



Salesiani

PER LA FORMAZIONE PROFESSIONALE
CNOS-FAP ETS

Pubblicata su *CNOS-FAP* (<https://www.cnos-fap.it>)

[Home](#) > Intelligenza artificiale (AI)

Intelligenza artificiale (AI)

Definizione(i) contemporanea(e) È importante notare che non esiste una definizione univoca di Intelligenza Artificiale (AI) universalmente accettata. La definizione che viene utilizzata spesso dipende dal contesto specifico ed è necessario tenere presente che l'AI è un campo in continua evoluzione con definizioni che potrebbero non essere via via sufficienti a catturare tutte le sfumature e le potenzialità di questa tecnologia dato il suo continuo sviluppo. In generale, l'Intelligenza Artificiale è definita come la “capacità di un sistema artificiale di simulare l'intelligenza umana attraverso l'apprendimento, il ragionamento e l'azione autonoma” (Norma ISO/IEC 42001:2023, standard per il sistema di gestione dell'Intelligenza Artificiale). La Commissione Europea, nella relativa al Piano di Coordinamento per l'AI, propone la seguente definizione di Intelligenza Artificiale: “con Intelligenza Artificiale si intendono sistemi che mostrano un comportamento intelligente analizzando il loro ambiente e intraprendendo azioni, con un certo grado di autonomia, per raggiungere obiettivi specifici.” Entrambe le definizioni richiamano il concetto di sistema AI o sistema basato sull'Intelligenza Artificiale che, seguendo la proposta definitoria dell'OCSE, viene definito come “un sistema automatizzato progettato per funzionare con diversi livelli di autonomia e che può mostrare capacità di adattamento dopo l'installazione e che, per obiettivi espliciti o impliciti, deduce, dagli input che riceve, come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali” (articolo 3, comma 1). Intelligenza Artificiale e Sistemi di Intelligenza Artificiale sono spesso usati come concetti intercambiabili ma c'è una differenza fondamentale: l'Intelligenza Artificiale è un campo ampio di studio che mira a sviluppare sistemi capaci di simulare l'intelligenza umana; i Sistemi di Intelligenza Artificiale sono realizzazioni concrete di AI funzionali a svolgere e/o risolvere compiti/problemi specifici (programmi o macchine che incorporano algoritmi). A tal proposito, il Regolamento europeo (2024/1689) auspica, e precisa, che la nozione di «sistema di AI» dovrebbe essere definita, oltre che in maniera chiara, in modo strettamente allineato al delle organizzazioni internazionali che si occupano di AI, al fine di garantire la certezza del diritto, agevolare la convergenza internazionale verso un'ampia accettazione del concetto, prevedendo, al contempo, la flessibilità necessaria per agevolare i rapidi sviluppi tecnologici. Il Regolamento, inoltre, fa riferimento, ai fini di un comune quadro definitorio, alla necessità di tenere conto delle principali caratteristiche dei sistemi di AI e, tra queste, soprattutto di quella fondamentale: la loro capacità inferenziale. Una capacità basata su tecniche che consentono l'inferenza attraverso approcci sia di automatico che consentano di imparare dai dati per conseguire determinati obiettivi, sia di quelli basati sulla logica e sulla conoscenza che possano trarre inferenze dalla conoscenza codificata o dalla rappresentazione simbolica del

