

# RUBRICA: INTELLIGENZA ARTIFICIALE

## Info Autori :

<sup>1</sup> Psichiatra, Docente a contratto di Psichiatria, Corso di Laurea in Logopedia, Università di Roma Tor Vergata

<sup>2</sup> Psichiatra, già Docente di Psicomatica, I Scuola Specializzazione Psichiatria, Università di Roma la Sapienza

Rosalba Spadafora <sup>1</sup>, Antonio Bernabei <sup>2</sup>

## LA MENTE E LE MACCHINE “...CHE FINE HA FATTO HAL 9000?”

### ABSTRACT

Alan M. Turing, the well-know British mathematician, in the early 1950s, posed the problem of computers intelligence, and if they could, beyond programmed applications of calculation, have an emotional-affective inner life. The response was operational, based on the linguistic behavior of the machine: if the judge can't tell the difference, based on a blind interaction, who is man and who is machine, the machine is as intelligent as man; if the machine refers to a question feelings, thoughts and emotions, because it was so programmed, the machine has an inner life like man.

While considering the progressive improvement of hardware and software, and while acknowledging the brilliance of the argument, the authors express perplexity.

Machine learning, or artificial intelligence, the science and engineering of making intelligent machines, is now realized in numerous applications, from robots to chatbots, in daily life, and in medicine. Particular attention is given to applications in clinical psychology and psychiatry.

### RIASSUNTO

Alan M. Turing, il noto matematico britannico, nei primi anni '50 pose il problema dell'intelligenza dei computer, e se essi, oltre le operazioni di calcolo programmate, potessero avere una vita interiore, affettivo-emotiva.

La risposta fu operativa, basata sul comportamento linguistico della macchina: se un giudice non riesce a distinguere, sulla base di una interazione alla cieca, chi sia l'uomo è chi la macchina, la macchina è intelligente come l'uomo; se la macchina riferisce a domanda sentimenti ed emozioni, perché così programmata, ha una vita interiore come l'uomo.

Pur considerando il progressivo perfezionamento di hardware e software, e pur riconoscendo la brillantezza dell'argomentazione, gli autori esprimono perplessità.

Il machine learning, o intelligenza artificiale, scienza e ingegneria di fare macchine intelligenti, si concreta ormai in numerose applicazioni, dai robot ai chatbot, nella vita quotidiana, e in medicina. Una particolare attenzione è data alle applicazioni in psicologia clinica e psichiatria.

## Keywords:

Intelligenza Artificiale, Apprendimento automatico, Vita emozionale intima della macchina, Applicazioni in psicologia clinica e psichiatria

## Corresponding author:

Antonio Bernabei  
Mail: ntnbernabei@gmail.com

## Contributions:

the author(s) made a substantial intellectual contribution, read and approved the final version of the manuscript, and agreed to be accountable for all aspects of the work.

## Conflict of interest:

the author(s) declare no potential conflict of interest.

## Ethics approval and consent to participate:

not applicable.

## Patient consent for publication:

not applicable.

## Funding:

none.

## Availability of data and materials:

data and materials are available from the corresponding author upon request.

Received: 18 December 2024.

Accepted: 20 December 2024.

Se il lettore è incuriosito, o studioso, della Intelligenza Artificiale, e se, casualmente, si trova a passare a Ucknall, cittadina del Nottinghamshire, UK, può recarsi, per un dovuto omaggio, nella chiesa di St. Mary Magdalene, dove, nella cripta dei Byron, non visitabile, accanto al padre George Gordon Byron (1788-1824), il noto irregolare poeta, è sepolta Augusta Ada Byron, contessa di Lovelace (1815-1852), più nota come Ada Lovelace, nobildonna e raffinata matematica <sup>(1)</sup>, autrice di un algoritmo capace di calcolare i numeri di Bernoulli, algoritmo applicabile, anche, come software della macchina progettata, e mai completamente costruita, dal matematico contemporaneo Charles Babbage (1791-1871) <sup>(2)</sup>.

La Macchina Differenziale di Babbage restò allo stato di progetto, solo parzialmente realizzata e sperimentata, anche a causa del mancato finanziamento della Royal Society. Ma la Macchina di Babbage è il prototipo del computer, almeno come progetto, e Lovelace è l'autrice del primo software applicabile a una macchina.

Ada Lovelace inoltre ipotizzò la possibilità di comunicare con la Macchina con schede perforate (ricordate un centro di calcolo degli anni '70 e '80 del '900?) e si pose, del tutto astrattamente, il quesito della possibilità di pensare della Macchina.

Come si ricorderà, Daniel Bernoulli (1700-1782), matematico e fisico svizzero, compì studi di fluidodinamica (il principio di Bernoulli è utilizzato anche nella medicina cardiocircolatoria: il principio consiste nell'affermare una relazione inversa tra velocità e pressione di un fluido incompressibile, per cui, per esempio, se il sangue scorre in un vaso più piccolo di calibro, la sua velocità aumenta, e viceversa) e studi matematici. Bernoulli è tuttora attuale tra i matematici, che studiano sulla somma di potenze di numeri interi successivi e relative applicazioni.

La nota G, fra le note lasciate da Ada Lovelace, pubblicata da Luigi Federico Menabrea (1809-1896) <sup>(3)</sup>, tratta appunto il tema, e la sua applicazione per il calcolo nella macchina di Babbage.

Alan Mathison Turing (1912-1954) è noto al grande pubblico soprattutto per aver decifrato il codice Enigma, che, operato da una macchina dalla forma di una macchina da scrivere, cifrava e decifrava messaggi militari tedeschi, durante la II Guerra Mondiale, servendosi di tre o più rotori, la cui disposizione quotidiana era da configurare giorno per giorno, secondo uno schema prestabilito, segreto, dato

diverso ad ogni unità operativa: per esempio, in una flotta di sommergibili, ogni sommergibile aveva una diversa, unica, istruzione di disposizione rotori, nota, per quel giorno, solo a chi cifrava e trasmetteva e a chi riceveva e decifrava <sup>(4)</sup>. Molte delle operazioni militari britanniche, adiuuate dal codice Enigma, sono tuttora segrete.

