



AI – INTRODUZIONE

L'INTELLIGENZA ARTIFICIALE AL SERVIZIO DELL'UOMO

AI INTRODUZIONE

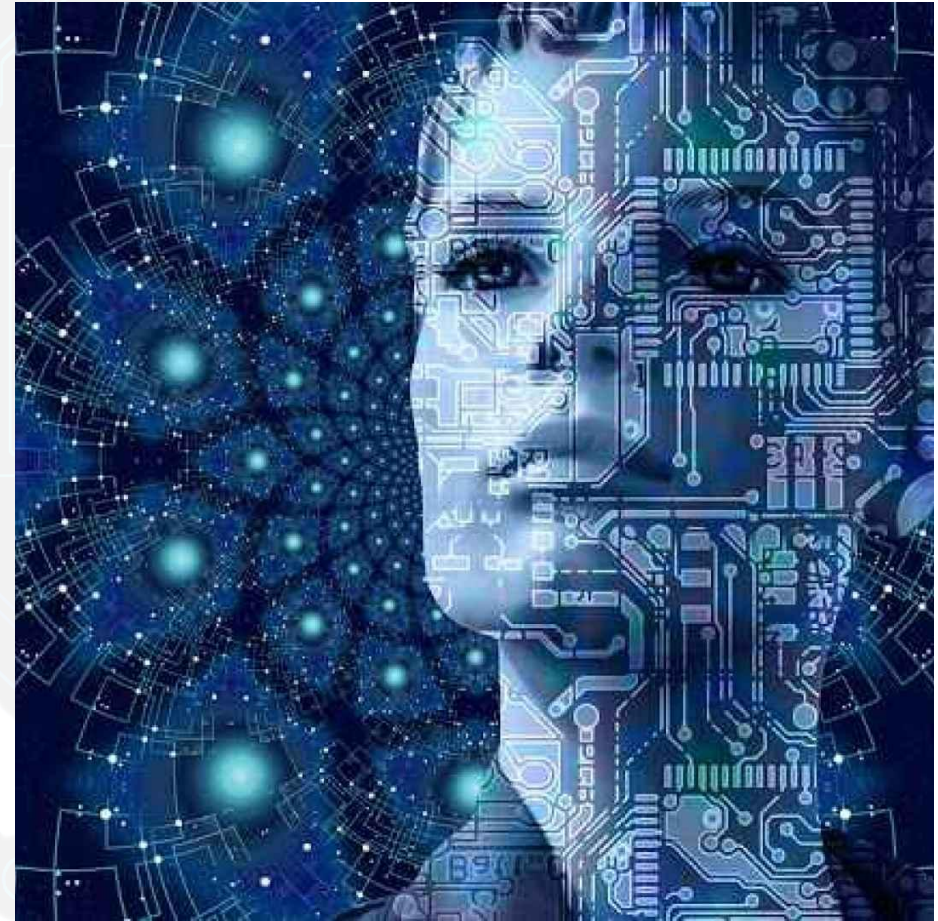
21/07/2025

I



ORDINE DEL GIORNO

- Corso Intelligenza Artificiale
- Obiettivo generale: offrire una preparazione teorico-pratica sull'IA, con strumenti e strategie per:
 - comprendere i meccanismi alla base dell'IA
 - stupire e coinvolgere
 - usare l'IA come alleato e non come sostituto



