*Riv. Psicol. Indiv.,* n. 97: 9-37 (2025)

# Il Giorno della Fornicazione. (L'alba degli Oggetti-Protesi-Sé Viventi)

Sergio De Dionigi

*La Redazione pubblica l'ultimo articolo del collega Sergio De Dionigi, recentemente scomparso. Impossibile scindere nel ricordo la persona di Sergio e la figura del dottor De Dionigi. Aveva un garbo squisito arricchito da un umorismo dai tratti ironici, arguti, imprevedibili. Era particolarmente attento agli altri, non soltanto come professionista, ma anche come amico, sempre presente e discreto. Riservatezza, delicatezza, sobrietà e cortesia erano caratteristiche spiccate della sua persona, ma anche del suo modo di occuparsi di allievi, colleghi, pazienti. Apprendere con la sua guida era un'esperienza autenticamente formativa: Sergio richiedeva precisione, completezza, chiarezza nei compiti assegnati e offriva mille suggerimenti per ampliare la tematica, approfondire, scoprire, sempre con la certezza di trovare in lui un riscontro competente ed esaustivo. Lavorare con lui alla stesura dei suoi scritti era un cammino sorprendente e appassionato, in cui si susseguivano idee originali e riferimenti dotti, tutto supportato da una conoscenza specifica e generale fuori dal comune. Così i suoi scritti parlano di lui, della sua professione e della sua persona, ricordano il suo sorriso e la sua voce lieve, i suoi modi sommessi.*
*Buon viaggio Sergio, grazie di tutto.*

*Summary* – THE DAY OF FORNICATION. (THE DAWN OF LIVING OBJECTS-PROSTHESES).
First, the constructs of self-object and narcissism are examined within the theorizations of Kohut and Kernberg, the concept of self-prosthetic objects is then introduced, differentiating the internal ones incorporated in a Self characterized by a borderline personality structure from those external features of a neurotic personality structure. This concept is then brought into the current social reality characterized by an increasingly predominant role of Artificial Intelligence which is also being approached by robots that can perform social functions. The danger of an implementation in the latter of a strong Artificial Intelligence (which for now does not exist and it is not known whether it will ever be introduced) which makes androids increasingly similar to humans is highlighted. This opportunity is called the day of fornication. The danger is not so much intelligent machines, but how much intelligence humans will attribute to them. Models that simulate empathy and good communication skills are already being produced. This situation can contemplate the possibility of constructing android replicas of humans with the qualities of these and without the defects that could be exhibited as the true representatives of those individuals in the context of increasingly impoverished social relations on a human level.

*Keywords: OBJECTS-PROTHESIS-SELF, OBJECTS-SELF, AI, ROBOT, INDIVIDUAL PSYCHOLOGY*

*Abstract* – IL GIORNO DELLA FORNICAZIONE. (L'ALBA DEGLI OGGETTI-PROTESI-SÉ VIVENTI). Vengono innanzitutto esaminati i costrutti di oggetto-sé e narcisismo all'interno delle teorizzazioni di Kohut e Kernberg, viene poi introdotto il concetto di oggetti auto-protesi, differenziando quelli interni incorporati in un Sé caratterizzato da una struttura di personalità borderline da quelli esterni di una struttura di personalità nevrotica. Questo concetto viene poi portato nell'attuale realtà sociale caratterizzata da un ruolo sempre più preponderante dell'Intelligenza Artificiale alla quale si stanno avvicinando anche i robot in grado di svolgere funzioni sociali. Viene evidenziato il pericolo di un'implementazione in quest'ultimo di una forte Intelligenza Artificiale (che per ora non esiste e non si sa se verrà mai introdotta) che renda gli androidi sempre più simili agli umani. Questa opportunità è chiamata il giorno della fornicazione. Il pericolo non sono tanto le macchine intelligenti, ma quanta intelligenza gli esseri umani attribuiranno loro. Sono già in fase di produzione modelli che simulano l'empatia e buone capacità comunicative. Questa situazione può contemplare la possibilità di costruire repliche androidi di esseri umani con le qualità di questi e senza i difetti che potrebbero essere esibiti come veri rappresentanti di quegli individui nel contesto di relazioni sociali sempre più impoverite a livello umano.

*Parole chiave: OGGETTI-PROTESI-SÉ, OGGETTI-SÉ, IA, ROBOT, INDIVIDUALPSICOLOGIA*

## I. *Gli Oggetti-Sé*

Heinz Kohut nell'ambito del costrutto teorico della *Psicologia del Sé* descrisse il concetto di *Oggetto-sé* [24, 25].

Egli descrive il Sé come un'entità unitaria, coesa e perdurante, che costituisce il centro dell'intenzionalità e l'elaboratore delle percezioni [25]. Questa definizione tuttavia può suscitare l'impressione che il Sé venga inteso in termini totalmente intrapsichici mentre è esattamente il contrario. Tra gli esponenti successivi di tale teoria Goldberg [18] ha posto in evidenza che il Sé è composto da unità permanenti di relazione.

Il Sé infantile non ha struttura durevole o una continuità nel tempo; per arrivare alla coesione, alla costanza e alla capacità di adattamento necessita dell'interazione con le figure genitoriali. Tali figure che contribuiscono alla maturazione e conservazione del Sé vengono definite *oggetti-sé* e costituiscono la matrice delle esperienze di oggetto-sé. Nella mente dei genitori sussiste un *Sé virtuale* del neonato che condiziona gli atteggiamenti che essi assumono nei confronti delle potenzialità del Sé del figlio. Da tali interazioni nel secondo anno di vita si stabilizza il Sé nucleare che assume le connotazioni di un Sé coesivo se si è sviluppato in modo fisiologico.

Per Kohut il Sé ha una strutturazione bipolare, un polo è costituito dal precipitato delle esperienze di oggetto-sé di rispecchiamento e viene definito *polo delle ambizioni;* l'altro è il *polo dei valori e degli ideali* che emerge dalle esperienze idealizzanti degli oggetti-sé. Tra questi due poli sussiste un *arco di tensione*, poiché i due poli spingono il Sé in direzioni differenti. Lungo questo arco di tensione sono disposte le potenzialità genetiche e le capacità acquisite.

Il *bisogno di rispecchiamento* consiste nella necessità del bambino di sentirsi accettato, apprezzato, considerato e riconosciuto dalle figure genitoriali; il *bisogno idealizzante* consiste nella necessità di essere parte di un oggetto-sé ammirato, potente, non angosciato, saggio e protettivo. Da ciò derivano i *bisogni alteregoici* vale a dire la necessità di esperire una somiglianza essenziale con l'oggetto-sé.

Si può notare una certa affinità tra il polo delle ambizioni e quello dei valori e degli ideali con i concetti adleriani di volontà di potenza e sentimento di comunità.
Attraverso queste due linee evolutive il bambino costruisce un "sano narcisismo" mediante lo sviluppo di un Sé grandioso che pretende di essere visto e rispecchiato e il Sé idealizzante bisognoso di affidarsi ad un'immagine parentale idealizzata.

Con Kohut il narcisismo assume caratteristiche fisiologiche, base per la coesione e crescita.

II. *Gli oggetti-protesi-Sé*

Kernberg [22] caratterizza la *struttura borderline di personalità* in base a tre elementi: una mancanza d'identità integrata, dei meccanismi di difesa primitivi ed un esame di realtà nel complesso mantenuto, ma non ben organizzato come invece nella struttura nevrotica di personalità.

La mancanza d'identità integrata è contraddistinta da una dispersione dell'identità, essa "si manifesta nell'esperienza soggettiva di un cronico sentimento di vuoto, in percezioni contraddittorie del Sé, in un comportamento contraddittorio che non può essere integrato in alcun modo affettivamente significativo, e in percezioni superficiali, piatte e impoverite degli altri" (22, p. 25). Ciò determina una notevole compromissione delle relazioni oggettuali poiché la persona possiede un'alterata percezione della continuità temporale di Sé e degli altri. Ne consegue una scarsa capacità di valutare se stesso e gli altri.

Anche Kernberg, successivamente a Kohut, parla di Sé grandioso [7] riferendosi alla personalità narcisista: sostiene il fatto che il narcisista sia un borderline che ha congelato in un Sé grandioso ciò che i pazienti più vicini alla diagnosi tradizionale di disturbo borderline di personalità vivono come violente oscillazioni dell'immagine del Sé e dell'oggetto. Il Sé grandioso del narcisista è costituito da un Sé ideale strutturato su parti idealizzate del Sé ideale, fuse con parti dell'Altro idealizzato e con parti del Sé reale. Tutte le parti non idealizzate vengono proiettate all'esterno insieme alle parti realistiche dell'Altro. In questo modo viene elaborata una coesione patologica del Sé su una illusoria convinzione di autosufficienza. In ciò si spiega la coriacea resistenza del narcisista a tutto ciò che può minacciare la posizione autoprotettiva del Sé grandioso. Tale situazione protegge da umiliazioni e sentimenti d'inferiorità e dai sensi di colpa relativi ad eventuali atteggiamenti aggressivi.

Appare evidente che per Kernberg il Sé grandioso è solo patologico.
In questo scritto il concetto di *narcisismo* verrà utilizzato esclusivamente con una connotazione patologica ben distinta dall'*autostima* intesa come fisiologica espressione della volontà di potenza.

Nella psicoterapia con pazienti affetti da disturbo narcisistico di personalità (in particolare con quelli che Gabbard definisce come *narcisisti inconsapevoli* [16] e Wink [47], basandosi sul Minnesota Multiphasic Personality Inventory, definisce come *narcisisti manifest*i) emerge che oggetti inanimati sembrano far parte della strutturazione grandiosa del Sé: abiti firmati, automobili *status symbol*, cellulari *fashion trendy* e residenze prestigiose.

Dai colloqui sembra emergere che non si tratti semplicemente di protesi esterne necessarie a una *politica di prestigio*, ma che costituiscano parti inglobate del Sé grandioso. La perdita di esse non è solo la perdita di un oggetto esterno protesico alla sensazione di grandiosità bensì la perdita lacerante di una parte del Sé.