Il lavoro di Turing sviluppò un lavoro iniziato dalla Intelligence polacca, e fu elaborato in collaborazione con altri matematici, polacchi, britannici e statunitensi. Alan M. Turing è poi noto per aver posto fine alla sua vita, nel 1954: bisessuale, scoperto, condannato, essendo, al tempo, delitto i rapporti omosessuali, alla prigione o, in alternativa, alla castrazione chimica, che scelse. Non tollerando le modifiche del corpo e della sessualità conseguenti, Turing si uccise. Turing è passato alla storia dello sviluppo intellettuale umano per aver proposto l'idea di una macchina immaginaria capace di eseguire ogni tipo di calcolo su numeri e simboli, detta poi Macchina di Turing (MdT), già nel 1936 <sup>(5)</sup>.

Una MdT <sup>(6)</sup> è un insieme di regole che definiscono il comportamento della macchina fisica su un nastro di lettura e scrittura: il nastro si immagina di lunghezza infinita, diviso in celle; una testina si sposta sul nastro e scrive o cancella simboli sulle celle del nastro. In accordo con il suo stato interno, la MdT, opportunamente programmata, legge in sequenza i simboli su nastro, e cambia il suo stato interno, e scrive un simbolo sul nastro, oppure sposta la testina a destra o a sinistra di una posizione. Dal punto di vista di uno psicologo comparato, la MdT, che concettualmente è alla base di ogni calcolatore, anche il più moderno, è un processo che simula il processo mentale umano, partendo dall'analisi degli elementi più semplici, mediante le subroutine, per poi calcolare il risultato finale di un programma, per esempio un calcolo matematico, eventualmente richiamando dati da una memoria.

La macchina, opportunamente programmata, è capace, dunque, di un processo analogo al pensiero, producendo un output, poi valutabile dall'uomo. La MdT è una macchina universale, una macchina che, opportunamente programmata, può lavorare su algoritmi differenti. Il pc sul nostro tavolo, acquistato già programmato per eseguire operazioni aritmetiche e test statistici, che noi adoperiamo fiduciarmente, come una scatola nera programmata dal fornitore per il calcolo aritmetico e statistico, è un esempio

di macchina universale. Così pure è una macchina universale il computer centrale di una Università, che, previa opportuna programmazione, già in memoria, o da inserire insieme ai dati da elaborare, serve numerosissimi terminali dei più vari Dipartimenti, oppure il computer centrale della Agenzia delle Entrate che, opportunamente programmato, accede a dati bancari e finanziari individuali. La differenza è nei tipi di algoritmi (programmi) forniti alla macchina, e nella estensione della memoria.

La struttura fisica della macchina fu detta da Turing hardware, e gli algoritmi, il programma, software. La MdT è l'evoluzione successiva delle macchine meccaniche, come il cifratore-decifratore Enigma, capaci di compiti limitati, non programmabili a piacere, ma con algoritmi fissi.

Nel 1950, Turing, trasferitosi dopo la guerra alla Università di Manchester, propose il "gioco dell'imitazione" <sup>(7)</sup>: un giudice, collocato in una stanza, interagisce con un essere umano e un computer, ciascuno in una stanza propria. Il Giudice deve decidere, interagendo per 5 minuti con ciascuno, chi sia il computer e chi l'essere umano.

Il computer è programmato in modo da simulare un essere umano, con paragonabili latenze ed errori nella risposta: con circa il 70% di probabilità, al test di Turing il giudice non era in grado di individuare la macchina. Turing prevede poi che nel giro di 50 anni l'evoluzione di hardware e software avrebbe reso impossibile distinguere l'uomo dalla macchina.

Al presente, si valuta che, facendo interagire 10 giudici umani con l'algoritmo di una macchina (una chatbot), nelle condizioni del test di Turing e per 5 minuti, almeno 3 giudici non sarebbero in grado di distinguere la macchina dall'uomo.

La nascita del calcolatore, prima meccanico, poi con valvole termoioniche, poi con semiconduttori e chip, e la sua evoluzione, avviene in circa 170 anni, in parallelo alla rivoluzione industriale, a due conflitti mondiali, a numerosi conflitti regionali, al progresso delle scienze: tutto questo produsse stimoli alla domanda e all'offerta dei calcolatori. Turing, a ben vedere, si pose il problema della intelligenza delle macchine in modo pragmatico e paradossale, da filosofo qual che egli era: se l'intelligenza può essere definita come la capacità di trovare nuove soluzioni a nuovi problemi, ovvero la capacità di attribuire un significato concettuale o pratico ai vari momenti dell'esperienza, nessuna di queste

definizioni si attaglia perfettamente a un calcolatore; se invece l'intelligenza attribuita si identifica con un giudizio di stima della 'intelligenza' da parte di un giudice interlocutore, come nel test di Turing, allora la definizione si fa interamente operativa, cioè il risultato di operazioni o processi che portano alla definizione significativa dei concetti, nel senso di Bridgman <sup>(8) (9)</sup>. La comunità scientifica, psicologica e filosofica in particolare, non percepì, al tempo di Turing, con particolare attenzione il problema dell'intelligenza dei calcolatori, sia perché la definizione implicitamente operativa di Turing non sollevava particolare allarme teorico, sia perché i calcolatori non erano dotati di capacità operative autonome, ma dovevano operare solo se programmati con un linguaggio di programmazione prescelto, inserendo, mediante schede perforate, con successiva traduzione in linguaggio macchina, il programma (diviso in main program e subroutines) e i dati.