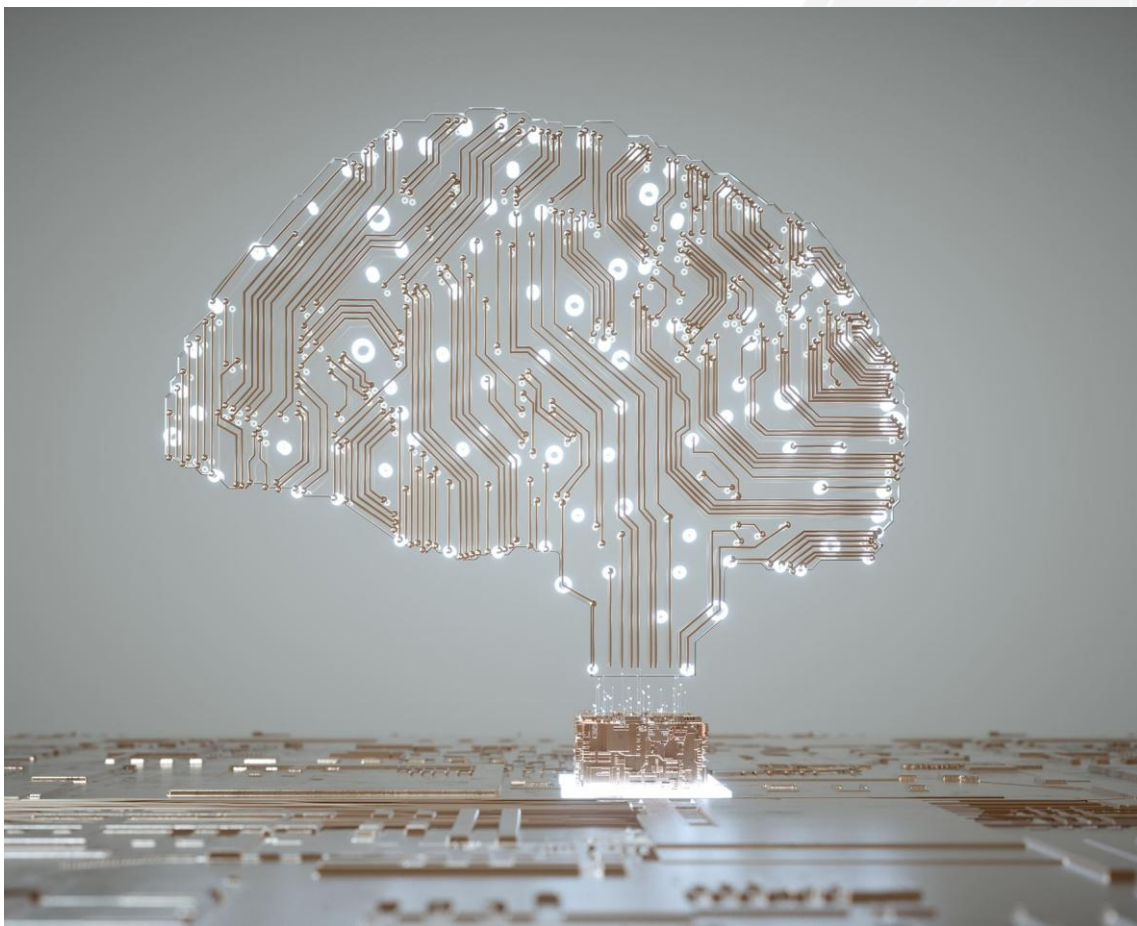


INTRODUZIONE ALL'INTELLIGENZA ARTIFICIALE

- Definizione di IA:
 - L'intelligenza artificiale (IA) è il ramo dell'informatica che si occupa di creare **sistemi in grado di svolgere compiti che normalmente richiederebbero l'intelligenza umana**, come il ragionamento, l'apprendimento, il riconoscimento del linguaggio o delle immagini.
 - Sistemi o macchine che imitano l'intelligenza umana per eseguire compiti e possono migliorare iterativamente basandosi sulle informazioni raccolte.
- Storia dell'IA:
 - Dalla nascita negli anni '50 con Alan Turing e il test di Turing, fino ai moderni sviluppi come il deep learning.
- IA anche in altri settori:
 - Traffico sulle mappe di Google,
 - Previsioni meteo, ecc.



COS'È L'INTELLIGENZA ARTIFICIALE?



- L'intelligenza artificiale (IA) simula l'intelligenza umana tramite computer e algoritmi.
- Permette alle macchine di apprendere da dati e migliorare nel tempo.
- Include campi come apprendimento automatico, elaborazione del linguaggio e visione artificiale.
- L'IA viene applicata in vari settori, dalla medicina all'educazione.
- L'obiettivo è supportare e potenziare le capacità umane, non sostituirle.