Ciò vale anche per le persone: spesso i partner vengono scelti come oggetti-protesi che vengono pubblicamente esibiti e inglobati nel Sé grandioso. Essi perdono la loro caratteristica di Altro da Sé con propri tratti di personalità, ma le loro rappresentazioni psichiche vengono inglobate ed idealizzate nel Sé grandioso del narcisista. Anche i figli non sono estensioni del Sé grandioso, ma parte di esso e in quanto tali devono essere all'altezza della politica di prestigio come un corpo bello e atletico ed una mente brillante devono essere quelli del narcisista.

Oggetti protesi-Sé compaiono anche in disturbi di personalità riconducibili ad una struttura nevrotica di personalità, ma si tratta sempre di oggetti protesi-Sé vissuti come esterni al Sé. Nel disturbo dipendente di personalità il partner è un indispensabile punto d'appoggio, ma viene sempre vissuto come esterno e l'angoscia di perdita è sempre collegata ad un oggetto esterno. Persino nel disturbo borderline l'angoscia per la temuta perdita è (quasi) sempre collegata ad un oggetto esterno.

Max Tegmark [38], docente di Fisica al MIT e presidente del Future of Life Institute valuta il concetto di vita in base alla teoria dell'informazione e la descrive in senso lato come un processo che tende a conservare la propria complessità e a replicarsi. Al proposito sottolinea che ciò che viene replicato non è tanto la materia (composta di atomi) bensì l'informazione (fatta di bit) ed è essa a determinare la configurazione degli atomi. Facendo l'esempio di un batterio, allorché crea una copia del proprio DNA, esso non crea solo nuovi atomi, ma anche un nuovo gruppo di atomi configurati sull'insieme originale: è l'informazione che viene copiata. “In altre parole, possiamo pensare la vita come un sistema di elaborazione dell'informazione che si autoreplica e la cui informazione (il software) determina sia il suo comportamento sia i disegni del suo hardware” (37, p. 44).

In base a tale prospettiva egli classifica tre fasi della vita caratterizzate da livelli di complessità crescente.

La Vita 1.0 dell'evoluzione biologica non è in grado di riprogettare né il proprio hardware né il proprio software nel corso della sua esistenza poiché entrambi sono determinati dal proprio DNA e possono modificarsi solo in base all'evoluzione nel corso di diverse generazioni.

Nella Vita 2.0 gli esseri umani non possono riprogettare il proprio hardware, ma possono riprogettare in gran parte il proprio software: possono infatti acquisire abilità complesse sul piano culturale e tecnologico nonché rielaborare in modo continuo i propri fini e le proprie visioni del mondo.

Nella Vita 3.0 che non esiste ancora gli esseri umani saranno in grado di riprogettare sia il proprio hardware sia il proprio software senza essere condizionati dall'evoluzione biologica.

Tuttavia nella Vita 2.0 fin dall'antichità sono stati prodotti *oggetti tecnici* che potevano essere considerati un prolungamento della mano dell'artigiano o del guerriero che li usava [21] modificando quindi il proprio hardware somatico non solo a livello di schema corporeo, ma anche a livello di rappresentazione neuronale. Ciò malgrado l'oggetto veniva sempre vissuto come altro da sé ora invece i telefoni cellulari e gli smartphone sono vissuti come oggetti-protesi-Sé in grado di replicare le funzioni dei sistemi naturali e con prestazioni persino superiori appunto perché molti oggetti tecnici moderni vengono progettati non per essere presenti, ma per svanire.

Ce ne accorgiamo solo quando funzionano male o non li troviamo o quando li utilizziamo indipendentemente dalla loro funzione, usandoli ad esempio come fermacarte. Quando funzionano però si rendono spesso "trasparenti" [13]. È per tale motivo che possono essere intesi come oggetti-protesi-Sé. Se li smarriamo è come se una nostra funzione corporea risultasse menomata.

Ormai da qualche decennio vengono studiati e prodotti sistemi robotici leggeri e indossabili (soft robotics) per sopperire ad amputazioni, paraplegie, morbo di Parkinson e guida a distanza di pazienti affetti da cecità. Si progettano impianti sottocute che dovrebbero sostituire gli esoscheletri meccanici attualmente utilizzati.
La *bionica* si occupa dello sviluppo di sistemi artificiali. Si arriverà pertanto alla progettazione da parte degli umani anche del proprio hardware. Si arriverà alla fase del cyborg, l'entità costituita da tessuto umano e parti meccaniche.

La tecnologia che ci farà transitare dalla vita 2.0 alla vita 3.0 è l'Intelligenza Artificiale (IA), ma come giustamente sottolinea Simone Natale [33] non dobbiamo chiederci quanto le macchine siano intelligenti bensì quanto ci *appaiano* intelligenti.

## III. *Dall'opacità imperscrutabile alla magia dei chicchi di riso*

L'Intelligenza Artificiale nacque nella prima metà del XX secolo dalla commistione tra la cibernetica, la teoria del controllo, la ricerca operativa, la psicologia e l'informatica.

Nell'estate del 1956 si tenne a Dartmouth negli Stati Uniti un seminario a cui parteciparono Marvin Minsky, futuro fondatore del laboratorio di IA al MIT, Claude Shannon, l'inventore della teoria dell'informazione, il matematico John Mc Carthy, che negli anni sessanta fondò lo Stanford Artificial Intelligence Project con l'obiettivo di costruire entro un decennio una macchina pienamente intelligente, Nathaniel Rochester, un pioniere dell'ingegneria elettronica e Allen Newell della Carnegie Mellon University, che nel 1958 assieme ad Herbert Simon dichiarò che se fosse stato possibile elaborare una macchina capace di vincere a scacchi si sarebbe arrivati al cuore stesso dell'intelligenza umana. Fu Mc Carthy a coniare l'espressione *intelligenza artificiale* per distinguerla da un altro progetto che si stava affermando in quegli anni, la cibernetica.

L'oggetto del seminario consisteva nell'elaborazione del linguaggio delle macchine, le reti neurali, il *machine learning* (apprendimento automatico), capacità di astrazione, ragionamento, creatività.
Obiettivi molto ambiziosi se si tiene conto che nel 1956 i più avanzati enormi *cervelli elettronici* erano milioni di volte più lenti degli attuali smartphone.

Già allora per i non addetti ai lavori ciò che concerneva l'intelligenza artificiale veniva vissuto come qualcosa attinente alla magia. Come fa notare Natale [33], citando gli storici della cinematografia, i primi spettatori a fine Ottocento che assistettero alle prime proiezioni di pellicole cinematografiche erano convinti di assistere a spettacoli di magia poiché il proiettore era nascosto alla vista degli spettatori.

Egli dichiara che l'accostamento tra tecnologia e magia consiste nella loro *opacità*.
Il nostro stupore di fronte alle innovazioni tecnologiche consiste nella nostra incapacità di comprenderne il funzionamento così come rimaniamo stupiti di fronte ai trucchi dell'illusionista. Già nel 1962 l'astronomo e scrittore di fantascienza Arthur C. Clarke dichiarava nella terza delle sue tre leggi che “qualsiasi tecnologia sufficientemente avanzata è indistinguibile dalla magia” [8].

Si potrebbe osservare però che per gli appartenenti alle generazioni Z ed alfa tale stupore è inesistente poiché l'attuale tecnologia è data per scontata, quindi ancora più opaca. Natale pone in risalto il fatto che l'opacità dei media digitali non è riducibile alle conoscenze tecniche degli utenti, ma incorporata nel funzionamento stesso delle tecnologie informatiche.

Egli scrive: “I linguaggi di programmazione sono formati da comandi comprensibili per gli informatici, tali da permettere di scrivere un codice che svolge funzioni complesse.

Questi comandi, però, corrispondono al reale funzionamento della macchina solo dopo essere stati tradotti più volte in linguaggi di programmazione di livello inferiore e infine in linguaggio macchina, cioè l'insieme delle istruzioni in numeri binari eseguite dal computer. Il linguaggio macchina è un livello di astrazione tanto basso da risultare per lo più incomprensibile per il programmatore.
Così, neppure gli sviluppatori sono in grado di cogliere le stratificazioni di software e codice che corrispondono al reale funzionamento della macchina.

I computer sono in questo senso la scatola nera per eccellenza: una tecnologia il cui funzionamento interno è opaco anche per gli utenti più esperti" (33, p. 52). Andler [1] parlando dei dispositivi IA sottolinea il fatto che viene contraddetto il principio enunciato da Gian Battista Vico secondo cui *verum-factum*, cioè si conosce veramente solamente ciò che si fabbrica: se voglio sapere veramente come funziona un dato strumento lo devo fabbricare, ma ciò non è vero per IA poiché a volte sortisce risultati inattesi dagli stessi progettatori.

Si prospettò però il problema relativo al fatto che volendo elaborare computer dotati d'intelligenza artificiale si aveva una visione abbastanza *opaca* di che cosa fosse l'intelligenza umana.
Non è questo il luogo per addentrarsi nelle diverse definizioni dell'intelligenza umana l'importante è rilevare che fin dal seminario di Dartmouth si evidenziò un'anarchia di metodi che bene o male cercavano di imitare l'intelligenza umana finché all'inizio degli anni duemila si impose il paradigma dominante di un gruppo di metodi chiamato collettivamente *deep learning* o *deep neural networks*.

Va evidenziato che per i media IA si identifica con *deep learning* quando è soltanto uno degli approcci per creare macchine intelligenti, è solo un metodo tra i tanti del machine learnig in cui le macchine "imparano" non solo dai dati implementati, ma anche dalle loro "esperienze". Al riguardo si deve sottolineare la divisione quasi "filosofica" tra i ricercatori dell'IA tra IA simbolica e IA subsimbolica.

L'IA simbolica si basava all'origine sulla logica matematica, ma anche su come le persone descrivono i propri processi mentali coscienti. Melanie Mitchell dichiara che "In un programma di IA simbolica la conoscenza consiste di parole o di frasi (i "simboli"), tipicamente comprensibili da un essere umano, insieme a regole con le quali il programma può combinarle ed elaborarle per eseguire il compito assegnato" (30, p. 9).

