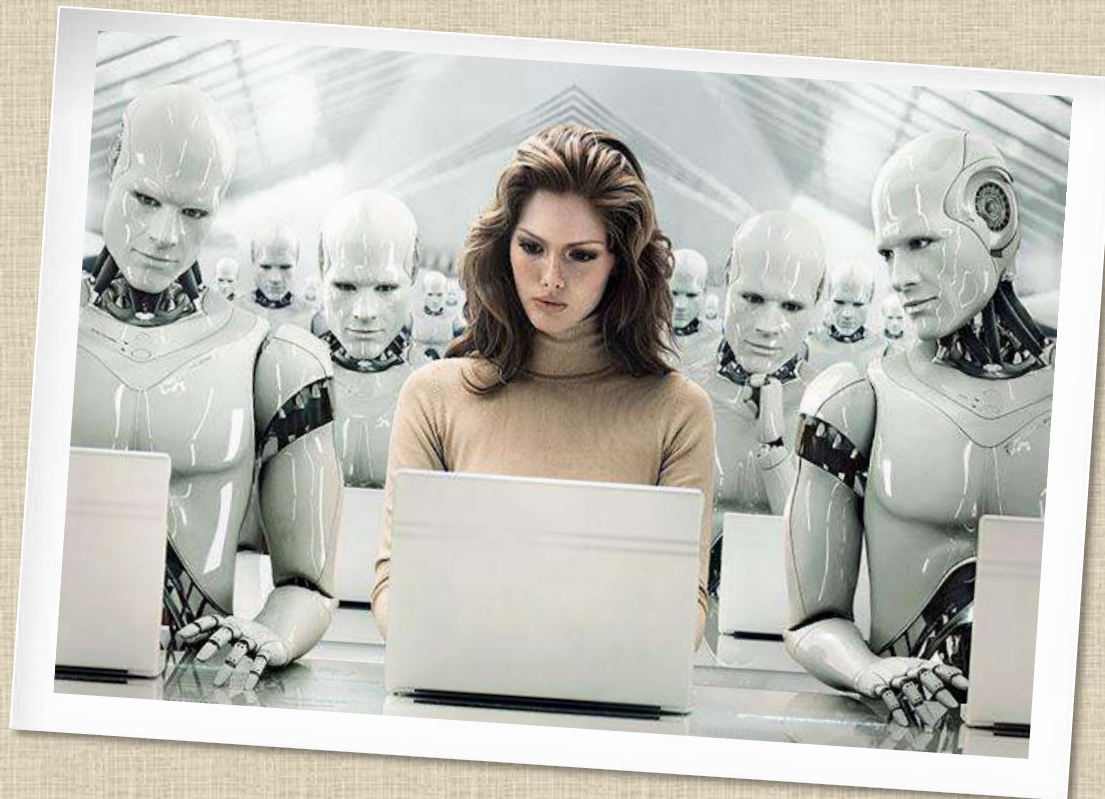


L'Intelligenza Artificiale (IA) è reale

I confini tra reale e artificiale diventano sempre più labili, ma **l'impatto dell'IA** sulla vita delle persone e delle società è un'evidenza incontrovertibile.

Non vi è ambito in cui l'IA non abbia introdotto profondi cambiamenti.

E' possibile stendere un elenco delle **applicazioni dell'IA** che copre tutte le lettere dell'alfabeto (automotive, bioinformatica, chatbot, domotica, ..)



IA e lavoro

Alcuni anni fa, l'economista capo della Banca d'Inghilterra aveva previsto che l'IA avrebbe **distrutto il 50% dei posti di lavoro nel Regno Unito.**

Un report del World Economic Forum del 2020 ([The Future of Jobs Report 2020](#)) prevedeva entro il 2025 la sostituzione di circa **85 milioni di posti di lavoro con l'IA** (e la creazione di altri 97 milioni).

IA e finanza

In presenza di situazioni complesse e anomale, le decisioni prese dai sistemi di IA, in modo estremamente rapido e in assoluta autonomia, possono rafforzare le tendenze negative.



Questa foto
BY-NC-ND

CC

IA e società

L'IA consente di acquisire informazioni su singoli individui e gruppi di individui per analizzarne le caratteristiche e dividerli in categorie (**profilazione**) e poter effettuare valutazioni o previsioni.

Forme di profilazione sempre più raffinate e non facilmente monitorabili (dal tracciamento delle interazioni con la rete ai sistemi di riconoscimento facciale e agli assistenti vocali) rendono le persone sempre più esposte al rischio di condizionamenti e manipolazioni.