TIPI DI INTELLIGENZA ARTIFICIALE

- **IA Debole (Narrow AI)**
 - Sistemi progettati per eseguire un compito specifico
 - Non hanno capacità cognitive generali
 - Esempi: Assistenti vocali come Siri e Alexa, sistemi di raccomandazione di Netflix, algoritmi di riconoscimento facciale
- **IA Forte (General AI)**
 - Sistemi con capacità cognitive simili a quelle umane
 - Possono comprendere, imparare e applicare conoscenze in diversi contesti
 - Ancora teorica e non realizzata
 - La ricerca è in corso per sviluppare sistemi con queste capacità
- **Superintelligenza Artificiale (Superintelligent AI)**
 - Sistemi che superano le capacità cognitive umane in tutti gli aspetti
 - Inclusa la creatività, la risoluzione dei problemi e l'intelligenza emotiva



IA DEBOLE VS IA FORTE

Tipo di IA	Descrizione	Esempi
IA Debole (o ristretta)	Progettata per eseguire un compito specifico. Non ha coscienza o comprensione reale.	Assistenti vocali (Siri, Alexa), ChatGPT, traduttori automatici
IA Forte (o generale)	Ipotesi teorica: un'IA con intelligenza pari o superiore a quella umana . Può ragionare, pianificare, apprendere in modo autonomo e generico.	Non esiste ancora. È il soggetto di molti film di fantascienza (es. HAL 9000, Terminator, Her)

Esempi reali di IA Debole

1. Assistenti Vocali (Siri, Alexa, Google Assistant) → Riconoscono la voce e rispondono con azioni o informazioni.
Esempio: “Accendi le luci” → interpreta la frase → invia il comando a un dispositivo smart.
2. Traduttori automatici (Google Translate, DeepL) → Usano modelli di apprendimento profondo per comprendere e tradurre testi da una lingua all'altra. Si basano su milioni di frasi e strutture grammaticali.
3. Raccomandazioni intelligenti (Netflix, Spotify, Amazon) → Analizzano i tuoi comportamenti (cosa guardi/ascolti/acquisti) per suggerirti contenuti pertinenti.

 Domanda per gli studenti “In quali app o servizi che usi ogni giorno pensi ci sia dietro un'IA? Riesci a distinguere se si tratta di IA debole?”