Le schede perforate, come mezzo di comunicazione con la macchina, furono ideate da Babbage, come evoluzione del metodo di programmare i telai mediante lettura meccanografica, cui seguì, nella prassi informatica, la lettura ad aghi, e infine la lettura mediante luce, attraverso le perforazioni di schede, ultima evoluzione, ancora possibile, mediante appositi lettori, in centri di calcolo. Poi, vi furono computer capaci di ricevere informazioni provenienti da dischi rigidi, o da nastri magnetici, o da una tastiera, o da posizioni su uno schermo, come accade nel nostro palmare quotidiano, alimentato a batterie.

Nei primi anni '70, furono elaborate varie tecniche di elaborazione automatica in Psichiatria, capaci, a partire da scale ordinali (disturbi psicopatologici quantificati su una scala ordinale, in una cartella clinica predisposta in tal senso) o da punteggi a scale derivate dalla siglatura di test obiettivi o proiettivi, di offrire una ipotesi diagnostica, o una interpretazione narrativa di un test, o di fornire ed elaborare dati epidemiologici <sup>(10)</sup>. L'output, su carta stampata, era effettivamente stupefacente per chi assisteva alla generazione dell'output, ma concettualmente il sistema era semplice: operava, anche statisticamente, sui dati introdotti, sceglieva da una library i profili diagnostici o narrativi corrispondenti, eventualmente anche calcolando, con analisi discriminante, le probabilità di attribuzione diagnostica ad un gruppo fra alcuni

(per esempio, normalità, depressione, deviazione sociopatica, schizofrenia, ecc.); ovvero, sulla base dei punteggi a una cartella descrittiva, mediante scale ordinali, generare un rapporto clinico di status; ovvero calcolare incidenze e prevalenze di malattie psichiatriche e trattamenti in una popolazione, descritta con diagnosi, e dati clinici, personali, anagrafici e non. Nonostante la relativa semplicità dei programmi, e sottolineando la eseguibilità di un programma di analisi discriminante, capace di offrire la probabilità di appartenenza di un soggetto a un dato gruppo diagnostico fra alcuni, in un tempo quasi reale dall'immissione dei dati, si trattava di sistemi capaci di offrire un output simile, o anche qualitativamente superiore, a quello generato, in quegli ambiti limitati, da uno psichiatra o da uno psicologo clinico.

Al test di Turing, il giudice avrebbe avuto difficoltà a distinguere se interagisse con un umano o con un computer, o avrebbe più facilmente individuato il computer per la rapidità operativa e per il linguaggio impiegato, tutte variabili facilmente modificabili, in maniera di renderle indistinguibili, per linguaggio e tempi di esecuzione, per il giudice.

Si può affermare che l'imitazione del processo diagnostico, narrativo, e statistico, dei sistemi impiegati, realizzava un sistema intelligente artificiale, applicato a un ristretto settore psichiatrico e psicologico-clinico.

Dal tempo del progetto della Macchina di Turing, e dal tempo delle prime, semplici, anche se concettualmente innovative e praticamente utili, applicazioni del computer in Psichiatria (formulare una ipotesi diagnostica, interpretare narrativamente la siglatura di un test, elaborare dati epidemiologici), seguite e accompagnate dalle applicazioni di utilizzo in molte specialità mediche, molto tempo e molta evoluzione dell'hardware e del software è intervenuta, in ambito non medico e non psicologico esclusivamente. In California, è attivo, nel Lawrence Livermore National Laboratory, Los Alamos, il computer El Capitan, dedicato alla sicurezza ed efficacia del deterrente nucleare degli USA, alle variazioni climatiche, alla difesa e deterrenza strategiche, capace di 1,7 quintilioni (2 seguito da 18 zeri) di operazioni al secondo: si calcola che se tutti gli abitanti della Terra eseguissero una operazione al secondo, occorrerebbero 8 anni per uguagliare le operazioni che El Capitan esegue in un secondo.

Prestazioni un poco inferiori hanno il computer Frontier (Oak Ridge National Laboratory, Tennessee), dedicato a problemi di transizione energetica, cybersicurezza, sicurezza nucleare, e Aurora (Argonne Leadership Computing Facility, Illinois) dedicato all'analisi di scoperte e tecnologie fortemente innovative, come l'Intelligenza Artificiale. Non in classifica il computer Leonardo, centro di calcolo di Cineca (Consorzio Interuniversitario Calcolo Automatico), Bologna, in fase di assemblaggio definitivo, dedicato a tecnologie di supercalcolo e sviluppo di applicazioni di frontiera. Cineca ha una presenza distribuita, oltre che a Bologna, a Roma, Milano, Napoli, Chieti, con 4992 nodi di calcolo, 5308 utenti, ai fini di ricerca, comunicazione tra gruppi di lavoro, e didattica. Hardware e software di questi computer sono di costruzione umana; anche, parte del software può essere scritta dal computer stesso. Ma l'interazione con queste macchine, e la loro velocità e esattezza nel calcolo, solleva un problema che Turing si pose ormai 75 anni fa: il computer è intelligente? Il fatto che, almeno ordinariamente, i computer non hanno sensori ed effettori in relazione con l'ambiente esterno, e dipendono dall'ambiente esterno, controllato dall'uomo, per ricevere energia elettrica e termoregolazione, non dovrebbe indurre a concludere, forse semplicisticamente, che essi non abbiano una intelligenza autonoma? Oltre il "gioco dell'imitazione" ideato da Turing <sup>(7)</sup>, che più che valutare se un computer sia intelligente, valuta la perspicacia del giudice umano, rispetto però a un comportamento della macchina, senza entrare nei suoi processi cognitivi, si può parlare di intelligenza della macchina? Giunti, Pinna e Garavaglia <sup>(11)</sup> schematizzano le obiezioni possibili riguardo l'intelligenza della macchina, secondo Turing stesso <sup>(7)</sup>: esse possono essere matematiche o riguardanti l'argomento dell'autocoscienza.

L'obiezione matematica è che alcune limitazioni delle macchine non sono condivise dall'intelligenza umana, per cui l'intelligenza della macchina non sarà mai paragonabile a quella umana.

