

EXORDIUM

CORPI AL TEMPO DEL POSTUMANO.
L'EDUCAZIONE DELLE MACCHINE NUOVA
FRONTIERA DELLA PEDAGOGIA?

BODIES AT THE TIME OF POST-HUMAN.
MACHINE EDUCATION THE NEW FRONTIER
OF PEDAGOGY?

*Umberto Zona (Università di Roma Tre),
Martina De Castro (Università di Roma Tre)**

Il corpo, nel suo incessante processo di mutazione, riflette la complessità delle dinamiche che oggi concorrono a definire il concetto di identità, non più circoscrivibile nella logica binaria fondata sulle varianti di sesso e genere ma destinato ad assumere una dimensione fluida e poliforme. Pur se con difficoltà, la cultura occidentale inizia a recepire questo mutamento di orizzonte, che ha evidenti ripercussioni sul piano dei comportamenti sociali, dei rapporti e delle relazioni umane, mentre riserva scarsissima attenzione alle trasformazioni tecnologiche del corpo, avviato a emanciparsi dalle catene della carne e a ibridarsi sempre più profondamente con le macchine. Eppure, i cyborg, attraverso una intricata rete di innesti, protesi meccaniche e elettroniche, sono da tempo fra noi mentre, grazie ai progressi continui registrati nel campo dell'intelligenza artificiale e dei nuovi materiali, gli androidi rappresentano ormai il futuro prossimo della condizione postumana. Il saggio si prefigge di analizzare gli sviluppi dell'interazione uomo/macchina riflettendo in particolare sulle sue ripercussioni nel campo dell'identità

* Il presente lavoro è il frutto delle riflessioni e del lavoro congiunto degli autori. Tuttavia, ai fini dell'attribuzione delle singole parti, si specifica che i paragrafi 1 e 4 sono a opera di Martina De Castro e i paragrafi 2 e 3 a opera di Umberto Zona.

e sugli interrogativi etici e filosofici che essa introduce. L'educazione delle macchine è dunque la nuova frontiera della pedagogia?

Body, in its incessant process of mutation, reflects the complexity of the dynamics that today contribute to defining the concept of identity, which can no longer be circumscribed within the binary logic based on sex and gender variants, but is destined to assume a fluid and multiform dimension. Although with difficulty, western culture is beginning to acknowledge this change in horizon, which has obvious repercussions on the level of social behaviour and human relations, while it dedicates very little attention to the technological transformations of the body, which is set to emancipate itself from the chains of the flesh and to hybridise itself ever more profoundly with machines. Yet cyborgs, through an intricate network of grafts, mechanical and electronic prostheses, have long been among us, while, thanks to the continuous progress made in the field of artificial intelligence and new materials, androids now represent the near future of the post-human condition. The essay sets out to analyse developments in human/machine interaction, reflecting in particular on its repercussions in the field of identity and the ethical and philosophical questions it introduces. Therefore, is machine education new frontier of pedagogy.

1. Introduzione

Quello del rapporto tra reale e virtuale è un tema sul quale già Morin, nella prefazione a *Lo spirito del tempo* (2006), si interrogava: «È possibile che gli esseri umani siano ormai manipolati dalle macchine infernali che essi stessi hanno creato? Si tratta di un tema basilare, angoscioso, assolutamente delicato, che attiene alla nostra vita quotidiana» (p. 19). Per quanto ci riguarda, spinti probabilmente da quello stesso senso di urgenza che segnalava Morin, ci siamo interrogati più volte circa i residui margini di autonomia decisionale conservati dall'uomo nei confronti della macchina, e

siamo giunti alla conclusione che il rapporto tra macchine e umani non può essere letto esclusivamente come conflitto fra corporeità e tecnologia. Sia perché tale conflitto, come suggerito da molti autori (Caronia, 2007; Clark, 2003; Haraway, 1991), è destinato a risolversi in un'ibridazione irreversibile fra umano e post-umano, ma soprattutto perché i dispositivi digitali rappresentano ormai un'estensione esterna della nostra memoria, un prolungamento dei nostri arti superiori, e ci consentono di proiettare le nostre identità all'interno di agorà virtuali globali. Dunque, una fitta trama connettiva fa sì che la conoscenza non abbia più come sede esclusiva la mente individuale, ma sia distribuita fra il corpo, l'ambiente sociale e tutti quei dispositivi cognitivi, linguistici e culturali di cui l'uomo quotidianamente si serve e che ormai, per larghissima parte, operano in Rete. Questa convinzione ci deriva dalla definizione di *artefatto culturale* che Vygotskij (1987) ci ha fornito:

[t]utti gli strumenti artificiali, tutta la situazione culturale servono “alla amplificazione dei nostri sensi”, e l'uomo civile moderno può permettersi di avere le peggiori proprietà della natura, arricchendole con congegni artificiali e con il loro aiuto far fronte al mondo esterno in modo migliore di quanto lo facesse il primitivo, utilizzando direttamente le proprie doti naturali. [...]. Nel processo evolutivo l'uomo ha inventato strumenti e ha creato un ambiente produttivo civile; ma questo ambiente produttivo ha trasformato lui stesso, ha prodotto, al posto delle forme primitive di comportamento, forme complesse, culturali (pp. 166-167).

La macchina, in tal senso, appare in Vygotskij ancora un ibrido tra un prodotto storico/sociale – creato grazie alle intuizioni di più soggetti e trasformatosi e perfezionatosi con il passare del tempo e con il mutare delle esigenze – e uno strumento in grado di amplificare facoltà e potenzialità individuali. Questo tipo di lettura si ritrova in molti studiosi della comunicazione, a partire da Marshall McLuhan (1964; 1981), che guarda alle macchine elettroniche come estensioni del corpo, per finire ad Andy Clark (2003; 2011), che considera i dispositivi digitali degli artefatti in grado di estendere le capacità della mente umana. Se fino a qualche tempo fa,

però, questo potenziamento delle facoltà avveniva attraverso congegni in grado di agire all'esterno dell'organismo umano, i progressi registrati negli ultimi decenni nel campo dell'elettronica, della robotica e, più in generale, della tecnologia stanno accelerando, di fatto, il meticciamiento tra i nostri corpi di carne e le macchine – si pensi, ad esempio, a tutte e tutti coloro che vivono in omeostasi con un pacemaker o che hanno scelto di farsi inserire chip sottocutanei. Attraverso una intricata rete di innesti, protesi meccaniche e elettroniche, perciò, ciascuno dei nostri corpi sta – neanche troppo lentamente – migrando verso la dimensione cyborg:

[n]el modello del cyborg il corpo non è né fisico né meccanico, e non è nemmeno puramente testuale. Il cyborg è piuttosto un controparadigma che descrive l'intersezione del corpo con una realtà esterna molteplice e complessa: è una lettura moderna non solo del corpo, non solo delle macchine, ma di quello che passa e succede tra di loro (Haraway, 2018, p. 30).

In quest'ottica, riuscire a delimitare il confine tra ciò che è naturale e ciò che è artificiale/culturale diviene sempre più complicato o, per meglio dire, superfluo. E il corpo, nel suo incessante processo di mutazione e ibridazione, riflette la complessità delle dinamiche che oggi concorrono a definire il concetto di identità, non più circoscrivibile in una logica binaria fondata sulle varianti di sesso e genere ma destinato ad assumere una dimensione fluida e poliforme. Le identità si fanno nomadiche, performative, radicali, non più frutto di logiche naturali, ma tendenti a rappresentarsi sul palcoscenico della vita secondo ruoli e copioni imposti dalle dinamiche di potere:

[m]aschile/femminile e maschio/femmina sono le categorie funzionali a celare il fatto che le differenze sociali dipendono sempre da un ordine economico, politico e ideologico. [...]. Ma il punto è che non esiste alcun sesso. Esistono solo un sesso oppresso e un sesso oppressore. Ed è l'oppressione a creare il sesso; non il contrario (Wittig, 2019, p. 22).

In sostanza, tutto ciò che ancora consideriamo come naturale – il sesso così come l’eterosessualità – sarebbe secondo alcune studiose un insieme di caratteri culturali oppressivi, in base ai quali si riproducono leggi generali che hanno la presunzione di potersi adattare a tutti i soggetti, in qualsiasi società ed epoca. Da Donna Haraway (2018) con il suo costrutto di cyborg a Judith Butler e la sua performance di genere (2017), il femminismo moderno sembra aver recepito queste istanze provenienti dal basso e pare sempre più interessato agli spazi di rivendicazione aperti dall’ibridazione, dagli eccessi corporei e dalla devianza alle norme naturali. In questo senso, uno scenario interessante all’interno del quale elaborare le nostre riflessioni ci pare quello offertoci dagli androidi, esseri extra-corporei ed extra-umani.