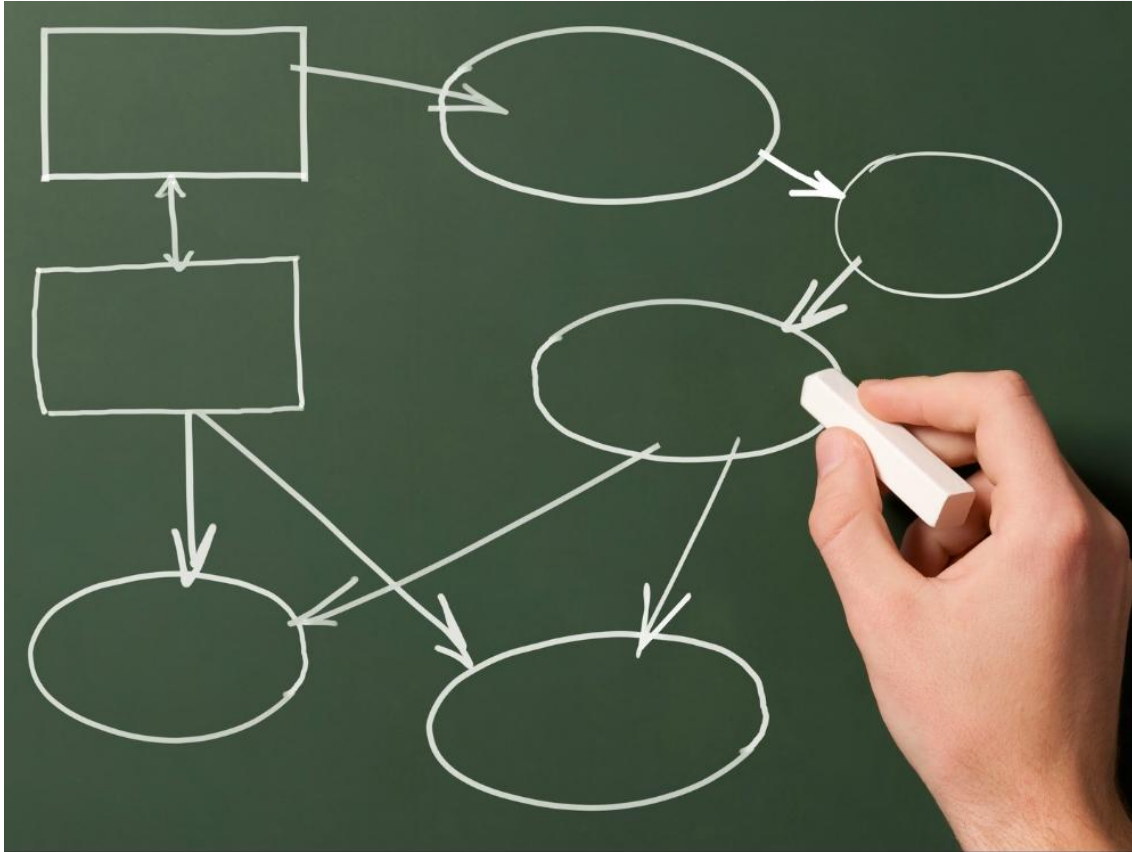
CLASSIFICAZIONE BASATA SULLE FUNZIONALITÀ

- **Macchine Reattive**
 - Rispondono a stimoli specifici senza memoria o capacità di apprendimento
 - Esempio: IBM Deep Blue - Wikipedia
- **Memoria Limitata**
 - Utilizzano esperienze passate per decisioni future per un periodo limitato
 - Esempio: Auto a guida autonoma
- **Teoria della Mente**
 - Comprendono emozioni e stati mentali umani
 - Permettono interazioni sociali più naturali
 - Stato attuale: In fase di ricerca e sviluppo
- **Autocoscienza**
 - Consapevolezza di sé e comprensione della propria esistenza