Questa foto
BY-NC-ND

CC



Questa foto
BY-SA-NC

CC

IA e guerra

Più di 10 anni fa, migliaia di ricercatori firmarono una lettera aperta in cui prevedevano che l'IA avrebbe trasformato la guerra e dato inizio a una corsa agli armamenti caratterizzata da **robot killer**.

Il futuro è adesso (la guerra in Ucraina e non solo).

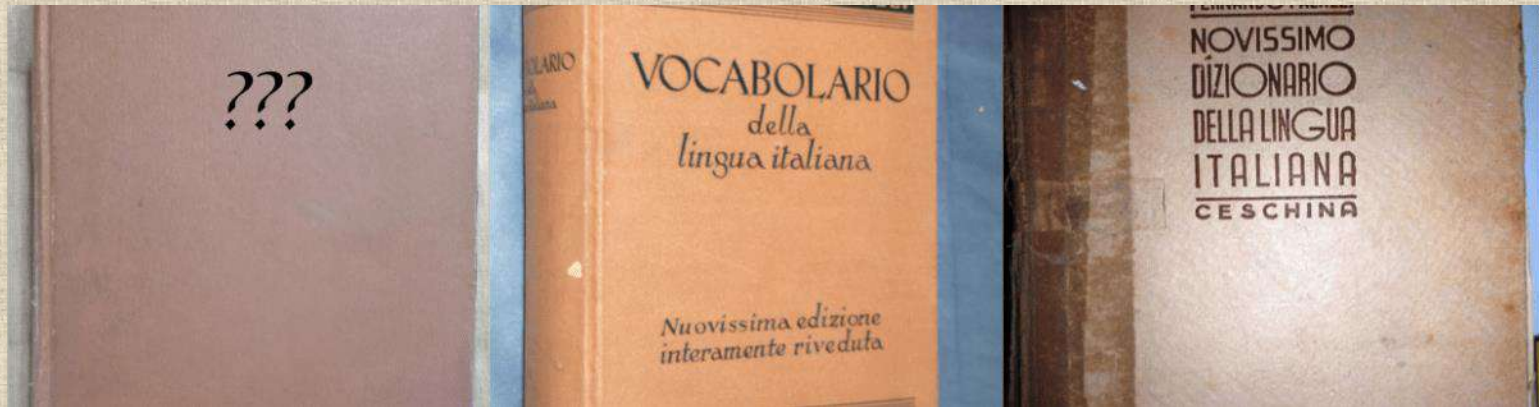
L'intelligenza Artificiale Generale

- L'Intelligenza Artificiale Generale (**Artificial General Intelligence**, abbreviata **AGI**) ha l'obiettivo di costruire dei sistemi artificiali in grado di eseguire ogni compito di natura intellettuale che l'essere umano è in grado di eseguire.
- Non sistemi pensati per realizzare una specifica attività di natura intellettuale, anche molto avanzata, ma strumenti versatili, in grado di apprendere e ragionare, capaci di affrontare e, per quanto possibile, risolvere problemi di qualsiasi natura, al pari degli esseri umani.

Un breve viaggio nell'IA

- Che cos'è l'IA?
- Ragionamento (la logica) e memoria (i dati)
- Intelligenza artificiale e esseri umani
- Considerazioni finali

Una questione preliminare



Questa foto di Autore sconosciuto è concesso in licenza da [CC BY-NC-ND](#)

Uso di un **vocabolario antropomorfo** nella descrizione delle caratteristiche e del funzionamento dei sistemi informatici

- è particolarmente evidente nel caso dei sistemi di IA (**intelligenza**, conoscenza, apprendimento, ragionamento),
- ma si è verificato in misura più o meno rilevante in molti altri ambiti dell'informatica (**memoria**, comunicazione, interrogazione).

Ragioni ..

Ragioni dell'uso di un tale vocabolario:

- può creare una certa **familiarità** (un rapporto positivo) con la macchina;
- può aiutare a comprendere alcune caratteristiche strutturali o funzionali delle macchine “per **analogia**” con le corrispondenti caratteristiche degli esseri umani (e/o degli animali);
- In diverse aree disciplinari, esseri umani e animali sono stati presi programmaticamente quali **modelli** di riferimento per la costruzione di sistemi artificiali.

N. Wiener: Cybernetics or Control and Communication in the Animal and the Machine, MIT Press, 1962).



.. e rischi



L'uso di un vocabolario antropomorfico presenta un grande rischio:

assumere che struttura e funzionamento dei sistemi di IA **riproducano** esattamente struttura e funzionamento degli esseri umani e/o degli animali.

E' questa la posizione della cosiddetta **IA forte** criticata da Searle in un famoso articolo di più di 40 anni fa e in numerosi successivi interventi.