compito da risolvere. In altri termini, sia la capacità dei sistemi AI di ottenere degli output (quali previsioni, contenuti, raccomandazioni o decisioni, che possono influenzare gli ambienti fisici e virtuali) che quella di ricavare modelli e/o algoritmi da input o dati. Il termine Intelligenza Artificiale, inoltre, contiene un esplicito riferimento alla nozione di intelligenza. Tuttavia, poiché l'intelligenza (sia nelle macchine che negli umani) è un concetto vago, i ricercatori in AI utilizzano principalmente la nozione di razionalità: «[...] la capacità di scegliere la migliore azione da intraprendere per raggiungere un certo obiettivo, dati determinati criteri da ottimizzare e le risorse disponibili». L' AI da un punto di vista operativo è, dunque, un "sistema razionale" di apprendimento che percepisce, attraverso alcuni sensori, l'ambiente nel quale il sistema è immerso, che raccoglie e interpreta dati, che decide qual è la migliore azione da intraprendere, che agisce attraverso alcuni attuatori, che modifica l'ambiente di riferimento e ne valuta il suo nuovo stato per determinare il successo della sua azione. L'Intelligenza Artificiale presenta, dunque, le caratteristiche distintive di una disciplina scientifica, dati il suo approccio rigoroso, l'enfasi sull'oggettività e la riproducibilità, la crescita continua del volume di , il suo impatto tangibile sul mondo reale e in svariati settori (ad esempio, medicina, ingegneria, finanza, trasporti, produzione), le connessioni con tante altre discipline (ad esempio, informatica, matematica, filosofia, psicologia, biologia, neuroscienze, ecc.), le sue ricadute in termini di progresso continuo. Provando a esemplificare in che modo l'Intelligenza Artificiale soddisfa i criteri chiave di un metodo scientifico rigoroso, dobbiamo tenere conto che l'AI opera attraverso un processo articolato nelle seguenti fasi: formulazione dell'ipotesi su come affrontare un tema, attuare una soluzione, ecc.; raccolta di dati oggettivi e concreti per testare le ipotesi, spesso in grandi quantità e da diverse fonti; sperimentazione attraverso il testing di modelli e algoritmi basato sui dati raccolti; analisi dei risultati per trarre conclusioni e valutare l'ipotesi originale; iterazione per affinare e testare nuovamente le ipotesi nella prospettiva di un processo di miglioramento continuo; riproducibilità dei risultati per assicurare che altri possano essere in grado di ottenere gli stessi risultati utilizzando gli stessi dati e metodi al fine di verificare e convalidare gli output/le soluzioni. In quanto disciplina scientifica, l'Intelligenza Artificiale comprende un'ampia gamma di approcci e tecniche, oltre ai ben noti machine learning e deep learning. Alcuni, tra i più importanti: il ragionamento basato su sistemi esperti guidati da un insieme di regole e conoscenze predefinite da un esperto umano per risolvere problemi in un dominio specifico (diagnosi mediche, sistemi di consigli finanziari, configurazione di reti di computer, ecc.); la rappresentazione della conoscenza e il ragionamento (logica simbolica) che utilizza simboli per rappresentare concetti e relazioni, e regole logiche per trarre inferenze e risolvere problemi (sistemi di pianificazione automatica, sistemi di tutoring intelligenti, robot che navigano in ambienti complessi); l'apprendimento automatico sia "supervisionato" (sistemi che apprendono da dati etichettati, come immagini con didascalie o testi con la risposta corretta) che "non supervisionato" (sistemi che identificano modelli e strutture da set di dati non etichettati e non strutturati) e "per rinforzo" (sistemi che apprendono attraverso prove ed errori, interagendo con un ambiente e ricevendo premi o penalità per le loro azioni); l'elaborazione del linguaggio naturale (NLP), distinguendo tra "comprensione del linguaggio naturale" (sistemi NLU che comprendono il significato del linguaggio umano, come intenzioni, entità e relazioni tra parole) e "generazione del linguaggio naturale" (sistemi NLG che generano testi simili a quello umano, come riassunti di documenti, risposte a domande o chat); la visione artificiale attraverso sistemi che estraggono informazioni (da immagini e video, da oggetti, scene e attività) che riconoscono, identificano e classificano gli oggetti (in immagini e video, in oggetti, scene e attività) che estraggono caratteristiche e informazioni da immagini (texture, forme e colori); la visione computazionale per analizzare e comprendere immagini e video; la robotica, in termini di "controllo" (sviluppo di algoritmi per controllare i movimenti e le azioni dei robot), di "percezione" (percezione dell'ambiente attraverso sensori come telecamere e radar), di "pianificazione e navigazione" (pianificazione dei percorsi e di movimento in ambienti complessi), di "apprendimento