Turing risponde all'obiezione che è possibile che anche l'uomo stesso abbia limitazioni dell'intelligenza simili a quelle della macchina. L'obiezione è comunque considerata da Turing non facilmente liquidabile, e interessante. L'obiezione riguardante l'argomento dell'autocoscienza consiste nella considerazione che per affermare che una macchina abbia le stesse capacità di un cervello umano è necessario che essa

abbia sentimenti, emozioni, sensazioni, cioè una vita interiore, impossibile per una macchina, per questo non capace di uguagliare un cervello umano, ovvero di essere intelligente. Se, per ipotesi, una macchina dichiarasse di provare emozioni, sentimenti e sensazioni, questa sarebbe una pura finzione. La risposta di Turing a questa obiezione è che, se la macchina supera il “gioco dell’imitazione”, non si potrebbe negare una interiorità a tale macchina, perché, se lo facessimo, dovremmo negare una interiorità ad ogni essere umano, tranne noi stessi, perché l’unico modo che abbiamo di constatare una interiorità è interrogare una persona su quello che sta provando e sentendo, cosa che anche una macchina, superato il “gioco dell’imitazione”, potrebbe rispondere in modo indistinguibile da quello di un essere umano. Dunque, sarebbe falsa la premessa per cui nessuna macchina potrebbe avere una vita interiore.

Come si vede, Turing fonda la sua conclusione esclusivamente sul comportamento linguistico, non dando valore alla nostra intuizione di una vita interiore esclusiva degli esseri senzienti. L’argomentazione di Turing è filosofica e psicologica, fondata sul comportamentismo. La conclusione può essere essenzialmente soggettiva, e destinata a complicarsi a fronte di macchine dotate di organi di senso, effettori, memoria a breve e lungo termine, e programmate in modo da esibire, nella risposta non solo verbale, una personalità, o, se si preferisce, una organizzazione cognitiva. Si potrà dire che la macchina è costruita, e programmata da un essere umano; si potrà rispondere che esattamente qualcosa di simile avviene nel concepimento, nella gestazione, nella vita post-natale di un umano, e nell’apprendimento, anche scolastico, successivo.

Gli autori di questo articolo tendono a condividere, su base intuitiva-emozionale, la posizione dell’interiorità impossibile alla macchina, e dunque della macchina non capace di intelligenza equiparabile a quella umana, ma riconoscono che le obiezioni, che Turing fece a sé stesso, come detto, al riguardo, non sono facilmente superabili. Il dubbio sulla capacità intellettuale di una macchina che abbia superato il “gioco dell’imitazione” è, ancora al presente, probabilmente irrisolvibile, restando per ora confinato negli ambiti della neuropsicologia, dell’informatica, del cognitivismo-comportamentismo.

Quello che invece è del tutto attuale, sul piano psicologico, psicopatologico, e sociologico, filosofico, è il problema dell’atteggiamento, sul piano cognitivo ed emozionale, degli umani, che interagiscono con le macchine che abbiano superato il “gioco dell’imitazione”, nei settori più vari.

Oltre 70 anni dopo l’articolo di Turing su Mind <sup>(7)</sup>, mentre gli interrogativi sulla possibile interiorità delle macchine restano invariati, i punti di contatto, e le loro funzioni nella vita di tutti i giorni, tra umani e intelligenza artificiale (IA), sono vari (le macchine reattive, le macchine con memoria limitata, la teoria della mente, l’autoconsapevolezza delle macchine e le loro funzioni nella vita di tutti i giorni degli umani), ma soprattutto l’apprendimento automatico (machine learning), che è quel settore dell’IA alla base della progettazione e costruzione di robot e di macchine che prevedono, riconoscono, agiscono (esempi, fra tanti: prevedere l’andamento dei mercati, riconoscere immagini, guidare un’auto senza conducente o un veicolo spaziale). I sistemi di apprendimento automatico imitano struttura e funzioni delle reti neurali del cervello umano, e usano algoritmi messi a punto su insiemi di dati, per creare modelli che auto-apprendono, capaci di classificare informazioni e predire l’esito di processi senza intervento umano. Le reti neurali dell’apprendimento automatico sono modelli computazionali ispirati, vagamente, alla semplificazione di una rete neurale biologica <sup>(12)</sup>.

John J. Hopfield e Geoffrey E. Hinton, per il loro contributo ultratrentennale, per lo sviluppo delle reti neurali artificiali, hanno ricevuto il Premio Nobel per la Fisica, nel 2024. Schematicamente, la rete neurale è composta di tre “strati” di algoritmi: il primo strato tratta gli ingressi, in maniera di adeguarli alle richieste dei neuroni del secondo strato (se gli ingressi sono già adeguati, può anche non esservi); il secondo strato, strato “hidden”, nascosto, effettua l’elaborazione vera e propria e può essere costituito anche da un altissimo numero di neuroni, il terzo strato “O” è quello di uscita, e raccoglie i dati e li adatta alle richieste della rete successiva. L’apprendimento automatico differisce dalla programmazione tradizionale perché la macchina è predisposta ad apprendere in maniera autonoma dai dati, senza che siano state date esplicite istruzioni. Le reti neurali elaborano in continuazione le informazioni che ricevono dall’esterno, così come le reti neurali biologiche.

Algoritmi di machine learning sono utilizzati, ad esempio, nella nostra interazione a distanza, per operazioni bancarie le più varie, o quando acquistiamo online su Amazon, o quando operiamo sui social media, come X, Instagram, Facebook, o quando un'auto è nel traffico senza un umano alla guida, o quando attiviamo un robot, o interroghiamo una chatbot.

Come è noto, gli psicologi dell'educazione riconoscono tre tipi di apprendimento: mnemonico, strutturale, generativo. Il machine learning implementa tutti e tre i tipi di apprendimento. Il tipo mnemonico immagazzina le informazioni utilizzando codici verbali e codici non verbali, come le immagini; organizza informazioni strettamente intercorrelate in caselle da cui possono essere richiamate; crea forti connessioni per associazione tra elementi diversi.

