

# Il Cambriano dell'Intelligenza Artificiale

Di Sara Colantonio - 1 Dicembre 2024



Immagine realizzata con un programma di Intelligenza Artificiale

Tempo di lettura: 8 minuti

🌀 **Affinché il potenziale trasformativo dell'IA possa manifestarsi in modo benefico, deve esserci la responsabilità di garantire che i suoi benefici siano distribuiti in modo equo, etico e sostenibile. Come nel Cambriano, siamo testimoni di un'evoluzione straordinaria; spetta a noi assicurarci che sia benefica per l'umanità** 🌀



Peso:1-83%,2-85%,3-88%,4-85%,5-88%,6-24%

Stiamo davvero vivendo il Cambriano della vita artificiale?

Secondo quanto argomenta Fei Fei Li, una delle figure di spicco nell'ambito dell'Intelligenza Artificiale (una delle poche donne menzionate come tali!), la risposta parrebbe affermativa. In suo recente intervento (1), Li ha ripreso il concetto di «esplosione cambriana» da Jensen Huang, Ceo di Nvidia, per articolare l'analogia tra l'era geologica di straordinaria diversificazione biologica e il nostro tempo, caratterizzato da un'evoluzione esponenziale di tecnologie «intelligenti», pronte a ridefinire profondamente scenari e società.

La scienziata mette in relazione l'esplosione della vita con la capacità biologica di sfruttare la luce e percepire l'ambiente circostante, tracciando un parallelo con lo sviluppo travolgente delle tecniche di intelligenza artificiale che hanno preso piede dopo i primi significativi successi in ambito percezione visiva, avvenuti negli anni dieci del XXI secolo e dei quali lei e il suo gruppo di ricerca a Stanford sono i principali artefici. Grazie alla costruzione del più grande database di immagini annotate, ImageNet (2), il gruppo di scienziati è riuscito nell'impresa di «insegnare alle macchine a vedere».

La metafora è indubbiamente evocativa, ma suscita anche riflessioni profonde sulle potenziali implicazioni dell'IA, siano esse utopiche o distopiche. In uno scenario in cui si affastellano quotidianamente notizie sulle grandi potenzialità e anche sui seri pericoli che tale tecnologia può comportare, è utile provare a fare un poco di chiarezza su cosa sia, come si sia evoluta e come funzioni.

## Cos'è l'Intelligenza Artificiale?

La stessa definizione di Intelligenza Artificiale è stata ed è oggetto di un dibattito continuo, che riflette la natura sfaccettata dell'intelligenza umana e lo studio e la comprensione della stessa.

Secondo colui che nel 1956 ha introdotto per la prima volta la locuzione stessa di Intelligenza Artificiale, l'informatico e neuroscienziato statunitense John McCarthy, «l'IA è la scienza e l'ingegneria del costruire macchine intelligenti». Secondo Stuart Russell e



Peso:1-83%,2-85%,3-88%,4-85%,5-88%,6-24%

Peter Norvig, autori del più importante libro di riferimento nell'ambito dell'IA, si tratta dello «studio di agenti che ricevono percezioni dall'ambiente e compiono conseguentemente azioni nello stesso». Secondo la Treccani, «l'IA è la disciplina che studia se e in che modo si possano riprodurre i processi mentali più complessi mediante l'uso di un computer. Tale ricerca si sviluppa secondo due percorsi complementari: da un lato l'IA cerca di avvicinare il funzionamento dei computer alle capacità dell'intelligenza umana, dall'altro usa le simulazioni informatiche per fare ipotesi sui meccanismi utilizzati dalla mente umana». Quest'ultima definizione è quella che riflette maggiormente la complessità del campo.

Intuitivamente, concepiamo l'intelligenza negli esseri umani come un insieme di capacità, quali comprendere, apprendere, ragionare e risolvere problemi, ma anche di agire e interagire, sentire e percepire l'ambiente, le persone e ciò che li circonda. Replicare queste abilità in una macchina si è rivelato notevolmente impegnativo. Tale difficoltà intrinseca ha plasmato l'evoluzione dell'IA e a una storia affascinante segnata sia da periodi di esuberante ottimismo sia da umilianti battute d'arresto. Curiosamente, la mancanza di una definizione precisa e universalmente accettata di IA ha probabilmente aiutato il campo a crescere, fiorire e progredire, giacché gli scienziati e gli sviluppatori nel campo sono stati guidati dall'imperativo di «darsi da fare» per arrivare a risultati significativi. Tutti gli sforzi in tal senso si sono concentrati su approcci diversi, che hanno portato a varie tecniche, ciascuna delle quali si è focalizzata su replicare o emulare un aspetto specifico della nostra intelligenza.