automatico" (miglioramento delle prestazioni attraverso l'apprendimento dall'esperienza). All'interno dell'ampio campo dell'Intelligenza Artificiale, si possono considerare due macro-categorie (o paradigmi) di AI: l'intelligenza artificiale debole detta anche AI ristretta (narrow AI), si concentra su un compito specifico o un insieme di compiti strettamente correlati, dunque addestrata e orientata a eseguire attività specifiche. L'AI debole è alla base della maggior parte dell'AI che ci circonda oggi, abilitando, ad esempio, applicazioni come il riconoscimento facciale, l'automatismo del gioco degli scacchi, la traduzione automatica dei testi, Siri di Apple, Alexa di Amazon, IBM Watson, la personalizzazione dei feed in Facebook, Twitter e TikTok, oltre che i veicoli a guida autonoma; l'intelligenza artificiale forte, nota anche come Intelligenza Artificiale Generale (AGI) o Intelligenza Artificiale a livello umano (o superintelligenza artificiale ASI), è un tipo ipotetico di Intelligenza Artificiale che possiede un'intelligenza equivalente o superiore a quella umana e che riguarda macchine in grado di esercitare qualsiasi tipo di intelligenza umana, dotate di una coscienza autoconsapevole rispetto alla capacità di risolvere problemi, imparare e pianificare il futuro. Al momento, l'AI forte non esiste e il suo sviluppo è ancora oggetto di ricerca e dibattito scientifico date le sue implicazioni profonde per l'umanità, le cautele e le attenzioni necessarie in relazione alle sue opportunità e potenzialità. Nel frattempo, i migliori esempi di ASI possono essere tratti dalla fantascienza, come HAL 9000 (l'acronimo sta per Heuristic ALgorithmic), il supercomputer di bordo della Discovery nel film di "2001: Odissea nello spazio." Infine, merita particolare attenzione l'Intelligenza Artificiale generativa (AI generativa), ramo dell'intelligenza artificiale che ha catturato la maggiore attenzione da parte del "pubblico" soprattutto dopo il rilascio online da parte di OpenAI nel novembre 2022 della più nota AI generativa "ChatGPT" o Generative Pre-Trained Transformer. L'Intelligenza Artificiale generativa si concentra sulla creazione di nuovi contenuti in diverse forme, tra cui: testi (articoli di notizie, poesie, script, testi pubblicitari, email, interi libri), immagini (fotografie realistiche, dipinti, illustrazioni, modelli 3D, nuovi stili artistici), musica (brani musicali di generi diversi, suoni, composizioni originali), codici (software, script, algoritmi, nuovi linguaggi di programmazione), dati scientifici (molecole, composti chimici, strutture proteiche, interi genomi). L'AI generativa utilizza modelli di automatico complessi, come le reti neurali generative (GAN) e gli autoencoder variazionali (VAE), per estrarre modelli dai dati esistenti e generare nuovi contenuti che siano simili, ma non identici, ai dati di addestramento. I modelli di AI generativa vengono, infatti, addestrati su grandi quantità di dati (immagini, testi o codici). Durante l'addestramento, il modello impara a riconoscere le relazioni all'interno dei dati e, una volta addestrato, può essere utilizzato per generare nuovi contenuti simili ai dati su cui è stato addestrato.

Evoluzione storica

Le prime tracce di Intelligenza Artificiale come disciplina scientifica risalgono agli anni Cinquanta (periodo di grande fermento scientifico sullo studio del calcolatore e il suo utilizzo per sistemi intelligenti). Nel 1956, al Dartmouth College (New Hampshire), si tenne un convegno al quale presero parte i maggiori esponenti del mondo informatico. In quell'occasione si raccolsero i principali contributi sul tema, ponendo anche l'attenzione sugli sviluppi futuri. Durante il convegno di Dartmouth ebbe un ruolo fondamentale il lavoro di Alan Turing, considerato uno dei padri dell'informatica moderna (nel 1936 Alan Turing aveva posto le basi per i concetti di calcolabilità, computabilità e per la macchina di Turing). Nel 1950 lo stesso Turing scrisse un articolo intitolato "Computing machinery and intelligence", in cui proponeva quello che sarebbe divenuto: "una macchina poteva essere considerata intelligente se il suo comportamento, osservato da un essere umano, fosse considerato indistinguibile da quello di una persona". Grazie al lavoro di Turing, il tema dell'Intelligenza Artificiale ricevette una forte attenzione da parte della comunità scientifica e nacquero diversi approcci oltre che nuove tecnologie, anche se è doveroso sottolineare che il termine "intelligenza artificiale" non è stato all'inizio universalmente accettato e che alcuni studiosi preferivano termini come "cibernetica" o "informatica cognitiva". Nel 1958, John McCarthy inventa il linguaggio di programmazione Lisp, ancora utilizzato oggi nella ricerca sull'intelligenza artificiale. Nel 1959, Arthur Samuel