Il tipo strutturale organizza i dati in mappe concettuali, diagrammi, flussi. Il tipo generativo incorpora nuove informazioni e nuovi contenuti all'interno della conoscenza già presente in memoria.

Gli stessi tre tipi sono operativi nel machine learning: l'apprendimento automatico, per esempio, calcola attivi e passivi del nostro deposito titoli; mette in memoria del nostro cellulare le foto scattate dal cellulare stesso e le ordina per data e ricorrenza; incorpora i dati GPS di un percorso, li paragona a informazioni già presenti in memoria, ci suggerisce l'itinerario più rapido in auto e a piedi; rileva pressione arteriosa e frequenza cardiaca e ci invia un segnale di allarme, se necessario; ordina dati di giurisprudenza in modo opportuno per argomento e data; e ancora. Spesso macchine che apprendono e IA sono usati come sinonimi, ma, a rigore, il termine IA si riferisce al tentativo generale di creare macchine capaci di eseguire processi cognitivi simili o identici a quelli umani, l'apprendimento automatico si riferisce specificamente all'uso di algoritmi e insiemi di dati per fare questo.

Sul piano psicologico è da osservare come noi utenti dell'apprendimento automatico lo accettiamo come un dato scontato, sicuramente esatto nelle conclusioni, operato da sistemi su cui operiamo a scatola nera, di cui ignoriamo il funzionamento, e conosciamo soltanto le procedure di accesso e uscita.

Nell'esaminare le possibili applicazioni dell'IA in psicologia e in psichiatria, sembra opportuno adottare la definizione di IA suggerita da Baker e Smith: "computers che eseguono compiti, di solito associati

con menti umane, particolarmente l'apprendimento e la soluzione di problemi" (13).

Il termine Intelligenza Artificiale è stato coniato da John McCarthy, nel 1955, in una conferenza, definendola "la scienza e l'ingegneria di fare macchine intelligenti" (14), e scrivendone poi (15).

La definizione suggerita da Baker e Smith, in evoluzione non contraddittoria con la definizione operativa di Turing, prima esaminata, guidò la ricerca, ambiziosa per il livello di sviluppo dei primi anni successivi, se l'IA si potesse impegnare in un ragionamento in più fasi, comprendere il significato del linguaggio naturale, generare nuovi piani per raggiungere obiettivi, ed effettuare complessi compiti motori. La ricerca non produsse in quegli anni i risultati sperati, e fu detto che l'IA conobbe "estati e inverni" (14), fasi di fioritura e sviluppo, e fasi di congelamento e arresto. Progressivamente, divenne chiaro che l'IA non è una tecnologia unitaria ed esattamente definita, ma un gruppo di tecnologie che imitano/emulano funzioni neurocognitive umane, ma che non ancora si avvicina alla effettiva intelligenza umana, che è il risultato di interazioni dinamiche tra molte componenti sottostanti sfaccettate.

Al presente, l'IA emula funzioni neurocognitive quali analisi visiva e classificazione, linguaggio ricettivo ed espressivo, ma è un analogo limitato dell'intelligenza umana, più complessa. Il software che simula ed elabora le conversazioni umane, in modalità scritta o parlata, è un chatbot: può elaborare in linguaggio naturale o in machine learning, apprendimento automatico.

I chatbot possono essere divisi in dichiarativi e predittivi.

I dichiarativi generano risposte automatizzate colloquiali alle richieste, verbali o scritte, dell'utente; questi chatbot, ad esempio, informano su orari di apertura e chiusura, film in programmazione, previsioni del tempo estratte da un data meteorologico. I chatbot, che funzionano con l'apprendimento automatico, sfruttano la capacità di comprensione del linguaggio naturale, sono consapevoli del contesto e, nei limiti dati, imparano sull'utente, la sue domande, la sua collocazione spaziale. Imparano le preferenze di un utente, forniscono raccomandazioni, fanno previsioni. Alexa di uso domestico è un chatbot predittivo; Google Search include, oltre la capacità di ricerca, l'apprendimento delle preferenze di ricerca che anticipa in offerta a fronte della digitazione di una domanda.

Assistenti digitali avanzati possono connettere più chatbot.

I chatbot hanno numerose applicazioni pratiche: una azienda può interagire con tutti i suoi clienti e dipendenti, una banca può interagire singolarmente con i suoi clienti, con proporzionale risparmio rispetto a una interazione umana da call center, o da sportello. Un chatbot è uno strumento di conversazione; un bot (abbreviazione di robot) è un programma autonomo che accede in rete: per esempio, nei social media l'utente umano può credere di interloquire con un altro umano, e invece interloquisce con un bot.

Un bot può essere un avversario virtuale in un videogioco, può seguire i link ipertestuali nei testi, per poi indicizzarli e depositarli nel database ipertestuale di un motore di ricerca, e poi fare spamming tra tutti coloro che entrano in una chat, ecc. Ancora, iniziano a comparire veri e propri robot (dal cane robot al robot che imita un domestico), che a volte sollecitano vissuti e relazioni simili a quelli reali, almeno quando sono in fase operativa: come dire che una fisionomia vagamente simile a quella naturale e una operatività paragonabile, a volte anche dotata di possibilità di espressione vocale sensata, possono portare a includere il robot, da parte del suo proprietario, in una sfera naturale o umanoide. Il futuro dirà.

Un chatbot potrebbe essere, opportunamente programmato, l'interlocutore di persone con disturbi del comportamento: per esempio, erogazione di suggerimenti in caso di craving dell'alcolista, di attacco di panico, di sconforto depressivo.