2. Maestri o ammaestratori?

Nel corso degli ultimi mesi, due eventi hanno portato il dibattito sull’intelligenza artificiale fuori dai confini della comunità scientifica, destando l’interesse dei media generalisti e la curiosità dell’opinione pubblica. Il primo risale al dicembre del 2021, quando la società di robotica *Engineered Arts*¹ ha rilasciato un video in cui si vede l’androide Ameca interagire con degli umani con grande naturalezza. Fin qui nulla di particolarmente sconvolgente: già da alcuni anni, le aziende impegnate nella ricerca sull’A.I. (*Artificial Intelligence*) hanno presentato prodotti robotici sempre più sofisticati, a partire dall’androide sociale Sophia, progettato dalla *Hanson Robotics* e presentato nel marzo del 2016 al *South by Southwest Festival* (Grasso, 13 ottobre 2016). Anche in quel caso i media riservarono all’evento ampio spazio, Sophia fu accolta nei maggiori talk show e, sull’onda del clamore suscitato dalle sue brillanti apparizioni televisive, l’11 ottobre 2017 fu invitata alle Nazioni Unite, dove ebbe un breve “scambio di vedute” con il vicesegretario generale, Amina J. Mohammed (UN News, 2017), e qualche giorno dopo, il 28 ottobre, le fu concessa

¹ Disponibile in: <https://www.engineeredarts.co.uk/> [10/09/2022].

addirittura la cittadinanza da parte dell'Arabia Saudita (Morosi, 2017).

Tornando ad Ameca, dunque, la ragione della risonanza mediatica scaturita dalla sua presentazione non risiede tanto nella perfezione tecnica dell'artefatto. Certo, l'androide della *Engineered Arts* può vantare un'eccellente mimica facciale, che riproduce in maniera molto accurata (pur se limitata a un set determinato di espressioni) quella umana. Anche la gestica è fluida, così come i movimenti oculari, ma, dicevamo, non sono tanto i traguardi tecnologici raggiunti dai progettisti ad aver attivato l'interesse generale quanto, piuttosto, una reazione "inaspettata" che Ameca ha manifestato durante la presentazione al CES, la fiera tecnologica di Las Vegas: quando, infatti, un membro dello staff ha ripetutamente avvicinato il proprio dito al volto dell'androide, fino a toccargli il naso, questi l'ha afferrato e respinto con modi risoluti, accompagnati da uno sguardo affatto rassicurante². Il team della *Engineered Arts* assicura di essere rimasto spiazzato e, in qualche modo, intimorito, dalla reazione di Ameca (Corriere Tv, 2021), ma pensiamo sia lecito nutrire qualche dubbio sulla genuinità di tali esternazioni, non fosse per il fatto che l'atto di "disobbedienza" di Ameca ha regalato all'azienda produttrice grande notorietà in tutto il mondo. Va detto che la comunità scientifica è rimasta sostanzialmente indifferente alla performance di Ameca, così come era accaduto per Sophia, ma ciò non ha dissuaso i media dal conferire grande visibilità all'evento, segno evidente che una certa narrazione pubblica del rapporto uomo-macchina, imperniata sul conflitto fra umano e artificiale, è dura a morire e ha generato un copione sempre uguale, dove l'atto della creazione, sottratto alle leggi della natura e affidato all'onnipotenza della scienza, genera il "mostro". In questo plot narrativo, l'uomo è prigioniero della sua volontà di potenza, si ingegna per creare macchine a sua immagine e somiglianza, fatte per servirlo docili e devote, pretende di entrare in un rapporto empatico con loro e di possederle, salvo poi scontrarsi con la loro disob-

² Di seguito il link per vedere la reazione di Ameca: https://www.youtube.com/watch?v=RiTfe-ckD_g [10/09/2022].

bedienza. Nella vulgata mediale sull'intelligenza artificiale, in considerazione del ruolo predominante ricoperto dall'addestramento, la ribellione dell'androide viene rappresentata come una sorta di fallimento educativo, quando invece, in virtù del lignaggio inferiore attribuito alla macchina, essa andrebbe probabilmente letta alla stregua di una ferita narcisistica inferta a un ammaestratore, costretto a misurarsi con le bizzarrie e le ritrosie di un animale. Skinner (2005), che riservava all'ammaestramento, di cui era straordinariamente esperto, il massimo apprezzamento, faceva confessare a Frazier, il suo alter ego in *Walden Two*:

[a] dominare i miei primi esperimenti era un desiderio frenetico ed egoistico. Mi ricordo la rabbia che provavo quando una previsione andava male. Avrei potuto gridare ai soggetti delle mie sperimentazioni: "Dannazione, comportatevi come dovrete!" (p. 47).