crea il programma "Samuel Learning Program", il primo programma di intelligenza artificiale a giocare a dama ad un livello competitivo. Nel 1966, Joseph Weizenbaum sviluppa la chatbot ELIZA, in grado di simulare una conversazione psicoanalitica. Nel 1969, Marvin Minsky fonda il MIT Artificial Intelligence Laboratory. Nel 1970, Edward Feigenbaum sviluppa il sistema MYCIN, un sistema di intelligenza artificiale per la diagnosi medica. Nel 1996 e nel 1997 il computer Deep Blue, progettato da IBM, sconfigge il campione del mondo di scacchi Garry Kasparov, dimostrando il potenziale dell'intelligenza artificiale nei giochi complessi. Nel 2001, un progetto internazionale utilizza l'intelligenza artificiale per sequenziare il genoma umano. Nel 2011, il computer Watson, basato su domande e risposte, vince il quiz televisivo americano "Jeopardy!", dimostrando la capacità dell'intelligenza artificiale di elaborare e comprendere il linguaggio naturale. Nel 2016, il programma di intelligenza artificiale AlphaGo sconfigge il campione del mondo di Go Lee Sedol, un gioco considerato molto più complesso degli scacchi. Dal novembre 2022, ChatGPT-3.5 diventa accessibile al pubblico e nel medesimo anno viene rilasciato Jurassic-1 Jumbo, uno dei modelli linguistici (LLM) più potenti e sofisticati al mondo, addestrato su un enorme set di dati di testo e codice, attualmente dotato di circa 180 miliardi di parametri. Il dibattito Negli anni più recenti è emerso come sia fondamentale avere un dialogo aperto e inclusivo sull'Intelligenza Artificiale, coinvolgendo esperti di diverse discipline, cittadini, governi e organizzazioni internazionali per definire linee guida etiche e normative, per un uso responsabile dell'AI a beneficio di tutta l'umanità, per ricercare e garantire un equilibrio organico e strutturale tra opportunità, vincoli e minacce. L'attuale dibattito è molto ampio e complesso, date le implicazioni etiche, sociali, economiche e filosofiche di vasta portata e riguarda anche la natura stessa dell'AI, la coscienza e il ruolo degli esseri umani in un mondo futuro sempre più plasmato, oltre che pervaso, dalla tecnologia. Le questioni chiave più discusse riguardano: l'impatto sull'occupazione, in quanto l'automazione alimentata dall'AI potrebbe generare una perdita di posti di lavoro e la necessità di una vasta, se non universale, riqualificazione dei lavoratori rispetto alle nuove competenze digitali; l'etica ed i pregiudizi, dato che gli algoritmi di AI, se poco trasparenti ed equi, potrebbero perpetuare pregiudizi esistenti nella società, discriminando gruppi di persone in base a fattori come razza, genere, religione o orientamento sessuale; la responsabilità in capo a coloro che progettano, sviluppano, attivano e dispiegano l'uso di AI; l'effettiva affidabilità delle informazioni e dei dati generati (output); la sorveglianza e la , dato che l'utilizzo dell'AI, in assenza di limiti universali chiari e garanzie per l'utilizzo di queste tecnologie, per la sorveglianza di massa e il riconoscimento potrebbe violare la privacy individuale e la libertà civile; il controllo e il potere, perché il dominio dell'Intelligenza Artificiale da parte di poche organizzazioni, aziende o governi potrebbe portare a concentrazioni di potere senza precedenti e minacciare la democrazia e la libertà individuale; il superamento dell'intelligenza umana da parte dell'Intelligenza Artificiale Generale (AGI) in tutti i campi apre interrogativi sul futuro dell'umanità, sul ruolo che l'uomo vorrà assumere in una società con macchine altamente intelligenti; il benessere e l'impatto ambientale; sebbene l'AI possa essere utilizzata per ottimizzare l'uso delle risorse, ridurre l'inquinamento e combattere il cambiamento climatico, vengono considerate le potenziali conseguenze negative sull'ambiente e la biodiversità (carbon footprint). Bibliografia Turing A.M., On Computable Numbers, with an Application to the Entscheidungsproblem, in Proceedings of the London Mathematical Society, 58, novembre 1936, pp. 230-256. Iwathaki A. – V. Anuradha (a cura di), The Institute for Other Intelligences, Los Angeles CA, X Artists Books, 2022. Howe A., Artificial Intelligence for Beginners: An Introduction to Machine Learning, Neural Networks, and Deep Learning, Astrid Howe, 2023. Pancioli C. – P.C. Rivoltella, Pedagogia algoritmica. Per una riflessione educativa sull'Intelligenza Artificiale, Scholé, 2023. Varanini F., Splendori e miserie delle intelligenze artificiali alla luce dell'umana esperienza, Milano, Guerini & Associati, 2024. Pasquale F., New Laws of Robotics: Defending Human Expertise in the Age of AI, Cambridge MA, Harvard University Press, 2020 (trad. it. Le nuove leggi della robotica.