Tuttavia, la percezione delle sfumature emotive e dei sentimenti umani non sembra al momento possibile per il chatbot, che potrebbe interagire solo in forma standardizzata, anche audio, con una data persona <sup>(16)</sup>. Tuttavia, questo può essere, al limite, preferibile all'assenza di qualsiasi interlocutore che dia sostegno e ascolto. Impieghi della IA in settori limitrofi alla psicologia clinica, come l'elaborazione di dati di trattamenti psicologici individuali, raccolti in forma standardizzata in un database, sono realizzabili, e utili a fini di ricerca, ma sollevano evidenti problemi di riservatezza dei dati personali.

Da osservare che la ricerca ha documentato <sup>(17)</sup> effetti psicologici avversi su umani che interagiscono con algoritmi messi a punto da ditte che generano profitti, allo scopo di ottimizzare l'attaccamento cognitivo, comportamentale, emozionale con il loro servizio

(gioco on line, social media). Uno dei primi chatbot è stato ELIZA <sup>(18)</sup>, che simulava una conversazione con uno psicoterapeuta, sviluppato nel 1966 da Joseph Weizenbaum, basato su un sistema di sostituzione di stringhe e pattern matching, che dava, con questi accorgimenti semplici di ripetizione, spesso interrogativa, delle frasi del paziente e di generici inviti a proseguire e a ripetere precisando, una pura illusione di comprendere l'interlocutore. Weizenbaum stesso, da subito, dichiarò che ELIZA non era veramente intelligente. Ma ELIZA rappresenta il prototipo di una possibile relazione terapeutica tra macchina e mente: l'impiego di ELIZA in una psicoterapia reale non è ovviamente proponibile, mancando, in ELIZA, una qualsiasi teoria della mente e della psicopatologia, la capacità di percepire la comunicazione non verbale e di valutare comportamenti riferiti, la capacità di modularsi secondo le caratteristiche cognitivo-emotive del paziente, la sua storia, gli stressor di rilievo. Tuttavia, l'interazione con ELIZA, da parte di uno psicologo clinico o di uno psichiatra, era effettivamente stupefacente, e in un certo modo affascinante.

Nel 1973, fu presentato PARRY <sup>(19)</sup>, chatbot che intendeva simulare uno schizofrenico nel colloquio: l'interesse di mettere a punto una tale simulazione era nell'intento di formalizzare in algoritmi un pensiero schizofrenico tipico. Non è escluso che, in futuro, la RMN funzionale possa offrire un correlato morfofunzionale di una interazione con chatbot, dotata di algoritmi idonei a imitare la comunicazione umana, normale e patologica.

Una prospettiva nuova della intelligenza delle macchine è l'IA generativa, che è in grado di generare testo, musica, immagini, video, in risposta a prompt, richieste, che in informatica altro non sono che, in un sistema operativo a interfaccia testuale, simboli che compaiono sul video a indicare che il computer è pronto a ricevere un comando successivo, proveniente da un umano o da un altro computer.

L'IA generativa utilizza modelli statistici di una variabile osservabile e di una variabile dipendente. Modelli linguistici di grande dimensione, come quelli, per esempio, che redigono a richiesta un testo complesso (da un tema in classe a un articolo di giornale, a un lavoro scientifico descrittivo) si basano su un dataset (raccolta strutturata di dati archiviati e organizzati Per l'analisi e per l'elaborazione) di addestramento utilizzato per crearli.

Esistono attualmente molti sistemi di IA generativa, da Chat GPT a Bard Google, da xAI di E.Musk a Jais in lingua araba. Dolly2.0 è un modello linguistico di grandi dimensioni, accessibile a tutti, senza restrizioni, per qualsiasi finalità. OpenChatAI è addestrato con sette miliardi di parametri offline. Stable Diffusion è uno dei tanti sistemi capaci di generare immagini tridimensionali. L'IA generativa è in grado di applicazioni nel marketing, nello sviluppo di software, nella predizione di struttura proteica a partire da catene di aminoacidi, nell'ipotizzare strutture di farmaci, ecc., e ha ricevuto investimenti di molti miliardi di dollari, specie da Microsoft.

Recentemente, si è sentita la necessità di identificare univocamente la provenienza dei contenuti digitali prodotti dall'IA: una coalizione di Adobe, Microsoft, Nikon e altri, ha convenuto di creare una etichetta da far comparire su contenuti/immagini digitali prodotti da IA applicata su Adobe Photoshop, Adobe Premiere, Bing Image Creator <sup>(20)</sup>.

Accorgimento evidentemente utile, per esempio ma non solo, a limitare la diffusione di immagini modificate del reale, prodotte a fini malevoli o a corredo di fake-news. È da osservare, comunque, che se sette produttori hanno aderito al suddetto sistema Content Credentials, moltissimi non hanno aderito, e tra questi infinite entità politiche, di intelligence, finanziarie, criminali, ecc.

Anche, se, provvisoriamente, e su base intuitiva, si può ammettere che queste macchine, che operano mediante algoritmi, non abbiano una vita interiore paragonabile a quella degli umani, è da osservare che al test di Turing la distinzione tra uomo e macchina sarebbe difficile, o addirittura paradossalmente possibile perché la macchina, per rapidità ed esattezza, sarebbe superiore all'uomo. Inoltre, una macchina dotata di IA generativa potrebbe rispondere in modo convincente a domande riguardanti i suoi sentimenti, sue emozioni, il suo passato, le prospettive future.

Nella macchina, tutto questo dipende da algoritmi opportunamente programmati e dal data set, e nell'umano da algoritmi neuropsicologici, in parte geneticamente determinati e in parte dipendenti dall'esperienza: sostanzialmente, la differenza è che la macchina può essere rapidamente programmata (sviluppo di software e progetto/costruzione dell'hardware), mentre l'umano può essere costruito, fisicamente e intellettualmente, solo nell'arco di decenni.