Che cos'è l'Intelligenza Artificiale?



L'IA è un settore dell'**informatica** (già presente negli scritti di Alan Turing, il padre riconosciuto dell'informatica).

Lo specifico dell'informatica: la capacità di **modellare** e di **risolvere problemi**, dove la prima abilità è almeno tanto importante quanto la seconda.

Concetto chiave: **algoritmo**



Questa foto
BY-ND

CC

L'algoritmo

Un algoritmo è una descrizione finita e non ambigua di una sequenza di operazioni che consente ad un dato agente (un uomo o una macchina) di risolvere un determinato problema.

Esempi molto famigliari di algoritmi: le ricette.

I sistemi di intelligenza artificiale

Da sempre le macchine fanno delle cose che l'uomo non è in grado di fare (pensiamo alla rivoluzione industriale e ai mezzi di trasporto).

La novità è che i sistemi di IA sono in grado di sostituire l'essere umano e, spesso, di migliorarne le prestazioni in compiti ritenuti da sempre di sua competenza esclusiva.

Quando i sistemi artificiali toccano la sfera dell'intelligenza diventiamo tutti molto più sensibili..



IA: ragionamento e memoria

L'intelligenza naturale si basa sulle capacità di ragionamento e di memorizzazione degli esseri umani (ciò vale, in una certa misura, anche per gli animali).

A **ragionamento** e **memoria** possono essere ricondotti i due filoni di ricerca da sempre presenti in IA:

- IA guidata dai modelli (IA simbolica);
- IA guidata dai dati (IA sub-simbolica).

IA simbolica e IA sub-simbolica

IA simbolica: rappresentazione della conoscenza e ragionamento automatico.

Un caso paradigmatico: la pianificazione automatica (di robot, satelliti, ..).

IA sub-simbolica: mimare il comportamento del cervello umano e la sua complessa rete di neuroni interconnessi.

I primi modelli di rete neurale risalgono agli anni '40 (il modello dei neuroni di McCulloch e Pitts, 1943).

Il ragionamento automatico

Uso della logica (matematica) per la rappresentazione della conoscenza e il ragionamento automatico.

Teoria.

La dimostrazione automatica dei teoremi e la verifica automatica delle dimostrazioni.

Applicazioni.

La pianificazione automatica, la robotica collaborativa, la gestione dei processi aziendali.

La robotica collaborativa

I robot collaborativi (**cobot**) sono dei robot industriali (ma vi sono applicazioni anche nell'ambito dell'assistenza domiciliare e della domotica) di nuova generazione progettati per lavorare assieme all'essere umano, fianco a fianco e in sicurezza, senza la necessità di barriere o gabbie protettive che separino gli uni dagli altri.



La memoria

La memoria è una componente essenziale dell'intelligenza (umana).

La dimensione temporale dell'intelligenza: memoria e **temporalità**.

Il passato: i dati.

Il futuro: le previsioni e i piani.

L'apprendimento (automatico) sulla base dei dati disponibili: non può esserci apprendimento senza dati.

La memoria e la rete

Dalla memoria del calcolatore alla rete: un'enorme memoria distribuita e condivisa (i meriti di Internet).

Una **memoria senza soggetto** (una memoria impersonale).

Quali sono le ragioni dell'**esplosione** della ricerca e delle applicazioni **del machine/deep learning** (IA guidata dai dati)? Enorme potenza di calcolo ed **enorme disponibilità di dati**.

Applicazioni. Tecniche di profilazione, upselling predittivo, manutenzione predittiva, agenti conversazionali (chatGPT).

La manutenzione predittiva



Monitoraggio in tempo reale del comportamento di un sistema per rilevare in anticipo potenziali guasti o malfunzionamenti e prevenire il fermo macchina

.. dall'ambito della produzione industriale a quello della salute umana

Ragionamento, memoria e intelligenza umana

L'intelligenza (umana) consiste unicamente in ragionamento e memoria?

Il test di Turing (o gioco dell'imitazione): una macchina può essere definita intelligente se riesce a convincere una persona che il suo comportamento, dal punto di vista intellettuale, non è diverso da quello di un essere umano medio.

Computing Machinery and Intelligence, A. M. Turing, *Mind*, New Series, Vol. 59(236), October 1950, pp. 433-460.



Intelligenza e linguaggio

Il Test di Turing assume uno stretto legame tra intelligenza e **linguaggio** (naturale): l'intelligenza si manifesta nell'interazione /comunicazione attraverso il linguaggio.

Il linguaggio occupa un ruolo fondamentale nel rudimentale **sistema esperto ELIZA** (psicanalista digitale) proposto da Weizenbaum negli anni '60.