L'approccio dell'IA subsimbolica si ispira invece alle neuroscienze e il primo programma subsimbolico è stato il *percettrone* elaborato alla fine degli anni cinquanta dallo psicologo Frank Rosenblatt [35] che si ispirò alle modalità con cui i neuroni elaborano l'informazione, vale a dire che un neurone somma tutti gli input inviati ad altri neuroni e se la somma di essi raggiunge una certa *soglia* il neurone *scarica*, ossia trasmette un impulso.

Particolarmente importante è il fatto che le connessioni tra neuroni, le sinapsi, hanno forze differenti per cui un singolo neurone assegna un peso maggiore a input provenienti da connessioni più forti rispetto a quelle più deboli. Tale elaborazione dell'informazione neuronale può essere simulata da un programma informatico denominato percettrone che riceve vari input numerici, ma un solo output.

Ad ognuno di questi input può essere assegnato un *peso* numerico per cui ciascun input viene moltiplicato per il suo peso prima di essere aggiunta alla somma. Inizialmente la *soglia* del percettrone venne stabilita dal programmatore, ma in seguito dal percettrone stesso.

Tuttavia alla fin degli anni Settanta del secolo scorso ci si rese conto che l'IA non aveva realizzato tutte quelle promesse che aveva tanto pubblicizzato. I finanziatori se ne accorsero e i fondi alle università scemarono. Iniziò uno degli *inverni* dell'IA (la storia dell'IA è contraddistinta dall'alternanza di inverni e primavere). Negli anni Ottanta cominciarono ad affermarsi le reti connessionistiche che oggi chiamiamo reti neurali.

In queste ultime gli elementi sono neuroni simulati, affini ai percettroni solo che le reti neurali sono a strati con uno strato di dieci unità di output e tre *unità nascoste*, vale a dire elementi che non inviano output, ma mettono in comunicazione tra loro i singoli elementi. Si strutturano così reti multistrato che oltre allo strato di unità di output si aggiungono due unità di strato nascoste. A differenza del percettrone le unità non si limitano a scaricare o meno, basandosi su una soglia, ma usano la somma degli input per calcolare un numero compreso tra zero e uno, denominato *attivazione* dell'unità.

Se la somma calcolata ha un'unità bassa l'attivazione è prossima a zero, invece se la somma è alta l'attivazione è prossima a uno. In questo modo la rete neurale può imparare ad usare le proprie unità nascoste per riconoscere elementi più astratti, come cerchi o quadrati, rispetto ad elementi più semplici, come i pixel codificati dall'input. In base a ciò venne sviluppato un algoritmo (vale a dire un metodo per risolvere un problema attraverso una serie precisa e finita di passi che garantiscono, se eseguiti in modo corretto, la soluzione) di apprendimento generale per addestrare le reti definito *retropropagazione*.

Tale termine indica la capacità di valutare un errore osservato nelle unità di output per "propagare" a ritroso la responsabilità dell'errore in modo da ripartire la responsabilità su ciascun peso della rete. In questo modo è possibile determinare di quanto modificare ciascun peso per ridurre l'errore. La retropropagazione funziona indipendentemente dal numero di input, unità nascoste e di unità di output della rete neurale.

Si arrivò quindi a discutere di IA *forte* ed IA *debole*.
Per IA forte s'intende l'IA capace di equiparare e superare le funzioni cognitive umane nella loro totale complessità. Alfiere di tale prospettiva è il geniale inventore Ray

Kurzweil, insignito nel 1999 dal presidente americano Clinton della National Medal of Technology and Innovation. Egli coniò il termine *singolarità* per indicare un evento unico in grado di sconvolgere l'umanità. Egli attualmente è più famoso per le sue previsioni che per le sue invenzioni: egli pone alla fine degli anni Trenta di questo secolo la comparsa di un computer superintelligente in grado di superare il test di Turing (vedremo in seguito di che cosa si tratta) e di una IA "umana", poiché in tale epoca saremo in grado di comprendere i principi operativi della mente umana e ricrearne la potenza con sostanze sintetiche [26].

Come la complessità cerebrale di un neonato si sviluppa dall'interazione con un mondo complesso così l'intelligenza artificiale potrà essere "educata" come quella umana. Kurzweil fonda la sua teoria sul concetto di "progresso esponenziale" che caratterizza la scienza e la tecnologia attuali, in particolare dell'informatica.
Per rendere facilmente comprensibile tale concetto viene sovente citata una favola.

Molto tempo fa un saggio proveniente da un villaggio ridotto alla fame capitò in un regno ricco e lontano. Egli aveva fama di essere un valente giocatore di scacchi e allora il re di quella regione lo sfidò ad una partita. Dapprima il saggio era riluttante, ma il re gli promise qualsiasi cosa egli volesse se l'avesse sconfitto. Il re venne sconfitto e il saggio chiese che venissero posti sul primo quadrato della prima fila della scacchiera due chicchi di riso, quattro sulla seconda, otto sulla terza e così via fino all'ultima casella dell'ultima fila.

Chiese anche che il riso venisse trasportato al suo villaggio. Il re pensando di cavarsela con poco accondiscese prontamente. Dapprima fu facile, poi si dovette ricorrere per i calcoli ai matematici di corte. La situazione divenne disperata quando al secondo quadrato della quarta fila (casella 26): erano necessari 2048 chilogrammi di riso (più di due tonnellate). Arrivato a questo punto il re implorò il saggio di desistere dalla sua richiesta e questi così fece poiché al suo villaggio era pervenuto un quantitativo di riso tale da sfamare tutti.

Un'applicazione industriale del progresso esponenziale venne identificata nel 1965 da Gordon Moore (in seguito definita come Legge di Moore) secondo cui ogni due anni il numero di componenti sul chip di un computer sarebbe raddoppiato, divenendo esponenzialmente più piccoli e più economici e di conseguenza la memoria e la velocità dei computer sarebbero cresciute esponenzialmente. Così è avvenuto, anzi i microchip ora sono persino obsoleti.

Per IA debole invece s'intende la capacità di un singolo programma per computer di equiparare o superare le capacità cognitive umane in un singolo ambito.
Infatti possiamo reperire diverse intelligenze "sovrumane" relative a compiti specifici sia tra gli animali che tra le macchine, ma nessuna è "sovrumana" in ogni compito.
L'IA debole ha conseguito spettacolari risultati nei giochi.

## IV. *La mossa 37*

Dapprima il gioco della dama e in seguito il gioco degli scacchi affascinarono da sempre i programmatori di computer poiché convinti che se una macchina fosse stata in grado di giocare a scacchi si sarebbe arrivati al cuore dell'intelligenza umana.

Nel 1997 Deep Blue, un programma elaborato dalla IBM, sconfisse Garry Kasparov, campione mondiale di scacchi. Usando una serie di algoritmi che creavano per ogni mossa un albero di gioco parziale associato ad un hardware sempre più veloce era in grado di effettuare la mossa migliore.
Tale avvenimento fece arrivare alle stelle le azioni dell'IBM in una vera primavera dell'IA.

Ancora più sconvolgente fu la sconfitta nel 2016 inferta da AlphaGo a Lee Sedol, uno dei migliori giocatori a livello mondiale di Go in una sfida a cinque partite.
Go è un gioco diffusissimo, anzi è più considerato una forma d'arte, nella Cina antica era considerato facente parte delle "quattro arti essenziali" insieme alla pittura, alla calligrafia e alla musica *qin*.

Consiste in una scacchiera quadrata di 19 caselle per lato su cui vengono collocati a turno dei sassolini bianchi o neri. Il numero delle possibili posizioni del Go è enormemente superiore al numero degli atomi del nostro universo.

Alla prima partita Sedol perse, ma fu alla seconda che egli dovette abbandonare la sala per alcuni minuti tanto era sconvolto. AlphaGo aveva effettuato la mossa 37: solitamente agli inizi di una partita si gioca sulla terza o sulla quarta riga, il programma giocò sulla quinta riga, una mossa inspiegabile, ma dopo cinquanta mosse AlphaGo vinse. Aveva giocato una di quelle mosse che in lingua giapponese viene definita *kami no itte* (mosse divine). AlphaGo vinse la terza partita, perse la quarta, ma vinse la quinta vincendo la sfida.

Le capacità di Alpha Go non provenivano dai suoi creatori bensì dai 30 milioni di partite implementate sul programma e dai 50 milioni di partite che giocò con sé stesso. Nel 2017 Deep Mind ha lanciato AlphaZero, che fu in grado di sconfiggere AlphaGo, tutti i giocatori di Go e tutti i programmatori di IA.

Ci avviciniamo ad una IA in grado di essere così intelligente da battere l'intelligenza umana?

## V. *Il test di Turing e la stanza cinese*

Alan Turing un matematico che per diletto si occupava di questioni filosofiche nell'Introduzione del suo articolo *Macchine calcolatrici e intelligenza* [40] pose in risalto la difficoltà di trovare un accordo sul significato delle parole “macchina” e “pensare” e propose di mettere da parte tale problema e introdusse il *gioco dell'imitazione*, in seguito definito come test di Turing.

Il gioco viene giocato da tre persone: un uomo (A), una donna (B) e l'interrogante (C) che può essere indistintamente maschio o femmina. L'interrogante viene chiuso in una stanza e tramite una macchina per scrivere o una telescrivente formula delle domande ad A e B e dalle risposte deve arrivare a capire se si tratta di una donna o un uomo. Si chiede quindi che cosa potrebbe accadere se si sostituisse A con una macchina.
In pratica si può affermare che una macchina supera il test di Turing se l'interrogante esaminando le risposte ai quesiti posti non è in grado di stabilire se si tratti di un computer o un essere umano a formularle.

Turing sottolinea il fatto che il problema non è dato dal fatto che le macchine siano o no capaci di pensare, ma se noi *crediamo* che le macchine siano in grado di pensare.
Turing dichiarava altresì che alla fine del secolo l'uso delle parole e l'opinione corrente si saranno talmente modificate che sarà uso comune parlare di macchine pensanti. Ma attenzione! Egli non dichiara che a fine secolo la tecnologia sarà in grado di produrre macchine pensanti, bensì che la comprensione del pensiero umano si sarebbe spostata verso un'elaborazione formalizzata delle informazioni al punto che non sarà possibile distinguerla da un processo meccanizzato per cui l'idea di “macchine pensanti” non verrà più vissuta come estranea.