ALGORITMI DI IA

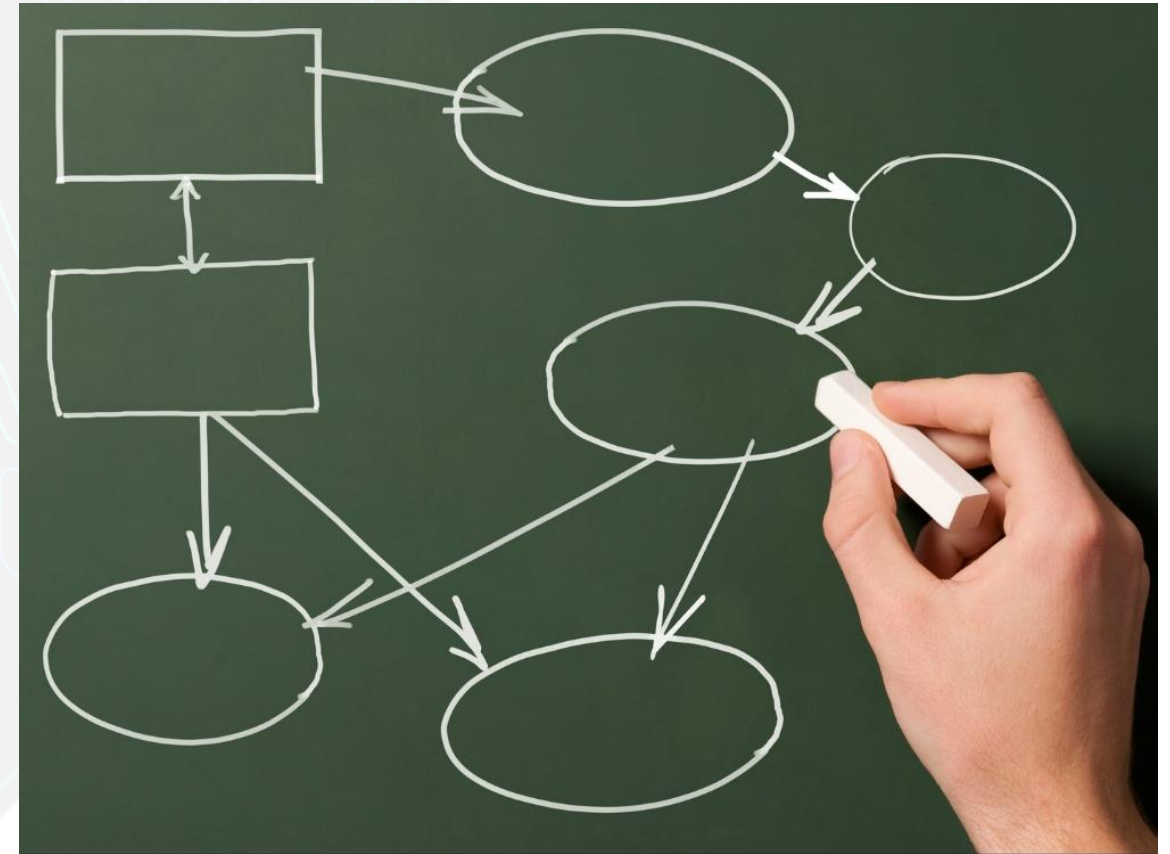


- Algoritmi di Apprendimento Supervisionato
 - Regressione Lineare: Previsione di valori continui (es. prezzi delle case)
 - Regressione Logistica: Classificazione binaria (es. rilevamento di email spam)
 - Alberi Decisionali: Decisioni basate su regole (es. rischio di credito)
 - Support Vector Machines (SVM): Separazione delle classi (es. riconoscimento facciale)
 - K-Nearest Neighbors (KNN): Classificazione basata sui vicini (es. raccomandazioni di prodotti)



ALGORITMI DI IA

- Algoritmi di Apprendimento Non Supervisionato
 - K-Means Clustering:
Segmentazione dei dati (es. segmentazione del mercato)
- Algoritmi di Apprendimento per Rinforzo
- Algoritmi di Reti Neurali e Deep Learning
- Algoritmi di Ensemble





LIMITI DELL'INTELLIGENZA ARTIFICIALE



Dipendenza dai Dati

Richiede grandi quantità di dati di alta qualità
Prestazioni compromesse senza dati adeguati
Esempio: Sistemi di riconoscimento facciale con immagini di bassa qualità



Bias nei Dati

Pregiudizi nei dati di addestramento
Decisioni discriminatorie
Esempio: Algoritmi di selezione del personale che discriminano



Costi Elevati

Sviluppo e implementazione costosi



Mancanza di Creatività



Problemi Etici



USO RESPONSABILE DEI MODELLI AI

Definizione: applicare l'IA con attenzione a principi etici, trasparenza, verifica e controllo umano costante.

Principi di riferimento (da Microsoft, ISO, Cisco, Google) Cisco:

- **Fairness:** trattare tutti equamente
- **Privacy:** salvaguardare i dati personali
- **Trasparenza:** rendere interpretabili le uscite
- **Responsabilità:** responsabilità umana sui risultati

Strumenti tecnici utili:

- **Model cards:** schede che descrivono origine dei dati, performance, limitazioni leonfurze.com
[Wikipedia](https://en.wikipedia.org/wiki/Model_card) [microsoft.com](https://www.microsoft.com/en-us/ai) [wired.com](https://www.wired.com) arxiv.org.
- **Toolkit di fairness** (es. IBM AI Fairness 360) [nypost.com](https://www.nypost.com).
- **Valutazione etica preventiva:** applicare checklist da ISO AI o DIU Responsible AI .

Best practice:

- **Audit regolari** dei dati e dei risultati
- **Coinvolgimento umano** nelle decisioni critiche
- **Documentazione trasparente** su obiettivi e limiti
- **Aggiornamento continuo** del modello con dati aggiornati e diversificati



I. ETICA DELL'IA

L'**etica dell'intelligenza artificiale** è il campo che studia *come progettare, usare e regolare l'IA* in modo che porti benefici alla società, prevenendo al contempo rischi e danni ([ibm.com](https://www.ibm.com)).