Forse, solo la dimensione del flusso di tempo può distinguere, a parte le caratteristiche somatiche, uomo da macchina. Un sistema di intelligenza artificiale generativa si costruisce assemblando un sistema di apprendimento automatico non supervisionato e un insieme di dati: è possibile, in questo modo, ottenere: elaborazione del linguaggio naturale, traduzione del linguaggio naturale (digitare su Google "traduttore" e sperimentare); generazione di un codice sorgente per nuovi programmi informatici; generazione di immagini da testo; sulla base di dati biologici, costruzione di proteine e di strutture di nuovi farmaci; generazione di campioni musicali, a partire da un testo descrittivo di una melodia e sulla base di forme d'onda audio registrate, con un testo di commento annotato; clip video; elaborazione di domande e risposte complesse (vedi Alexa); costruzione di altri sistemi "figli" di intelligenza artificiale. Naturalmente, nessun sistema di IA generativa è capace di tutto questo, perché non è apparso conveniente e operativamente agevole concentrare tutto questo in un sistema unico, ma concettualmente è possibile creare un Golem del genere. Ci sono varie valutazioni sull'impatto economico e sociale della IA, che variano da uno stimolo all'occupazione alla distruzione di posti di lavoro. Quello che qui rileva è la possibile conseguenza psicologica e psicopatologica della frequentazione della IA: è un fatto che molte applicazioni della IA sono, per così dire, sotterranee, con scarso o nessun allarme sociale: dall'uso domestico di Alexa, percepito con simpatia, al Traduttore di Google percepito come utile; dagli impieghi in economia e finanza, poco avvertiti dai più, alla generazione di immagini modificate a scopi spesso illeciti. È prevedibile che un uso più esteso della IA, dal lavoro allo svago, dalla creazione di contenuti all'orientamento delle preferenze politiche, personali, morali, possa condurre a vivere in una sorta di realtà virtuale, parallela alla realtà effettiva, che non è, o non era, dominata da IA.

Sulle conseguenze sulla organizzazione cognitiva psicosomatica della frequentazione della realtà virtuale, è stato osservato <sup>(21)</sup> che essa rende più difficile il riconoscimento degli stati interni, emozionali e somatici; che le emozioni sono più difficilmente riconosciute e discriminate, in una sorta di alexitimia indotta, scoraggiando l'ascolto del corpo e la discriminazione delle emozioni, innescando il rifugio in realtà virtuale. Ancor più allarmante, se possibile,

è una eventuale crisi dei costrutti morali, mediata da sistemi di IA opportunamente costruiti, e la perdita della libertà personale di autodeterminarsi, per affidamento a sistemi automatici che consiglino scelte e comportamenti, con l'autorità di sistemi (pseudo) scientifici.

Secondo un rapporto di Zara Abrams <sup>(22)</sup>, American Psychological Association (A.P.A.), l'interesse degli psicologi, inclusi gli psicologi clinici, per l'evoluzione della IA generativa, è vivo e articolato in vari settori. Lo scopo dichiarato principale è il risparmio di tempo e risorse umane in dati compiti: secondo Kurt Gray, Università della North Carolina, GPT può sostituire l'umano in vari compiti complessi, per esempio dando giudizi morali sovrapponibili a quelli umani in 500 diversi scenari (è da chiedersi cosa avverrebbe a fronte di convulsioni sociali veloci, se percepite, o meno dalla IA generativa); secondo Abrams, ChatGPT e altri modelli di IA generativa stanno apportando cambiamenti all'insegnamento della psicologia, alla conduzione della ricerca, alla cura dei pazienti. Gli psicologi stanno accettando l'IA come una realtà con cui lavorare, piuttosto che lavorare contro il suo impiego.

L' IA produce rapidamente testi originali, immagini, audio, e rimarrà con noi in futuro, e non sembra possibile ignorare questo. Il Presidente del Mental Health Technology Advisory Committee (/practice/mental-health-technology-advisory-committee) dell'A.P.A. ha affermato: "Noi non possiamo controllare questo, così come possiamo formare una collaborazione strategica che ci aiuti ad accettarlo ed ottimizzarlo?".

Dunque, psicologi clinici e psichiatri devono fare ricerca sulla credibilità, competenza e moralità delle nuove tecnologie di IA, con le quali, nel futuro prevedibile, dovranno convivere nella pratica professionale, nella ricerca, nell'insegnamento. Al presente, riguardo l'interazione tra uomo e computer, tra menti e macchine, molta ricerca è centrata sulla tecnologia: è prevedibile, e auspicabile, una maggiore attenzione sull'umano in interazione. I chatbot in terapia possono essere utili nell'addestramento di psicologi e psichiatri in formazione, facendoli interagire con pazienti simulati dal computer, in comunicazione reciproca per iscritto, o meglio in audio.

L'interazione del soggetto in formazione potrà poi essere supervisionata da un tutor esperto.

Eventuali errori avverranno in virtuale, e non coinvolgeranno il paziente reale. È verosimile che in futuro la macchina possa sostituire, nei servizi pubblici che scarseggiano di personale, i terapeuti umani, almeno in parte e sotto supervisione. Secondo il rapporto A.P.A. citato, un uso del chatbot come terapeuta è stato fatto nei conflitti coniugali, nelle turbe del sonno, e nello stress da lavoro e da vita scolastica. Anche se le dimensioni trattate non sono tra le più gravi, il grado di perfezionamento e adeguatezza delle chatbot è evidentemente notevole.

Naturalmente, tra una sessione e l'altra di terapia computerizzata, sarà necessaria supervisione e controllo; è stato proposto anche un rapporto in chat da telefono cellulare, con interlocutore umano o automatico, tra una seduta e l'altra. Si può commentare, da parte di chi scrive, che soggetti depressi più facilmente potrebbero sviluppare sentimenti e idee di inaiutabilità durante la relazione con la chatbot, avvertita come una entità deumanizzata, in fondo indifferente all'umano.

Il sistema può funzionare, allo stato, fino a quando il paziente accetta la simulazione di un rapporto umano come fosse un rapporto umano reale.