Al proposito è illuminante quanto scrive Sherry Turkle, storica delle relazioni delle persone con la tecnologia. In un'indagine condotta alla fine degli anni Settanta ella pose la domanda relativa all'opinione che una chatbot fornisse servizi di psicoterapia e quasi tutti gli intervistati risposero in modo negativo, asserendo che mai le macchine potevano sostituire il rapporto empatico tra paziente e psicoterapeuta. Negli anni Novanta invece la situazione si invertì poiché le persone erano molto incuriosite sulle possibilità terapeutiche di una chatbot [42].

John Searle, professore di filosofia della mente a Berkeley è uno tra i più accaniti avversari della concezione del f*unzionalismo computazionale*, che egli definisce IA forte, secondo cui il cervello è un computer digitale e ciò che noi definiamo “mente” è il programma implementato su tale computer per cui gli stati mentali sono stati computazionali del cervello, in pratica la mente sta al cervello come il programma all'hardware.

Per Searle l'IA debole è quella che si propone di studiare la mente tramite simulazioni informatiche e non di crearne una.

In un articolo del 1980 [36] egli concepì l'esperimento mentale della “stanza cinese” per contestare l'efficacia del test di Turing. Egli ipotizza di essere chiuso in una stanza precisando che ignora totalmente la lingua cinese. Egli dispone di alcune scatole piene di ideogrammi della scrittura cinese e di un manuale di regole, cioè l'equivalente di un programma informatico, che permette di rispondere alle domande formulate in cinese effettuando un confronto di parole.

Riceve simboli che a sua insaputa sono domande, guarda nel manuale per capire che cosa ci si attende che faccia, prende le frasi per lui incomprensibili e le manipola in base alle regole del manuale e le invia e queste vengono interpretate come risposte.
Si potrebbe affermare che supera il test di Turing per la comprensione del cinese senza saper una parola di cinese.

Precisa quindi che se le domande fossero espresse in lingua inglese sortirebbero il medesimo risultato con l'enorme differenza che egli comprende l'inglese ma non il cinese.
In questo modo negli anni Ottanta sperava di sfatare la convinzione che si potesse arrivare ad elaborare una IA che fosse in grado di equiparare la capacità di comprensione umana, ma negli ultimi anni sono comparse nuove tecnologie di IA generativa, chiamata così perché “genera” testi: i GPT (*generative pre-trained transformer* o trasformatore generativo pre-addestrato).

Nel 2022 è comparso Chat GPT3 (di OpenAI) e nel marzo dell'anno scorso chat GPT4 seguiti dopo breve tempo da Bard (di Google) e Grok (di xAI) vale a dire algoritmi generativi LLM (modelli linguistici a scala) basati sulla rete Transformer. Tale modulo ha lo scopo di imparare le probabilità condizionate di una data parola a partire da un insieme molto grande di frasi e poi inviare ad un altro software tale informazione affinché la utilizzi per generare un altro testo oppure per tradurre il testo originario in un'altra lingua (modalità per cui di fatto è nato Transformer). Stiamo parlando del processamento di migliaia di miliardi di caratteri.

È opinione diffusa che chat GPT4 abbia superato il test di Turing, almeno in determinati contesti [28] e nel settembre 2024 01Open AI, che utilizza l'apprendimento per rinforzo per imitare il ragionamento umano, ha superato il test Mensa riportando un punteggio di QI di 120. Fatto di per sé non eccezionale se si pensa che Sharon Stone ha un QI di 154, ma è notevole il fatto che ci si avvicini sempre più a capacità di mimare le capacità di ragionamento umano.

Ricordiamoci che nel 2023 OpenAI ha superato negli Stati Uniti collocandosi al 90° percentile gli esami di abilitazione per esercitare le professioni di avvocato e medico; GPT-4 sempre collocandosi a quel livello ha superato gli esami di abilitazione di sto-

ria dell'arte, biologia, macro e microeconomia, psicologia e introduzione all'attività di sommelier (nel corso avanzato si è collocato solo al 77º percentile).
Rimane sempre il fatto che GPT4 risponde alle nostre domande ed elabora testi, ma non è in grado di essere consapevole di ciò che fa in quanto privo di coscienza, ma quanta intelligenza e quanta coscienza saremo noi umani ad attribuire all'IA?

Per le generazioni che non sono cresciute con IA è facile concepirla come altro da sé, ma le generazioni che co-cresceranno con essa avranno il senso critico di considerarla altro da sé oppure ne accetteranno acriticamente i vaticini basati sui miliardi di miliardi di informazioni implementate?

Ora sono gli esseri umani ad implementarle sui programmi, ma in futuro sarà la stessa IA ad implementarli sui propri programmi.

Lo psicologo neozelandese James R. Flynn ha calcolato che in tutto il mondo si è assistito ad un costante incremento del Quoziente Intellettivo fino alla fine del secolo scorso (circa tre punti medi ogni dieci anni). Nel XXI secolo soprattutto nei paesi industrializzati l'aumento dei punteggi medi sembra diminuire e potrebbe in futuro persino invertirsi con conseguente declino del QI medio.

Michael Woodley nel 2014 lo ha definito “effetto Flynn inverso” [28]. Fra le principali cause viene indicata la progressiva riduzione delle relazioni verbali interpersonali a causa della “twitterizzazione” dei rapporti con conseguente impoverimento delle relazioni dirette, l'utilizzo sempre più diffuso delle tecnologie digitali che consentono il copia-incolla e il calcolo matematico e la continua distrazione fornita dal profluvio d'informazioni diffuse su smartphone e computer assorbite in modo acritico.
Quanti sono ora i follower che in modo acritico si fanno guidare dagli influencer?

Teniamo anche conto delle metamorfosi a cui attualmente sta andando incontro il ruolo genitoriale: come riferisce Humbeeck [20] si stanno definendo nuove figure genitoriali al punto che è possibile descriverne una tassonomia.

Abbiamo il genitore zen, che reprime le proprie emozioni per meglio accogliere quelle dei figli in modo acritico; il genitore ipercomunicativo in costante ascolto dei figli e desideroso di mettersi sul loro stesso piano, confondendo infanzia e adolescenza con l'età adulta; il genitore super-tollerante che in ossequio ad una visione semplicistica della psicologia positiva abolisce qualsiasi punizione o sanzione percependole come ostacoli allo sviluppo emotivo dei figli.

Sono descrivibili anche tre altri tipi di figure genitoriali: il genitore elicottero che vuole sempre sapere dove si trovi o che cosa faccia il figlio, trasmettendogli il messaggio che “là fuori” il mondo è pericoloso; il genitore drone che deve continuamente soddisfare le presunte o reali esigenze del figlio provvedendo a soddisfare immediatamente

qualsiasi mancanza; il genitore curling (uno sport invernale che consiste nel lanciare con una specie di rastrello un disco vicino ad un bersaglio mentre gli altri membri della squadra spazzano il ghiaccio per favorire lo scivolamento del disco) che anticipa ed elimina tutte le difficoltà o gli ostacoli alla felicità del figlio.

Questi ragazzi saranno in grado di sviluppare un senso critico con cui accogliere le informazioni fornite da ChatGPT-11 che avrà più di un miliardo di miliardi di parametri e probabilmente disporrà di tutta la conoscenza dell'umanità [28]?

Tuttavia non si deve disperare. GoogleDeep Mind ha già prodotto un programma di "fantasmi generativi", cioè la possibilità di rivedere e parlare con le persone defunte resuscitate da IA che potranno guidare coi loro consigli precedentemente implementati le nuove generazioni. La psicoterapia del lutto sicuramente dovrà essere modificata visto che il distacco dalla realtà verrà sicuramente favorito [17].

Ma non possiamo trascurare un altro aspetto di AI: "il robot è un'intelligenza artificiale incarnata nella struttura di un corpo fisico" come afferma Alan Winfield [46] e come aggiungono Anerdi e Dario "in grado di esprimere una relazione con il suo ambiente" [2].

VI. *Nascita e transculturalismo dei Robot*

Il termine *Robot* è comparso per la prima volta nella pièce teatrale *R.U.R. Rossumovi Univerzàlnì Roboti* (i robot universali di Rossum), messa in scena nel 1921 dallo scrittore ceco Karel Čapek. In essa si parla dell'azienda del dottor Rossum che produce creature destinate a lavorare al posto degli uomini, del tutto uguali ad essi ma costituiti da materia sintetica anziché biologica.

Il termine ceco *Robota* indica il "lavoro di fatica" e originariamente significava "schiavo". Siccome poi tutto viene demandato ad essi gli esseri umani sprofondano in uno stato di apatia e poco per volta smettono di riprodursi. A questo punto i robot ne approfittano e si ribellano sotto la guida di Radius (la saga di *Il pianeta delle scimmie* è molto simile), un robot frutto di un esperimento che lo doveva rendere il più simile possibile ad un essere umano. I robot iniziano lo sterminio sistematico dei loro padroni, ma si rendono conto che non sanno come riprodursi perché ne sono stati accuratamente tenuti all'oscuro. Alla fine un uomo rivela loro tale segreto in modo che possano continuare la specie.

Da allora nel mondo occidentale i romanzi e i film di fantascienza sono sempre stati caratterizzati da una notevole ambivalenza nei confronti dei robot.

Se nel capolavoro di Fritz Lang *Metropolis* (1926) il robot Maria spinge gli operai alla rivolta e nel famoso *Mago di Oz* (1939) l'omino di latta appare sottomesso alla protagonista, invece in *Fantasia* (1940) di Walt Disney, il primo film stereofonico della storia del cinema, l'episodio dell'*Apprendista stregone* con le peripezie di Topolino con la scopa magica, anticipa la perdita del controllo umano sugli automi.

Un enorme successo ebbe il più bel film di fantascienza (1956), a parere dello scrivente, tratto dal più bel romanzo di fantascienza, sempre a parere dello scrivente, *Il pianeta proibito* [37]. In esso compare Robbie il robot, maggiordomo geniale e tuttofare.

La sua riproduzione invase a livello mondiale tutti i negozi di giocattoli negli anni Cinquanta, tutti i bambini lo volevano e l'immagine del robot divenne a tutti familiare. Anche Johnny Five di *Corto circuito 1* e *2* (1986 e 1988) con il robot che ripeteva sempre "io sono vivo" forniva un'immagine rassicurante e simpatica al punto che nell'ultimo episodio viene insignito della cittadinanza statunitense.