Principi chiave:

- **Equità:** evitare decisioni ingiuste o diseguali.
- **Trasparenza** (spiegabilità): capire come e perché un algoritmo prende una certa decisione.
- **Accountability** (responsabilità): stabilire chi risponde se qualcosa va storto.
- **Privacy e autonomia:** proteggere i dati personali e rispettare la libertà dell'individuo ([generazioneai.it](https://www.generazioneai.it)).



2. 🔍 BIAS ALGORITMICO

Un bias (pregiudizio) è un errore sistematico in un algoritmo che favorisce o penalizza in modo ingiusto alcune categorie.

Fonti comuni di bias:

1. Bias nei dati: se il dataset usato per addestrare un modello non è rappresentativo (es. poche persone di colore o donne), l'IA sarà meno accurata su quei gruppi (aisigil.com).
2. Bias di progettazione: scelte fatte dallo sviluppatore, come variabili utilizzate o metriche di ottimizzazione, possono introdurlo .
3. Bias di etichettatura: errori umani nella classificazione dei dati (es. segnale dialetti come offensivi).
4. Bias di interazione: chatbot che imparano stereotipi dai contenuti degli utenti (aisigil.com).

Effetti concreti:

- Errori nel riconoscimento facciale su persone con pelle scura (aisigil.com).
- Algoritmi di selezione del personale che penalizzano candidati non bianchi o donne (es. Amazon) .
- Sistemi di scoring creditizio o polizia predittiva che discriminano gruppi vulnerabili (ai4business.it).



3. 🏛️ DISCRIMINAZIONE ALGORITMICA

La **discriminazione algoritmica** si verifica quando il bias porta a trattare in modo ingiusto individui o gruppi protetti, spesso peggiorando le disuguaglianze sociali (aisigil.com).

Esempi reali:

- Sistemi sanitari che diagnosticano meno accuratamente certi gruppi etnici .
- Strumenti di selezione che favoriscono candidati bianchi o maschi .
- Sorveglianza predittiva poliziesca che colpisce quartieri con minoranze (ai4business.it).



4. ✓ USO RESPONSABILE DEI MODELLI AI

- **Uso responsabile** significa integrare considerazioni etiche *fin dall'inizio* dello sviluppo e in tutte le fasi del ciclo:
- **Ethics-by-design**: includere check su bias e trasparenza sin dalla progettazione ([protiviti.com](https://www.protiviti.com)).
- **Pulizia e bilanciamento dei dati** (pre-processing): rimuovere dati distorti, bilanciare la rappresentatività ([diritto.it](https://www.diritto.it)).
- **Algoritmi fairness-aware** (in-processing): usare tecniche che penalizzano le decisioni sbilanciate ([diritto.it](https://www.diritto.it)).
- **Correzione post-processing**: modificare gli output per ridurre distorsioni senza rifare tutto il modello ([aisigil.com](https://www.aisigil.com)).
- **Spiegabilità (XAI)**: rendere trasparenti le decisioni per utenti, sviluppatori e autorità ([ictsecuritymagazine.com](https://www.ictsecuritymagazine.com)).
- **Goverance e responsabilità**: definire chi è responsabile e con quali procedure si monitorano rischi, sicurezza e privacy .
- **Audit periodici**: valutare periodicamente se il modello continua a funzionare in modo equo.
- **Rispetto della normativa**: come l'**AI Act** europeo, che vieta IA con rischio troppo elevato, per esempio quelle di riconoscimento facciale in tempo reale ([ai4business.it](https://www.ai4business.it)).



PERCHÉ È IMPORTANTE

- • Consapevolezza: l'IA non è neutrale, ma riflette anche pregiudizi umani.
- • Responsabilità professionale: imparare a progettare sistemi equi e trasparenti.
- • Preparazione futura: l'AI Etica sarà sempre più regolamentata, e capire queste dinamiche sarà fondamentale nelle carriere.



GRAZIE