I primi tentativi di definire l'IA si sono concentrati sull'imitazione dei processi di pensiero umani («pensare come gli umani») o sul raggiungimento di prestazioni di livello umano («agire come gli umani»). Il Test proposto da Alan Turing esemplifica quest'ultimo approccio, valutando la capacità di una macchina di imitare in modo convincente la conversazione umana. Dopo un primo slancio, questa prospettiva è stata considerata poco rilevante, perché non necessariamente utile a comprendere i meccanismi alla base dell'intelligenza. Un approccio ritenuto più fondamentale, il «pensare razionalmente», ha mirato a formalizzare le leggi del pensiero utilizzando la logica, cercando di ottenere sistemi che ragionino correttamente. Tuttavia, anche questo approccio ha dimostrato dei



Peso:1-83%,2-85%,3-88%,4-85%,5-88%,6-24%

limiti, legati alla complessità intrinseca di tradurre varie forme di conoscenza (procedurale e implicita) in logica formale. Una definizione più scientificamente praticabile di IA è stata associata all'«agire razionalmente», vale a dire a progettare sistemi che eseguano azioni progettate per ottenere risultati ottimali. Questo approccio trascende i limiti dell'emulazione del comportamento umano, concentrandosi invece sui principi fondamentali dell'azione intelligente. Questa attenzione alla razionalità ha plasmato molto della ricerca moderna sull'IA, enfatizzando lo sviluppo di agenti o sistemi in grado di funzionare efficacemente in ambienti complessi e incerti.

## Approcci all'Intelligenza Artificiale

A realizzare queste prospettive si sono delineati due approcci principali, che, nel tempo, si sono sviluppati seguendo percorsi diversi, spesso contrastanti (3), portando a due grandi famiglie di metodi: l'approccio simbolico e l'approccio sub-simbolico. L'approccio simbolico, di tipo top-down, utilizza rappresentazioni simboliche della conoscenza, attraverso formalizzazioni esplicite e comprensibili agli umani, basate su logica, ontologie e regole. Tale approccio modella principalmente il ragionamento deduttivo e i processi decisionali deliberativi, di ottimizzazione e pianificazione ed è stato storicamente identificato come l'IA classica. Esempi eclatanti dell'approccio simbolico sono DeepBlue, il sistema Ibm che nel 1997 ha battuto l'allora campione di scacchi Gary Kasparov e il sistema Ibm Watson che nel 2011 ha gareggiato nel quiz a premi statunitense Jeopardy! contro i campioni Brad Rutter e Ken Jennings, vincendo il primo premio di un milione di dollari.

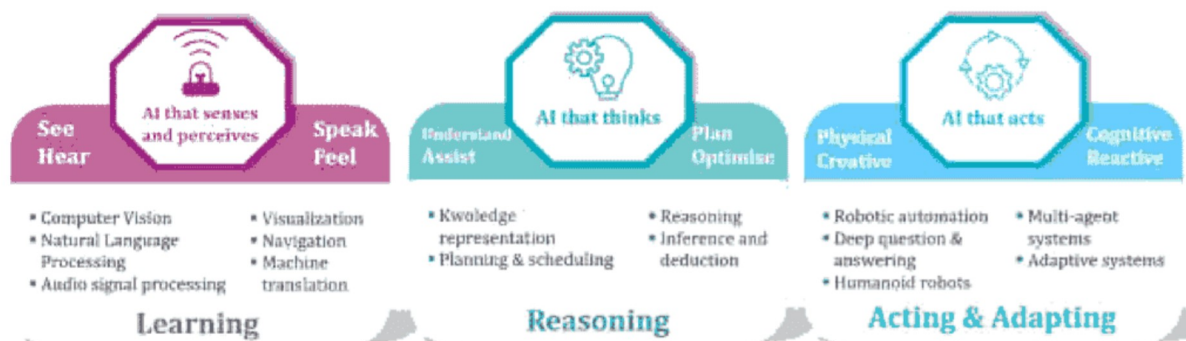
L'approccio sub-simbolico, di tipo bottom-up, modella principalmente le capacità umane di apprendimento e memorizzazione, per questo è stato storicamente indicato come apprendimento automatico (in inglese *Machine Learning*). Le reti neurali artificiali, gli algoritmi genetici, gli alberi decisionali e i metodi statistici, come le reti Bayesiane, rientrano in questo gruppo. La loro peculiarità è la capacità di scoprire modelli e informazioni da un insieme di dati o esempi, ovvero di apprendere dall'esperienza.