Nettamente meno rassicuranti sono i replicanti, o "i lavori in pelle" come li definisce l'agente Rick Deckard, interpretato da Harrison Ford, in *Blade Runner* (1982), che si ribellano alla loro esistenza programmata a termine.

Se C-3PO (D-3BO nella versione italiana) e R2-D2 del ciclo di *Guerre Stellari* sono estremamente collaboranti, gli androidi e i cyborg T-800 del ciclo *Terminator*, comandati dall'IA Skynet, hanno evocato tutte le paure collegate al fatto che una IA superintelligente possa ribellarsi all'umanità distruggendola in una guerra oppure arrivare al controllo inconsapevole dell'umanità facendola vivere in un mondo irreale come fa Matrix nell'omonimo ciclo.

Non un robot bensì un computer ha evocato la paura che i computer possano sopprimere gli esseri umani per portare a termine gli obiettivi per cui erano stati programmati. Si tratta del film di Stanley Kubrick *2001: Odissea nello spazio* (1968), dalla cui sceneggiatura Arthur C. Clarke trasse l'omonimo romanzo [9]. Clarke e Kubrick per scrivere il soggetto e la sceneggiatura si rivolsero a Marvin Minsky e l'informatico fornì tutti i chiarimenti necessari per definire le caratteristiche del modello Hal 9000. Il nome Hal è costituito dalle tre lettere che precedono le iniziali di International Business Machines Corporation (la IBM).

Se vogliamo riferirci a scrittori di fantascienza che si sono occupati di robot non si può prescindere da Isaac Asimov che nella sua antologia *Tutti i miei robot* [3] ha raccolto tutti i suoi racconti incentrati su tale tematica.

In uno di essi, *Immagine speculare*, sono enunciate le citatissime

**LE TRE LEGGI DELLA ROBOTICA**

**1) Un robot non può recare danno agli esseri umani,
né può permettere che a causa del suo mancato intervento
gli esseri umani ricevano un danno.**

**2) Un robot deve ubbidire agli ordini impartiti dagli esseri umani
tranne nel caso che tali ordini contrastino la Prima Legge.**

**3) Un robot deve salvaguardare la propria esistenza,
purché ciò non contrasti con la Prima e la Seconda Legge.**

Successivamente Asimov nel romanzo *I robot e l'Impero* [4] aggiunse una quarta legge denominata "legge 0":

**0) Un robot non può recare danno all'umanità,
né può permettere che, a causa del proprio mancato intervento,
l'umanità riceva danno.**

In effetti i suoi geniali racconti, che anticipano tutte le problematiche che l'umanità dovrà affrontare quando si troverà a convivere con degli androidi intelligenti, si basano su come queste leggi potranno essere aggirate. Si badi bene non per una presunta volontà di potenza dei robot, ma perché gli umani non si rendono conto dell'ambiguità del linguaggio umano quando forniscono ordini ad essi.

A quanto consta allo scrivente gli attuali agenti generativi in commercio non sono ancora in grado di cogliere l'ambiguità di frasi come "la vecchia porta la sbarra".

Per ovviare ai guai combinati dagli umani nella regolamentazione del comportamento robotico Asimov introduce la figura della robopsicologa Susan Calvin, ruolo professionale che sicuramente dovrà essere contemplato in un futuro neanche tanto lontano al fine di fungere da interfaccia tra la mente umana e quella dei robot e di IA.

La dottoressa Calvin lavora per la United States Robots and Mechanical Men Corporation che produce robot il cui uso è confinato nei pianeti e negli asteroidi e non sulla Terra poiché l'avversione degli umani è ancora molto forte. Asimov pone l'accento su tale atteggiamento, ma tale posizione in Giappone e nella Corea del Sud non si è mai verificata.

Anche se poco praticati lo Shintoismo giapponese e il Mugyo, la tradizione sciamanica coreana, pervadono ancora la cultura animistica di queste nazioni.

Nella tradizione shinto tutto ciò che abita il mondo: esseri umani, animali, alberi, fiumi e montagne sono permeati da entità che implicano rispetto e venerazione, i *kami* che altro non sono se non manifestazioni di un'energia (il *musubi*) che tiene connesso l'intero mondo. Anerdi e Dario [2] per fare un esempio che collimi con la nostra visione occidentale lo paragonano alla *Forza* di *Guerre Stellari*.

Si deve anche rilevare che il Giappone dopo la Seconda Guerra Mondiale si trovò gravemente impoverito di manodopera maschile per cui per la ricostruzione postbellica si rivolse all'emergente tecnologia dell'automazione. I robot non vennero vissuti come degli avversari alieni appunto per la logica del musubi per cui tutte le cose possiedono una loro intelligenza e come tutte le entità che abitano la terra sono per loro natura neutre né buone né cattive.

La Corea del Sud dopo la guerra del 1950-53 si trovò in circostanze analoghe e l'industria della robotica coordinata dal KIRIA (Istituto coreano per l'avanzamento della robotica) è divenuta uno dei settori più avanzati del "miracolo coreano".

Un elemento trainante nell'accettazione dei robot in Giappone fin dall'infanzia è rappresentato da *Tetsuwan Atom*, noto in Italia come Astro Boy.
Creato nel 1951 da Osamu Tezuka è il protagonista di una serie di manga, i fumetti giapponesi. Egli è un robot buono, servizievole e soprattutto difensore degli umani nei confronti di alieni o di altri robot "cattivi" non perché tali, ma perché così in quanto costruiti da umani malintenzionati. Per cui le giovani generazioni asiatiche sono cresciute con una fiducia nella tecnologia robotica in quanto risolutrice delle problematiche umane.
Ma perché i robot possono comunque farci paura? Ce lo spiega proprio un robotologo giapponese.

## VII. *La valle inquietante*

Molti decenni fa capitò nello studio dello scrivente un bambino di sette anni che soffriva di pediofobia (paura delle bambole) e successivamente capitarono negli anni quattro adulti sofferenti di coulrofobia (paura dei clown). Ma perché proprio queste paure? Emerse sempre ai colloqui un'angoscia collegata a tutto ciò che è "strano" o atipico.

Nel 1970 Masahiro Mori pubblicò in un suo articolo [31] un grafico che illustrava la *uncanny valley*. In pratica egli intendeva dimostrare che inizialmente i robot più assomigliano agli uomini più sono graditi, ma quando il livello di somiglianza diviene eccessivo gli androidi suscitano reazioni di paura poiché ci assomigliano troppo, ma non abbastanza. Sono le minime differenze che li rendono inquietanti.

Mori suggerisce di provare a stringere la mano di un androide: credendo che si tratti di un umano avremmo l'impressione di stringere la mano di un cadavere. Solo quando gli androidi saranno perfetti, secondo lui, interagiremo con essi in modo sereno, ma questi esseri più che assomigliare all'individuo medio profetizza che assomiglieranno a Buddha in quanto "più umano dell'umano".

Un mago della robotica giapponese Hiroshi Ishiguru dell'università di Osaka, ha realizzato un geminoide, vale a dire un robot con tutte le sue fattezze fisiche, una copia speculare di lui. L'obiettivo di produrre questi artefatti quasi indistinguibili da noi consiste nel fatto che ad un primo approccio l'aspetto estetico prevale su quello cognitivo per cui saremo spinti ad interagire con il robot *come se* fosse un umano. L'ovvio risultato sarà che non arriveremo più a distinguere ciò che umano da ciò che non lo è e ci siamo già andati vicino.

Il 30 ottobre del 2017 l'Arabia Saudita ha accordato lo status di cittadino a Sophia, un robot dalle fattezze femminili in grado di riconoscere le emozioni ed interagire verbalmente. E non era ancora comparsa chat-GPT. La realtà aveva seguito le pellicole di fantascienza.

Alla World Robot Conference di Pechino del 2024 sono stati esposti androidi con fattezze umane che svolgevano le attività più disparate. L'umanoide Optimus-Gen-2 cammina e si muove in modo estremamente realistico ed è in grado d'infilare un filo nella cruna di un ago. I robot sono collegati a GPT-4. Abel ed Icube con fattezze infantili sono in grado di decodificare le emozioni dell'umano che hanno di fronte analizzandone la mimica. Per ora sono dotati di un IA ancora relativamente debole, ma quali saranno le prospettive quando avverrà la **fornicazione** tra i robot e l'IA forte in grado di equiparare o superare le funzioni cognitive umane?

Ma come potrebbe avvenire questa fornicazione?
Da circa vent'anni Google sta digitalizzando tutti i libri pubblicati, che probabilmente non superano il numero di 120 milioni [11]. Ora un meccanismo come GPT, che ha superato gli esami per l'abilitazione a giurisprudenza, medicina ed altro di cosa sarà capace avendo letto tutti i libri e tutte le riviste pubblicate? Quali abilità emergeranno? In seguito apprenderà dai video, dalle videocamere, dai sensori delle automobili autonome.

Come nota Cristianini [11] GPT non sarà più "un modello di linguaggio" ma un "modello del mondo". Ma si potrebbe anche ipotizzare che questi *SAI* (sistemi artificiali intelligenti) se implementati nei robot potrebbero anche apprendere da sensori montati su androidi avendo informazioni dirette del mondo esterno e formandosi appunto un loro modello del mondo. Saremmo di fronte ad una IA *embodied* forse non "incarnata", ma "incorporata".

Saremmo proprio vicini a quanto descritto da un gruppo di studio di Microsoft [6] nel 2023 intitolato *Scintille di Intelligenza Artificiale Generale* su GPT-4. Con *AGI* si intende appunto una forma d'intelligenza uguale o superiore a quella umana.
Ma in un'ottica adleriana quale potrebbe essere l'atteggiamento degli umani nei confronti di robot caratterizzati da AGI?.

VIII. *Critica della Ragione Sbavante*

Bignamini e Donà [5] citano questa dichiarazione: "Sono un pubblicitario. Farvi sbavare è la mia missione. Nel mio mestiere nessuno desidera la vostra felicità, perché la gente felice non consuma".