Questo dipende, oltre che dalle caratteristiche individuali, dalla cultura sociale e dalle esperienze precedenti. La transizione, quando e se ci sarà, non sarà senza problemi. Tuttavia, nel futuro prevedibile, è possibile che la qualità percepita del rapporto uomo-macchina sia caratterizzato da un giudizio di competenza, da parte del paziente, pari o addirittura superiore a quello umano, ma che la credibilità della chatbot rimanga, a paragone, inferiore, in quanto non umana. Sembra fondamentale, perché la macchina resti credibile, che ci sia una supervisione umana che coinvolga anche il paziente, e il mantenimento di una possibilità di interazione verbale con un umano, anche in una chat via telefono cellulare.

Per completezza provvisoria, sono da citare tentativi di analizzare IA e i suoi comportamenti con costrutti psicoanalitici, giustificati dal fatto che procedimenti interni del software della macchina restano a volte non conosciuti; e tentativi di analizzare lo svolgersi di una psicoanalisi mediante IA.

A parte la connotazione “selvaggia” di tali tentativi, è da dire che essi sono del tutto iniziali, e pertanto di essi qui non si riferisce. Sarebbe comunque stimolante osservare la programmazione di una macchina come nevrotica, ed osservare la “cura” di essa da parte di uno psicoanalista, via interazione verbale audio e/o via interventi sul software. Ma questo, al momento, è soltanto immaginabile. Una macchina capace di imitare uno psicoanalista è fantasia fantascientifica, al momento. Infine si consenta l’auspicio che le macchine siano programmate in modo da rispettare le tre leggi della robotica, secondo lo scrittore di fantascienza Isaac Asimov <sup>(23)</sup>, che enunciano, formulate negli anni ’40 del secolo scorso, che un robot non può danneggiare un essere umano, o permettere che, causa il suo mancato intervento, un essere umano riceva danno; che il robot deve obbedire agli esseri umani, purchè gli ordini non siano in contrasto con la precedente legge; che un robot deve proteggere la sua esistenza, purchè questo non sia in contrasto con le leggi precedenti.

Nella pratica contemporanea, si può dubitare che, in vari contesti, i tre principi siano rispettati, ovunque, da tutti.

**Publisher’s note:** all claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.  
©Copyright: the Author(s), 2024

## BIBLIOGRAFIA

1. Baum J., *The Calculating Passion of Ada Byron*. Hamden, Archon Books, 1986.
2. Kim E., Toole B.A.T. , *Ada and the First Computer*, Scientific American, May 1999.
3. Menabrea L.F. , *Nota G di Ada Lovelace*. In *Sketch the analytical engine invented by Charles Babbage*. Bibliothèque Universelle de Genève, 1842.
4. Singh S., *Codici & Segreti*. Rizzoli, Milano, 1999.
5. Turing A., *On computable numbers, with an application to the Entscheidungsproblem\**. Proc. London Math. Soc., 42, pp 230-265, 1936 <https://www.abelard.org/turpap2/tp2-ie.asp> (\*problema della decisione, secondo Hilbert).
6. Brogi A., Romani F., *Le Macchine di Turing*, <https://pages.di.unipi.it/brogi>mdt>.
7. Turing A., *Computing Machinery and Intelligence*. Mind, LIX (236), 433-460, 1950 doi:10.1093/mind/LIX.236.433.
8. Bridgman P. W., *La logica della fisica moderna* , Bollati Boringhieri, Torino, 1952.
9. Somenzi V., *Operazionismo e relatività*, Rivista critica di storia della filosofia, X, 2, pp180-182, Franco Angeli, 1955
10. Cassano G.B., Castrogiovanni P., Pancheri P., Tansella M. *Tecniche di automazione in psichiatria*. Il Pensiero Scientifico, Roma, 1974
11. Giunti M., Pinna S., Garavaglia F. G., *Menti e macchine. Teorie filosofiche e scientifiche*, Mondadori Education, Milano, 2022
12. Burattini E., Cordeschi R. , *Intelligenza Artificiale*, Carocci, Roma, 2011
13. Baker T., Smith L., *Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges*. [https://mrdia.nesta.org.uk/documents/Future of AI and education v5 WEB.pdf](https://mrdia.nesta.org.uk/documents/Future%20of%20AI%20and%20education%20v5%20WEB.pdf)
14. Russel S., Norvig P., *Artificial Intelligence. A modern approach (2nd ed)*. Pearson, London, England, 2020
15. McCarthy J., Minsky M.L., Rochester N., Shannon C.E., *A proposal for the Dartmouth summer research project on artificial intelligence*. AI Magazine 27 (4) 12, 1955. <https://doi.org/10.1609/aimag.v27i4.1904>
16. Huang M. H., Rust R., Maksimovic V., *The feeling economy. Managing in the next generation of artificial intelligence (AI9 California Management Review, 61 (4), 2019*
17. Canadian Psychological Association, *Artificial Intelligence and Psychology*, <https://cpa.ca>File>CPD>
18. Weizenbaum J., *ELIZA---A Computer Program For the Study of Natural Language. Communication Between Man And Machine*. <https://oadoi.org/10.1145/365153.365168>
19. Network Working Group, RFC 439, PARRY encounters the DOCTOR , <https://tools.ietf.org/html/rfc439>, Internet Engineering Task Force, Internet Society, 1973
20. *Il simbolo Adobe che codifica i contenuti generate con AI*. <https://www.fastweb.it/fastweb-plus/digital-creativity/il-simbolo-adobe-che-codifica-i-contenuti-generati-con-AI/>
21. Bernabei A., Mollicella L., *Il rischio psicosomatico in realtà virtuale*. Medicina Psicosomatica 42, 4, 1997
22. Abrams Z. *What psychologists need to know about the evolution of generative AI*. American Psychological Association. 2024 Trends Report (/monitor/3034/01/trends-report)
23. *Three laws of robotics* (<https://www.britannica.com/topic/Three-Laws-of-Robotics>) Encyclopaedia Britannica, Inc.