Difendere la competenza umana nell'era dell'intelligenza artificiale, Roma, Luiss University Press, 2021). D'Acquisto G., *Intelligenza artificiale. Elementi*, Torino, Giappichelli, 2021. Tulchinsky I – C.E. Mason, *The Age of Prediction. Algorithms, AI, and the Shifting Shadows of Risk*, Cambridge MA, The MIT Press, 2023. Crawford K., *Atlas of AI. Power, Politics, and the Planetary Costs of Artificial Intelligence*, New Haven CT – London, Yale University Press, 2021. Floridi L. – F. Cabitza, *Intelligenza artificiale. L'uso delle nuove macchine*, Milano, Bompiani, 2021. Floridi L., *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide*, Milano, Cortina, 2022. Floridi L., *The Ethics of Artificial Intelligence. Principles, Challenges, and Opportunities*, Oxford, Oxford University Press, 2023. Pasquinelli M., *The Eye of the Master. A Social History of Artificial Intelligence*, London – New York NY, Verso, 2023. Pellerrey M. (a cura di), *La transizione digitale e i processi formativi: opportunità e pericoli*, Rapporto finale, CNOS-FAP, 2023. Cristianini N., *Machina Sapiens. L'algoritmo che ci ha rubato il segreto della conoscenza*, Bologna, Il Mulino, 2024. Bostrom N., *Superintelligence. Paths, dangers, strategies*, Oxford, Oxford University Press, 2014 (trad. it. *Superintelligenza. Tendenze, pericoli, strategie*, Torino, Bollati Boringhieri, 2018). Calzone O., *Machine learning. Introduzione all'apprendimento automatico*, Faraday Keen, 2018. Marmo R., *Algoritmi per l'intelligenza artificiale. Progettazione dell'algoritmo, dati e machine learning, neural network, deep learning*, Milano, Hoepli, 2020. Quintarelli S. (a cura di), *Intelligenza artificiale. Cos'è davvero, come funziona, che effetti avrà*, Torino, Bollati Boringhieri, 2020. Russell S.J. – P. Norvig, *Artificial Intelligence: A Modern Approach. Global Edition*, London, Pearson Education, 2021 (trad. it. *Intelligenza artificiale. Un approccio moderno*, Milano, Pearson, 2021). Schneider S., *Artificial You. AI and the Future of Your Mind*, Princeton NJ – Oxford, Princeton University Press, 2019. Davenport T.H. – J. Kirby, *Only Humans Need Apply: Winners and Losers in the Age of Smart Machines*, New York, HarperCollins, 2016. Walsh T., *Machines Behaving Badly: The Morality of AI*, Cheltenham, The History Press, 2022. Poggio T. – M. Magrini, *Cervelli Mente Algoritmi. Il mistero dell'intelligenza naturale, gli enigmi di quella artificiale*, Milano, Sperling & Kupfer, 2024.

URL di origine:<https://www.cnos-fap.it/parola-chiave/intelligenza-artificiale-ai>