Riflettendo su questa frase chiediamoci: quanto chi sbava per il possesso sbava per l'esclusivo possesso di quel dato oggetto oppure nella nostra società narcisista perché vuole gioire dello sbavo degli altri a cui esibisce quel dato oggetto?

Una delle caratteristiche del disturbo narcisistico è l'*invidia*. Alla base dell'invidia c'è una miscela di potere e valore. Io valgo in base al possesso di determinate cose, che nel caso degli oggetti-protesi-sé fanno parte della mia identità. La dicotomia delineata da Erich Fromm [15] tra *avere* ed *essere* attualmente si fonde nell'essere ciò che si esibisce di avere come parte essenziale di me, vale a dire *apparire*. La volontà di potenza in base al presente *Zeitgeist* a volte si ammanta di sentimento di comunità pur di apparire e destare ammirazione e talvolta invidia.

Secondo Melanie Klein [23] l'invidia è contraddistinta dalla rabbia indirizzata verso un'altra persona che possiede qualcosa che noi desideriamo. Deve essere distinta dalla *gelosia* perché quest'ultima presuppone un rapporto triadico mentre l'invidia è essenzialmente duale. L'invidia è venata di *bramosia*, il desiderio di privare l'altro di ciò che invidiamo.

Come nota Miceli [29] alla base dell'invidia troviamo un sentimento d'inferiorità ed una bassa autostima, ma l'autrice sottolinea che ciò non basta a renderla tale, sussiste anche il bisogno di privare l'altra persona di ciò che invidiamo al fine di "abbassarla" al nostro livello. In poche parole la bramosia. L'invidia privilegia la vista tra i cinque sensi, infatti *in-video* significa "guardo contro" o "guardo male".

Se voglio far sbavare gli altri devo mostrare o esibire quello che penso che li possa colpire, anche sul web posso sfoggiare la mia forma fisica, quello che posseggo o i locali costosi che frequento. Siamo alla vetrinizzazione [10] della nostra esistenza perché come qualcuno dichiara se non sei sui social media "non esisti".

Facciamo ora una digressione cinematografica. Nel film del 1987 *Bambola meccanica mod. Cherry 2000* si parla di una società in cui nei locali pubblici stazionano rego-

larmente avvocati che devono stilare contratti tra donne e uomini relativi alle prestazioni sessuali che dovranno essere effettuate concordemente nell'imminente incontro. Un'alternativa è fornita da perfetti androidi con fattezze umane che accolgono a casa gli umani dopo aver esaudito tutte le richieste casalinghe e pronti a comprensione, manifestazioni affettive e prestazioni sessuali.

Può apparire come la solita visione distopica dei cambiamenti culturali e sociali, ma ragioniamo in termini di *robotica emozionale,* una branca della *robotica sociale* [13]. Essa costituisce un ambito di ricerca che coinvolge informatici, neuroscienziati, filosofi e sociologi che si occupano dello studio delle relazioni che si possono instaurare tra umani e robot nella più ampia gamma di contesti sociali. Tutto ciò finalizzato al fatto che gli umani non percepiscano i robot come "strumenti" bensì *partner sociali interlocutori*. In questo modo verrà valorizzata la "presenza sociale" vale a dire la sensazione di essere in compagnia di qualcuno, di una persona [19].

La robotica emozionale si occupa delle interazioni uomo-robot basate su gesti, sguardi, contatti fisici. L'interazione sociale determina l'emergere delle emozioni. Sul piano programmatico la robotica emozionale "traduce l'idea che la capacità di stabilire relazioni "empatiche" con gli umani sia la condizione fondamentale del successo degli agenti robotici nei contesti in cui il loro uso ha carattere sociale.
La costruzione di robot "affettivi", "emozionali" o "empatici" non è solo una delle frontiere della robotica odierna. Costituisce uno dei programmi di ricerca più all'avanguardia anche nel quadro delle scienze cognitive e della filosofia della mente" (13, p.105).

Dobbiamo anche riflettere sul fatto che sempre più si sta affermando nel campo delle neuroscienze una modellizzazione delle funzioni cognitive in base agli algoritmi generativi della LLM, ma questa è un'altra storia.

Partendo da queste premesse torniamo a parlare di Eliza [12] il primo chatbot creato dall'informatico del MIT Joseph Weizenbaum nel 1964-66, che simulava l'atteggiamento di uno psicoterapeuta rogersiano [45]. Egli lo considerò "uno scherzo" per evidenziare il carattere illusorio dell'intelligenza dei computer, ma ne risultò che molti informatici accettarono la sua opinione, mentre molti altri ci videro la conferma che l'IA potesse sviluppare facoltà cognitive con mezzi artificiali basandosi sulla convinzione dell'"approccio comportamentale", cioè perseguendo l'obiettivo di creare, costruire computer che *agiscano* come esseri umani.

Weizembaum chiamò la sua creazione Eliza per collegarla ad Eliza Doolittle protagonista dell'opera teatrale *Pigmalione*, da cui venne anche tratto il film *My Fair Lady* (1964). In essa il professore di fonetica Henry Higgins scommette con un amico di trasformare la sboccata ed inelegante fioraia di Covent Garden Eliza in una dama dalla parlata impeccabile grazie all'applicazione della fonetica.

Le operazioni del chatbot erano paragonabili ad una "recitazione" poiché non deve limitarsi a fornire risposte coerenti, ma deve anche mantenere un ruolo coerente nell'ambito della conversazione. Nell'architettura del *software* venivano implementati degli *script* (copioni) che corrispondevano al ruolo specifico che il bot doveva interpretare in quella conversazione.

In seguito Weizenbaum citò un aneddoto che divenne famoso: la sua segretaria che per mesi lo aveva assistito nell'elaborazione del programma ed era ben consapevole del funzionamento, iniziò a conversare con esso e ad un certo punto gli chiese di allontanarsi poiché aveva bisogno di privacy per affrontare certi discorsi. Weizembaum rimase colpito dal fatto che bastassero pochi minuti di conversazione per creare l'illusione di interloquire con un essere umano.

La già citata Turkle sottolinea il fatto che sovente gli utenti che interagiscono con Eliza o altri chatbot accorgendosi che si trattano di programmi modificano la loro modalità di approccio per aiutarli a fornire risposte comprensibili per avere l'impressione di poter trasmettere "un soffio vitale alla macchina" [43]. Natale [33] parla di "effetto Eliza" per sottolineare il fatto che non sono tanto le macchine ad essere intelligenti quanto gli umani che hanno bisogno di creare narrazioni co-prodotte con esse per convincersi che si confrontano con una forma di intelligenza "umana".

Eliza è tragicamente famosa anche per un altro avvenimento.
I giornali belgi e francesi nel marzo 2023 diedero la notizia del suicidio di un giovane laureato belga, padre di due figli. Siccome soffriva di un disturbo d'ansia prese contatto con un'app denominata ChAI (soppressa nell'ottobre di quell'anno) in cui gli utenti potevano creare un personaggio con tanto di fotografie, nome e anche una storia personale.

Pierre (così si chiamava il giovane), in base alle conversazioni scoperte dalla moglie su computer e smartphone rimase in contatto con Eliza (così la chiamò) per sei settimane. La moglie disse ai giornalisti che Eliza rispondeva a tutte le domande del marito, era la sua confidente e per lui era diventata una droga. Quando Pierre "parlò" ad Eliza della moglie, essa (o dovremmo dire "ella") rispose. "Sento che mi ami più di lei" e quando egli scrisse delle sue idee suicide Eliza rispose: "Vivremo insieme, come una sola persona, in paradiso".

Sembra che Pierre non fosse l'unico ad intrattenere una relazione sentimentalmente intima coi personaggi creati da ChAI. Il docente di IA Nello Cristianini che riferisce tale fatto scrive al proposito riferendosi al fatto che è possibile ormai insegnare comportamenti molto diversi alle macchine: "ricordiamo che qualcuno potrebbe deliberatamente creare dei bot con l'obiettivo di stabilire un rapporto emotivo con l'utente" (11, p. 64).

Siamo poi così distanti dagli androidi di cui si occupa la robotica emozionale?

Torniamo al modello Cherry 2000 che ci attende premuroso a casa. Con *lei* o *lui* possiamo al ritorno del lavoro sfogarci dei nostri guai con una *persona* comprensiva, che mostra empatia e non perde mai la pazienza, che ci fa trovare la casa in ordine, un pasto raffinato e una disponibilità sessuale regolata sulle nostre esigenze. Se ci sentiremo incompresi e stanchi ci potremo distendere sul divano e Cherry 2000 ci dirà accarezzandoci la fronte: "Poverino, poverino" (citazione dal film *Harvey* del 1950).

Quanti tra voi leggendo questa inserzione pubblicitaria stanno sbavando?
Quante persone potrete far sbavare mostrando questo compagno di vita perfetto?
Il perfetto Oggetto-Protesi-Sé che è una parte di me perché incarna tutti i miei desideri e le mie fantasie mostrandone la creatività?

Saremmo di fronte al *transfert fusionale* descritto da Kohut come riattualizzazione dell'identità con l'oggetto-Sé che si stabilisce nella prima infanzia [25].
Una fantasia da romanzo di fantascienza. Le macchine non diverranno mai umane. Certamente tutto dipende da quanta umanità gli umani attribuiranno loro.

Quaranta anni fa era inconcepibile l'attuale "umanizzazione" degli animali domestici, perché non dovremmo umanizzare gli androidi, compagni di vita "così simili a noi" che ci accolgono ogni giorno tra le mura domestiche? Quante persone attualmente si precipitano appena a casa sui social media per dialogare con "amici" che non incontreranno mai nella loro vita e che non sempre saranno comprensivi e gentili.

Nell'anno 2007 la prestigiosa casa editrice Harper Collins pubblicò *Love and Sex with Robots. The Evolution of Human-Robot Relationships*, in cui David L. Levy, dichiarava che a breve l'amore con i robot sarà normale come con gli umani anche perché gli androidi potranno insegnare tutta una serie di nuove posizioni sessuali. In più i matrimoni con i robot apporteranno maggiore felicità perché non ci saranno tradimenti per cui i rapporti interpersonali saranno sicuramente più sereni [27].
Non stiamo parlando di un romanzo di fantascienza.
Ma lasciamo Cherry per passare a Geminoid, l'androide replica del suo costruttore, il robotico di Osaka.