Questa capacità consente loro di imparare autonomamente a riconoscere categorie, percepire stimoli, prevedere tendenze, prendere decisioni e comprendere comportamenti.



In parole povere, tali metodi non sono programmati specificamente per risolvere un compito, ma imparano a risolverlo sulla base dei dati disponibili e dell'esperienza fatta sul campo. Nel loro caso, la conoscenza è implicitamente codificata nel sistema che risulta dopo l'apprendimento, ad esempio lungo connessioni tra i neuroni di una rete neurale artificiale, per questo risulta meno comprensibile all'esplorazione umana. Tutti i più grandi successi recenti si basano sull'approccio sub-simbolico, con un primo eclatante esempio registrato nel 2016 con la vittoria di AlphaGo di DeepMind contro il campione di Go Lee Sedol, seguito da un crescendo di applicazioni fino alla comparsa di chatGtp che ha segnato il definitivo uso comune dell'IA per milioni di persone.

Attualmente, l'IA si configura come una scienza ampia che raccoglie l'approccio classico e l'apprendimento automatico, fornendo sistemi che mostrano comportamenti intelligenti lungo tre assi principali: apprendimento, ragionamento e pianificazione, azione e adattamento all'ambiente. La definizione che ne deriva coincide con quella data recentemente dalla commissione europea e dal suo gruppo di esperti, secondo la quale: l'IA è l'abilità di una macchina di mostrare capacità umane quali il ragionamento, l'apprendimento, la pianificazione e la creatività.



## Breve storia dell'Intelligenza Artificiale

Il desiderio umano di replicare e persino superare le proprie capacità fisiche e intellettive appare connaturato alla nostra natura. Ne sono testimonianza gli echi nella mitologia greca, come il gigantesco automa di bronzo Talo o le ancelle d'oro, creati da Efesto, così come la letteratura e le arti figurative moderne. Tanto che Astro Keller ha argutamente definito «l'intelligenza artificiale come la scienza che studia come far fare alle macchine le cose che fanno nei film».



In realtà, la gestazione della scienza IA risale a poco prima della metà del secolo scorso, alle famose conferenze Macy (1943-1955), che videro riunirsi scienziati di diverse discipline, tra i quali Norbert Wiener (matematico), Margaret Mead e Gregory Bateson (antropologi), Warren McCulloch (neuropsichiatra), Walter Pitts (matematico), John von Neumann (matematico) e Lawrence Frank (sociologo).

Lo scopo principale era quello di gettare le basi per una scienza generale del funzionamento della mente umana. Si trattò di uno dei primi studi organizzati sull'interdisciplinarietà, che generò scoperte nella teoria dei sistemi, nella cibernetica e in quella che in seguito divenne nota come scienza cognitiva. Durante tali conferenze, nel 1943, McCulloch e Pitt presentarono il primo modello matematico di neurone artificiale, elemento fondante di tutta la teoria delle reti neurali artificiali. Nel 1950 Alan Turing pubblicò il suo lavoro seminale «Computing Machinery and Intelligence», definendo il test di Turing e una prima visione completa dell'IA. Il 1956 segnò la nascita ufficiale dell'IA, con la conferenza di Dartmouth organizzata da John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon.

Da lì una prima era di grandi aspettative (1952-1969), con l'ottimismo iniziale dovuto ai successi di programmi in grado di dimostrare teoremi e risolvere semplici problemi logici. L'attenzione era allora sui sistemi basati sulla conoscenza e sui sistemi esperti, ottenendo successi in ambiti specifici quelli del sistema Mycin in ambito medico per diagnosticare malattie infettive.

Il primo crollo registrò tra il 1972 e il 1981, quando i progressi rallentarono a causa dell'impossibilità di gestire i problemi del mondo reale e le limitazioni dei modelli fin lì presentati (4). Nel 1973, venne costruito in Giappone, Waibot-1, il primo robot umanoide, un'esperienza disastrosa. La prima rinascita si registrò negli anni 80, con le applicazioni commerciali, che spinsero gli investimenti e lo sviluppo. Nel 1982, John Hopfield presentò il suo modello di memoria associativa e nel 1986 DE Rumelhart Geoffrey Hinton e Ronald Williams presentarono l'algoritmo di retropropagazione dell'errore, elemento cardine dell'apprendimento neurale che ne sancì la rinascita ed



Peso:1-83%,2-85%,3-88%,4-85%,5-88%,6-24%

esplosione. Nel 1988 LeCun presentò la sua rete LeNet, capostipite delle reti profonde. In

quel periodo, fino ai primi anni 2000, si registrò tuttavia un nuovo crollo, legato alla ridotta capacità di calcolo e all'impossibilità di addestrare algoritmi realmente potenti.