La comprensione di testi in linguaggio naturale è al centro del famoso **esperimento mentale della stanza cinese** di Searle

All'ambito dell'elaborazione del linguaggio naturale appartiene anche l'universo dei sistemi conversazionali (Chatbot), dei quali **ChatGPT** è uno dei rappresentanti più noti.

Il sistema esperto ELIZA

```
EEEE  LL  II  ZZ  AAAAAA
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 ?

Amico chatbot, almeno tu ascoltami (una recente ricerca mostra come gli adolescenti, pur sapendo di parlare con una macchina, non ci rinunciano).

Intelligenza e corporeità

Il modello di intelligenza sotteso al Test di Turing è un modello astratto/disincarnato dell'intelligenza.

Una delle acquisizioni più importanti della ricerca in IA degli ultimi decenni è la consapevolezza del ruolo cruciale che gli organi di senso svolgono nell'interazione dell'uomo col mondo e della conseguente impossibilità di un'intelligenza (artificiale) priva di “corporeità”. Ciò ha portato allo sviluppo di un rapporto sempre più stretto tra IA ("cervello senza corpo") e robotica ("corpo senza cervello").

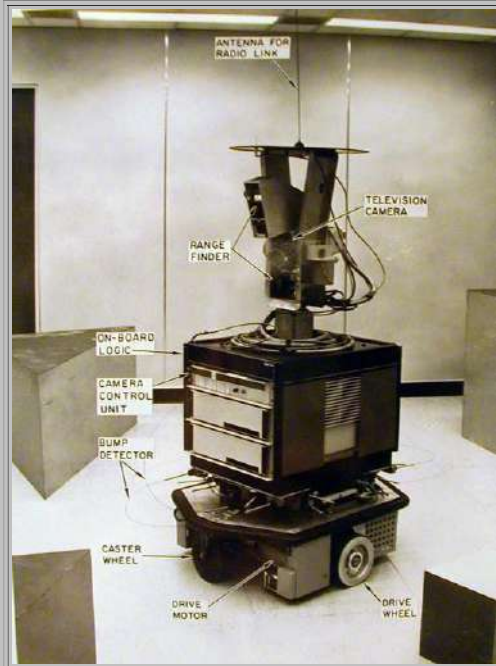
Per avvicinarsi all'intelligenza umana l'IA deve diventare un'intelligenza incarnata.

Anche gli strumenti di IA generativa sono privi di un corpo (che solo consente un'interazione diretta e non mediata col mondo).

Shakey ('60/'70), TMSUK 04 ('80/'90) e Justin ('09-)



Shakey the Robot - SRI (1968)



Intelligenza ed esseri umani

Linguaggio, corpo e .. coscienza.

Intelligenza e «umanità»: la coscienza è un tratto distintivo dell'intelligenza o dell'umanità o dell'intelligenza dell'essere umano?

L'intelligenza umana (naturale) è un'intelligenza soggettiva («in prima persona») ossia l'intelligenza di un soggetto dotato di (auto)coscienza e di **intenzioni**.

E' possibile un'intenzionalità delle macchine (una coscienza artificiale)?



L'argomento della stanza cinese

John Searle, *Minds, brains, and programs*, *Behavioral and Brain Sciences*, vol. 3, 1980.

Intenzionalità e macchine

John Searle: l'intenzionalità come caratteristica distintiva della coscienza e l'impossibilità per un sistema di IA, come quelli sin qui sviluppati, di possederla

La **dimostrazione** (un esperimento mentale): Searle immagina di sostituire un agente umano al calcolatore nel ruolo di esecutore di una specifica istanza di un programma e mostra come tale esecuzione possa avvenire senza alcuna forma di intenzionalità.

L'intenzionalità trova modo di esprimersi nella creazione di artefatti, ma non si trasferisce ad essi (**intenzionalità derivata**).

L'obiettivo: l'IA di cui ti puoi fidare

In molti casi, quali le applicazioni finanziarie e diversi sistemi critici dal punto di vista della sicurezza (velivoli, centrali per la produzione dell'energia elettrica, ..) non è possibile garantire la presenza di decisore umano (**human in the loop**).

Più una macchina diventa complessa/potente più risulta difficile controllarne il comportamento.

Il problema dell'**autonomia** e della responsabilità delle macchine. Si pensi alle auto a guida autonoma, ma anche all'uso dell'IA in ambito militare. Occorre bilanciare opportunità e rischi, caso per caso

Obiettivo: **trustworthy AI**. L'IA di cui ti puoi fidare, perché è in grado di rendere ragione/spiegare le conclusioni/previsioni formulate.