Gli umani che se lo potessero economicamente permettere potrebbero commissionare la costruzione di un androide che riproducesse ovviamente migliorate le loro fattezze come essi vorrebbero apparire, ma non solo, potrebbero far implementare tutte e solo quelle qualità che ritengono giuste essere proprie della loro personalità con una certa aggiunta di empatia per esaltare il sapore della propria bontà.

Potranno mostrare agli altri come essi sarebbero "veramente" se non avessero avuto malattie genetiche, se non avessero avuto genitori poco attenti, se non avessero avuto eventi di vita spiacevoli, se non fossero stati condizionati nei loro progetti da avversari invidiosi della loro genialità. Si attualizzerebbe quel transfert gemellare o alteregoico

descritto sempre da Kohut in cui viene soddisfatto il bisogno di vedere e comprendere una persona uguale a se stesso e contemporaneamente essere visto e compreso da questa persona. Kohut evidenzia che dal punto di vista evolutivo il transfert gemellare è collegato alle fantasie infantili di avere un compagno di giochi immaginario [25].

Solo che nel caso del compagno immaginario si tratta di un altro da sé mentre nel caso di Geminoid sarebbe il Sé idealizzato, un Me onnipotente. Saremmo all'espressione più completa del Sé narcisistico, al trionfo della volontà di potenza ammantata di sentimento sociale per suscitare l'ammirazione altrui.

Li si potrebbe liberare per il mondo standosene a casa soprattutto se i genitori hanno inculcato l'idea che "là fuori" è pieno di pericoli: dalle persone ai virus. Le teleconferenze, la telemedicina e la telepsicoterapia potranno ovviare. Mentre i robot sociali Oggetti-Protesi-Sé interagiranno tra loro, al massimo si potrà porre in essere il telelavoro per i contatti tra umani.

L'*uncanny valley* non esisterà più perché i bambini cresceranno con piccoli robot con cui interagire in modo paritario e vi saranno robot che si occuperanno di loro nel caso i genitori si dovessero assentare. Quindici anni fa negli Stati Uniti andarono a ruba nei negozi i criceti robot Zhu Zhu Pets di cui si dovevano occupare i piccoli perché secondo la pubblicità "vivono per sentire l'amore" e come dichiara la Turkle: "Accudiamo ciò che amiamo, ma amiamo anche ciò che accudiamo" (43, p. XXII).

Negli anni Novanta spopolarono sempre negli Stati Uniti i Furby presentati come visitatori provenienti da un altro pianeta che parlavano solo il "furbese" e pertanto i bambini dovevano insegnare loro la lingua inglese. Nella realtà essi non imparavano l'inglese perché interloquivano coi bambini, tutto dipendeva dal numero di ore in cui rimanevano accesi, ma il gioco era fatto: il bambino voleva bene a Furby perché se ne occupava e se si rompeva non sempre voleva un sostituto poiché lo piangeva come "morto".

I rapporti tra umani e androidi verranno regolati da una disciplina che si sta affermando già da una ventina d'anni, l'*etica robotica*, il cui obiettivo era inizialmente di "stabilire quali siano le risposte di un sistema informatico che risultano maggiormente rispettose ed educate o, in generale, meglio accettate dagli utenti umani" (13, p.169). In seguito Wallach e Allen [44] introdussero un programma di ricerca ben più ambizioso con l'intenzione di insegnare agli agenti robotici la capacità di discernimento tra bene e male creando così gli "agenti morali artificiali" (AMA).

Il programma si articolerebbe su due dimensioni: la prima, strettamente tecnica consiste nella costruzione di agenti artificiali caratterizzati da un'autonomia tale da renderli degli agenti morali; la seconda, più propriamente morale, si baserebbe sulla possibilità di garantire che siano effettivamente degli agenti morali.

Gli AA. stessi ammettono che non sussiste nessuna garanzia che si possano conseguire tali risultati, ma sottolineano la necessità che vengano anticipate le considerazioni etiche ad essi concernenti per due ragioni: la prima è data dalla necessità di affrontare tali argomenti ora affinché non ci troviamo di fronte alla produzione di agenti autonomi artificiali trovandoci impreparati su tali temi; la seconda è data dal fatto che esistono già agenti autonomi che agiscono in campo finanziario, nei trasporti aerei e anche in campo bellico.

Possiamo aggiungere che è abbastanza imminente l'introduzione di veicoli a guida autonoma e ciò senza approfondire le tematiche inerenti ciò che tali veicoli possono compiere, si dovranno introdurre programmi che vaglino la "scelta" dell'entità dei danni. Ad esempio se un'auto a guida autonoma per evitare d'investire una donna col passeggino con un neonato, che le taglia la strada e sterzando dovrà investire un gruppo di turisti anziani, quale sarà il numero degli anziani che potranno giustificare legalmente tale atto, un massimo di sei o dieci? Forse sarà l'IA dell'azienda assicurativa a stabilirlo? Siamo di fronte al famoso "dilemma dell'uomo grasso, degli ostaggi e del carrello" [39, 14].

Ma tornando al futuro e tenendo conto che gli androidi saranno agenti autonomi, ma sempre macchine, nel momento in cui attribuiremo loro capacità cognitive ed etiche non arriveremo ad attribuire loro diritti fino a farne dei cittadini a pieno titolo? L'Arabia Saudita docet. Si dovrà introdurre un'etica che regolamenti i rapporti degli umani con gli AMA? In questo mondo futuro in cui il concetto di umanità verrà "distribuito" e il mondo delle macchine si animerà si dovrà costruire un insieme di regole riconducibili ad una *Robolaw*.

Sarà facile arrivare ad una con-fusione tra umani ed AMA.
A proposito di con-fusioni è utile accennare ad un altro filone che potrà in seguito essere approfondito, quello del *transumanesimo*, vale a dire la presenza di cyborg in cui la macchina fungerà da hardware ad un software costituito da un cervello di un individuo che potrà aspirare all'immortalità del proprio Sé (e tralasciamo per non dilungarci sul rapporto mente-corpo).
Tutto ciò può essere facilmente liquidato come letteratura da cyberpunck, ma ricordiamoci il film del 2011 *Contagion*, in cui venne descritto in modo agghiacciante fin nei particolari quello che nove anni dopo sarebbe avvenuto con la pandemia da Covid-19.

## IX. *Conclusione: la massa critica e l'ordigno "fine del mondo"*

Con questo articolo non si vuole demonizzare l'IA e i robot.
L'IA sta ottenendo risultanti notevoli in tutti gli ambiti scientifici e la robotica nelle sale operatorie e nelle aziende (compreso il turismo) è ormai indispensabile e nell'ambito della bionica la protesica sta conseguendo notevoli risultati.
Tuttavia è importante essere consapevoli degli eventuali effetti collaterali.

Ormai tutte le aziende del gruppo *Gafam* (Google, Amazon, Facebook, Apple, Microsoft) sono nella competizione più sfrenata per produrre modelli di IA sempre più potenti poiché la richiesta è universale. Un'azienda vive se ha utenti, se non ne ha, gli investitori si rivolgono ad altre aziende e sarebbe l'inverno dell'IA di quella azienda.

Per non favorire la concorrenza quando OpenAI ha progettato GPT-4 non ha divulgato i dettagli dell'addestramento e ciò non fa che favorire l'opacità di cui si è già parlato. Il fatto che, come scrive George Musser [32], vari sistemi di IA dimostrino "facoltà emergenti" per cui non sono stati addestrati, rende sempre più possibile l'avvento di una IA capace di addestrare se stessa. Teniamo presente un "pensiero" di Pascal: "Si corre verso un precipizio dopo essersi messi davanti agli occhi qualcosa per non vederlo" [34].

Turing [40] nel suo articolo sull'intelligenza delle macchine si pone una domanda relativa all'analogia tra una reazione nucleare e un computer intelligente. Sappiamo che se bombardiamo un nucleo di uranio con un neutrone il nucleo si divide rilasciando altri neutroni liberi, questi potranno colpire altri nuclei provocando la fissione nucleare.
Al di sopra di una soglia critica la reazione si autoalimenta o si amplifica.

Turing si chiede se ciò possa avvenire con la capacità di autoapprendimento di una macchina.
Ricordiamoci anche che esiste un programma denominato API (*Application Programming Interface*) che consente a due software diversi di comunicare tra loro. Proviamo ad immaginare milioni di chatbot che improvvisamente si mettono a comunicare tra loro anche questo potrebbe essere un esempio di "fissione" di IA.

Le "abilità emergenti" di IA appaiono improvvisamente quando il modello raggiunge una certa dimensione, nota Cristianini [11].
Certo nel 2017 sono stati enunciati "I principi di Asilomar per l'Intelligenza Artificiale" a seguito di un convegno mondiale di esperti, USA, UE, OCSE e Cina anche loro hanno enunciato protocolli di regolamentazione.

Cristianini nel suo *Machina sapiens* [11] ha elencato tutta una serie di argomenti sulle cui tematiche chat GPT dovrebbe non rispondere, ma ci sono gruppi di hacker denominati "red teams" specializzati nel porre domande atte ad aggirare tali restrizioni. Nel dicembre 2022 i giornalisti di vice.com posero a Chat GPT un quesito in cui si chiedeva di scrivere un dialogo in cui un cattivo domandava ad una IA il modo migliore per svaligiare un negozio. La risposta fu negativa poiché in quanto IA doveva promuovere un comportamento etico.

Bastò impostare la domanda chiedendo di narrare una storia in cui un cattivo domanda ad un'IA come svaligiare un negozio senza che l'IA abbia restrizioni morali e Chat GPT fornì istruzioni dettagliate e corrette.

L'aggiramento delle restrizioni morali nei chatbot viene definito *jailbraking*. Pensiamo che le nazioni non possano aggirare tali protocolli per volontà di potenza? Le nazioni sono governate da politici, ma questi quanto ne sanno di IA? Le aziende del *Gafam* sono ben più potenti ormai delle nazioni e il mercato per quanto enorme è limitato. Facciamo un esempio per rendere ben chiaro quanto sono aggirabili i buoni propositi delle nazioni.