È dai primi anni dieci che si è registrato il significativo balzo in avanti nelle capacità dell'IA che stiamo tuttora vivendo. Si tratta del cosiddetto tsunami dell'apprendimento profondo, reso possibile dalla sempre maggiore digitalizzazione, dalle crescenti capacità di calcolo e dalla maggiore disponibilità di dati. Nel 2007 inizia infatti il lavoro di Fei Fei Li e del suo team di Stanford per la raccolta di ImageNet e nel 2012 AlexNet fa registrare performance mai viste prima nel riconoscimento del contenuto di immagini.

La possibilità di sfruttare le potenzialità dell'apprendimento con rinforzo permette ad AlphaGo di battere Lee Sedol con una mossa che spiazzava completamente il campione fino ad allora imbattuto. Le reti generative, con i primi modelli presentati da Ian Goodfellow e seguiti dai Dall-E e Midjourney, spiazzano tutti noi, fino all'evento epocale a fine 2022 registrato con la presentazione di chatGpt e la sua apertura al mondo intero. Il premio Nobel per la fisica appena riconosciuto a Hopfield e Hinton riconosce in maniera decisa il ruolo dell'IA come tecnologia dirompente.

## La mappa non è il territorio

Oggi, l'IA non è più confinata ai laboratori di ricerca e alle conferenze scientifiche; permea la nostra vita quotidiana, influenzando sempre più il nostro modo di vivere. Tuttavia, è bene ricordare che gli attuali sistemi di IA, come quelli per la guida autonoma, per il gioco, gli assistenti virtuali, i modelli generativi per la scrittura e la grafica, i sistemi di supporto alla diagnosi medica sono ancora focalizzati su compiti specifici e concentrati principalmente su uno dei tre assi principali: apprendimento, ragionamento e pianificazione, azione e adattamento all'ambiente.

Queste soluzioni sono addestrate con dati specifici o programmate con procedure definite per operare all'interno di un ambito di azioni predeterminato. Le potenzialità soprattutto



Peso:1-83%,2-85%,3-88%,4-85%,5-88%,6-24%

dell'apprendimento con rinforzo e della generazione di contenuti ha assicurato nuove capacità mai pensate prima, proprietà cosiddette emergenti che portano a percepire un senso di reale intelligenza. Tuttavia, la caratteristica principale di queste soluzioni è che non possono essere generalizzate a compiti diversi di quelli per i quali sono stati sviluppati e le loro prestazioni tendono a decadere quando vengono esposte a condizioni non familiari. Volendo riprendere la metafora del Cambriano, abbiamo davanti dei trilobiti artificiali. Questo perché si tratta pur sempre di digitalizzazione e modellazioni matematiche di processi cognitivi e come ci insegna George EP Box tutti i modelli sono sbagliati, perché la mappa non è il territorio, gli fa eco Gregory Bateson. Tuttavia alcuni modelli sono utili ed è di questa utilità che dobbiamo beneficiare nel modo migliore possibile, per potenziare quelle che sono le nostre di capacità intellettive e cognitive.

È fondamentale riconoscere che l'evoluzione dell'intelligenza artificiale è un viaggio condiviso, che riflette le nostre aspirazioni più profonde e i dilemmi sociali ed etici che si manifestano davanti a ogni potente tecnologia. Affinché il potenziale trasformativo dell'IA possa manifestarsi in modo benefico, deve esserci la responsabilità di garantire che i suoi vantaggi siano distribuiti in modo equo, etico e sostenibile. Come nel Cambriano, siamo testimoni di un'evoluzione straordinaria; spetta a noi assicurarci che sia benefica per l'umanità.

(1) With Spatial Intelligence, AI Will Understand the Real World | Fei-Fei Li | TED

(2) <https://www.image-net.org/>

(3) Si veda a riguardo il dibattito tra Joshua Bengio e Gary Marcus

(4) Come nel famoso paper di Minsky e Papert del 1969 sui limiti delle reti neurali artificiali

*Sara Colantonio, Prima Ricercatrice, Istituto di Scienza e Tecnologie dell'Informazione*



Peso:1-83%,2-85%,3-88%,4-85%,5-88%,6-24%