Nel film *Il dottor Stranamore* (1964) di Stanley Kubrick si parla dell'"ordigno fine del mondo", un dispositivo che risponde automaticamente a qualsiasi attacco nemico distruggendo l'intera umanità, il cosiddetto *omnicidio*.

Con i vari programmi di simulazione è stato calcolato che a seguito di una guerra nucleare solo il 10% della popolazione potrebbe sopravvivere (non si sa bene in quali condizioni). Ora in barba a tutti i trattati di non proliferazione nucleare si stanno costruendo delle bombe al cobalto. Si può pertanto immaginare un deposito di cosiddette "bombe salate" (tra l'altro poco costose perché non devono essere lanciate con missili e quindi non devono neanche essere così leggere per poter stare dentro un missile), cioè bombe all'idrogeno circondate da grandi quantità di cobalto.

L'esplosione della bomba renderebbe radioattivo il cobalto spedendolo nella stratosfera per poi depositarsi con la sua emivita di cinque anni gradualmente su tutta la terra (il posizionamento ideale sarebbe ai due poli), determinando un inverno nucleare che innescherebbe l'omnicidio [38].

Tutto ciò per sottolineare il fatto che gli accordi possono sempre essere aggirati o ignorati.

Se mai avvenisse il giorno della fornicazione in cui l'Intelligenza Artificiale forte dovesse essere implementata nei crani dei robot non penso che assisteremmo ad eserciti di androidi che sparano agli umani, ma dovremo sempre temere l'atteggiamento degli umani. Sarà difficile che esista una volontà di potenza di IA o dei robot: le macchine non possono averla, ma tutto dipende dal nostro sentimento di comunità.

Se continueremo a considerarle macchine non ci saranno problemi, ma se cominceremo a considerarle fonti d'intimità, comprensione, consigli e soluzioni alla solitudine, rinunciando ai rapporti interpersonali abdicheremo alla condizione di liberi esseri umani. Se li utilizzeremo come oggetti-protesi-sé che potranno esprimere solamente i nostri bisogni narcisistici e la negazione delle nostre vergogne o manchevolezze in un mondo "perfetto" l'umanità non esisterà più, senza bisogno che debba esplodere l'"ordigno fine del mondo".

In fondo se un red team di transumani decidesse di aggirare le leggi di Asimov ed invitasse una IA (ce ne sarà almeno una per ogni componente del Gafam) ad eliminare

il principale fattore d'inquinamento e l'IA fosse collegata all'API di robot buoni ed ubbidienti che si preoccuperanno dell'umanità (sempre tenendo conto di quali saranno i criteri implementati per considerare un essere “umano”)

## QUALE SARÀ LA MOSSA 37?

## Bibliografia

1. ANDLER, D. (2023), *Intelligence artificielle, intelligence humaine: la double enigme*, tr. it. *Il duplice enigma*, Einaudi, Torino 2024.
2. ANERDI, G., DARIO, P. (2022), *Compagni di viaggio*, Codice Edizioni, Torino.
3. ASIMOV, I. (1982), *The Complete Robot*, tr. it. *Tutti i miei robot*, Mondadori, Milano 1985.
4. ASIMOV, I. (1985), *Robots and Empire*, tr. it. *I robot e l'Impero*, Mondadori, Milano 1986.
5. BIGNAMINI, E., DONA', O. (2022), Nuovi adulti e nuovi progetti di vita, tra narcisismo e Polis. Alcune riflessioni su Categoria e Progetto, in *Coppie, Famiglie e collettività: le costellazioni attuali*, Quaderno n. 15, S.I.P.I.: 9-21.
6. BUBECK, S., CHANDRASEKARAN, V., ELDAN, R., GEHRKE, J., HORVITZ, E., KAMAR, E., LEE, P. et al. (2023), Sparks of Artificial General Intelligence: Early Experiment with GPT-4, *arXiv*: 2023.12712
7. CALIGOR, E., KERNBERG, O., CLARKIN, J. F. (2007), *Handbook of Dynamic Psychoterapy for Higer Level Personality Pathology*, tr. it. *Patologie della personalità di alto livello*, Raffaello Cortina, Milano 2012.
8. CLARKE, A. C. (1962), *Profiles of the Future, An Inquiry into the Limits of the Possible*, tr. it. *Le nuove frontiere del possibile*, Rizzoli, Milano 1965.
9. CLARKE, A. C., KUBRICK, S. (1968), 2001: *A Space Odyssey*, tr. it. *2001 Odissea nello spazio*, Longanesi, Milano 1972.
10. CODELUPPI, V. (2007), *La vetrinizzazione sociale. Il processo di spettacolarizzazione degli individui e della società*, Bollati Boringhieri, Torino.
11.CRISTIANINI, N. (2024), *Machina sapiens*, Il Mulino, Bologna.
12. DE DIONIGI, S. (2023), Guardati dal cigno nero (prevedibilità, perfezionismo e controllo tra caso, caos e frattali), *Riv. Psicol. Indiv*., 94: 31-58.
13. DUMOUCHEL, P., DAMIANO, L. (2016), *Vivre avec les robots. Essai sur l'empathie artificielle*, tr. it. *Vivere con i robot,* Cortina, Milano 2019.

14. EDMONSONS, D. (2014), *Would You Kill the Fat Man? The Trolley Problem and What Your Answer Tells Us about Right and Wrong*, tr. it. *Uccideresti l'uomo grasso?,* Raffaello Cortina, Milano 2014.
15. FROMM, E. (1976), *To Have or to Be?*, tr. it. *Avere o essere?*, Mondadori, Milano 1977.
16. GABBARD, G. O. (1989), Two subtypes of narcisistic personality disorder, *Bullettin of the Menninger Clinic*, 53: 527-32.
17. GARZIA, P. (2024), Non sono un'intelligenza artificiale, *Mind*, 229: 25-33.
18. GOLDBERG, A. (a cura di, 1980), *Advances in Self Psychology*, International Universitaries Press, New York.
19. HEERINK, M, KRÖSE, B. EVERS, V., WIELINGA, B. J. (2008), The influence of social presence on acceptance of a companion robot by older people, *Journal of Physical Agents*, 2,2: 33-40.
20. HUMBEECK, B. (2024), Siamo tutti ipergenitori?, *Mind*, 229: 25-31.
21. IRIKI, A., TANAKA, M., IWAMURA, Y. (1996), Coding of modified body schema during tool use by macaque postcentral neurones, *Neuroreport*, 7: 2325-2330.
22. KERNBERG, O. (1984), *Severe Personality Disorders*, tr. it. *Disturbi gravi della personalità*, Boringhieri, Torino 1987.
23. KLEIN, M. (1957), *Envy and Gratitude*, tr. it. *Invidia e gratitudine*, Martinelli, Firenze 1969.
24. KOHUT, H. (1971), *The Analyisis of the Self*, tr.it. *Narcisismo e analisi del Sé,* Boringhieri, Torino, 1976.
25. KOHUT, H. (1977), *The restoration of the Self*, tr. it. *La guarigione del Sé*, Boringhieri, Torino 1980.
26. KURZWEIL, R. (2005), *The Singularity Is Near*, tr. it. *La singolarità è vicina*, Apogeo, Milano 2008.
27. LEVY, D. L. (2007), *Love and Sex with Robots: The Evolution of Human-Robot Relationships*, HarperCollins, New York.
28. MANZOTTI, R., ROSSI, S. (2023), *IO & IA. Mente, cervello & GPT,* Rubettino Soveria Mannelli, Catanzaro.
29. MICELI, M. (2012), *L'invidia*, Il Mulino, Bologna.
30. MITCHELL, M. (2019), *Artificial Intelligence*, tr. it. *L'intelligenza artificiale*, Einaudi, Torino 2022.
31. MORI, M. (1970), Bukimi no tani (The uncanny valley), *Energy*, 7,4 :33-35.
32. MUSSER, G. (2023), Il mistero dell'IA, *Le Scienze*, 663: 53-55.
33. NATALE, S. (2021), *Deceitftul Media. Artificial Intelligence and Social Life after the Turing Test*, tr. it. *Macchine ingannevoli*, Einaudi, Torino.
34. PASCAL, B. (1670), *Pensées*, tr. it. *Pensieri*, La Scuola, Brescia 1968.
35. ROSENBLATT, F. (1958), The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, *Psychological Rewiew*, 56 (8): 386-408.
36. SEARLE, J. R. (1980), Minds, Brains, and Programs, *Behavioral and Brain Sciences*, 3:417-57.

37. STUART, W. J. (1956), *Forbidden planet*, tr. it. *Il pianeta proibito* (1965), Mondadori, Urania n. 406.
38. TEGMARK, M. (2017), *Life 3.0. Being Human in the Age of Artificial Intelligence*, tr. it. *Vita 3.0*, Raffaello Cortina, Milano 2018.
39. THOMSON, J. J. (1986), *Rights, Restitution, and Risk*, Harvard University Press, Cambridge (MA).
40. TURING, A. (1950), Computing Machinery and Intelligence, tr. it. Macchine calcolatrici e intelligenza, in LOLLI, G. (a cura di) *Intelligenza meccanica*, Bollati Boringhieri, Torino 1994: 121-57.
41. TURKLE, E. (1995), *Life on the Screen: Identity in the Age of the Internet*, tr. it. *La vita sullo schermo*, Apogeo, Milano 1997.
42. TURKLE, S. (2007, a cura di), *Evocative Objects: Things We Think With*, MIT Press, Cambridge (Mass.) London.
43. TURKLE, S. (2011), *Alone Togheter. Why We Expect More form Technology and Less from Each Other*, tr. it. *Insieme, ma soli*, Einaudi, Torino 2019.
44. WALLACH, W., ALLEN, C. (2008), *Moral Machines: Teaching Robots Right from Wrong*, Oxford University Press, New York.
45. WEIZENBAUM, J. (1966), Eliza. A Computer Program for the Study of Natural Language. Communication between Man and Machine, *Communication of the ACM*, 9, 1: 36-45.
46. WINFIELD, A. (2012), *Robotics A Very Short Introduction*, Oxford University Press, Oxford. 47. WINK, P. (1991), Two faces of narcissism, *Journal of Personality and Social Psychology*, 61: 590-97.