Dal suo punto di vista, Skinner – che considerava la libertà un lusso perfino per il genere umano (Skinner, 2002) – aveva tutte le ragioni per dolersi delle intemperanze di topi e piccioni: l'ammaestramento, in effetti, si fonda su un rapporto di potere, sul principio in base al quale il più forte – l'ammaestratore – può avere libero accesso al corpo del più debole, sia esso un animale o una macchina o tutto ciò che si muove ai confini dell'umano. Ameca, dunque, avrebbe procurato a Skinner una crisi di sconforto? Siamo convinti del contrario. A tale riguardo, ipotizziamo – anche se su un piano puramente speculativo – tre chiavi di lettura dell'atto di disobbedienza di Ameca.

2.1. Animali meccanici

Quando l'androide si trova a fronteggiare il dito del ricercatore sembra adottare una condotta tipicamente animale, propria in particolare dei mammiferi, che comunicano attraverso una complessa rete di segnali acustici, tattili, visivi e olfattivi, attraverso i quali intercettano altri esseri viventi o oggetti presenti nell'ambiente e, quando ricevono segnali rassicuranti sul loro conto, azzardano tal-

volta comportamenti esplorativi, che sembrano dettati dalla curiosità. In realtà non si tratta di curiosità intellettuale vera e propria, perché l'animale non è in grado di discernere la natura dei soggetti, degli elementi naturali o degli artefatti in cui si imbatte, quanto, piuttosto, di una manifestazione di quel Sé biologico che sarebbe caratteristico di tutti gli organismi viventi e sarebbe finalizzato a preservarne l'omeostasi. Se si osserva bene la sequenza diffusa dalla *Engineered Arts*, Ameca segue con lo sguardo il dito che si avvicina al suo viso e, quando questo sta per toccare il suo naso, si protende verso il dito, quasi volesse annusarlo, come può fare un animale che cerca di identificare l'odore di un suo simile o di testare le proprietà di un elemento sconosciuto che ha attirato la sua attenzione. È solo quando il ricercatore viola senza preavviso l'ultimo diaframma e stabilisce unilateralmente un contatto tra i corpi, che Ameca arma le sue difese e allontana l'intruso. La reazione di Ameca, dunque, può essere equiparata a quella di un animale? E, in quanto tale, rappresenterebbe un primo gradino evolutivo verso l'acquisizione di processi psichici più complessi? In realtà, quando un animale si spaventa mette in campo due strategie: fugge dalla minaccia, allontanandosi o, più raramente, attacca la fonte del pericolo. L'androide, invece, non appare spaventato né assume comportamenti ostili verso l'essere umano: semplicemente lo blocca, facendogli intendere che la sua sortita non è gradita. Dunque, Ameca sembrerebbe più che altro intenzionata a difendere il proprio habitat, il che ci suggerisce la seconda ipotesi interpretativa.

2.2. Il contatto come confine regolatorio

L'antropologo statunitense Edward T. Hall (1982), nel suo *La dimensione nascosta*, individua quattro zone della comunicazione interpersonale: una zona intima (0-45 centimetri), una personale (45-120 cm), una sociale (120-350 cm) e una pubblica (oltre i 350 cm). Queste zone sono presidiate da rigide prescrizioni culturali (che talvolta danno luogo a veri e propri divieti e sanzioni), espressione delle consuetudini, delle usanze e delle pratiche sociali in vigore presso una data comunità. Comunemente si ascrivono alla zona

intima quelle prassi dettate da un forte rapporto di confidenza, familiarità, dimestichezza, come quelle affettive, sessuali o che potremmo definire “cameratesche”. Il presupposto di tali pratiche è, in genere, il *contatto* fra i corpi e una profonda conoscenza reciproca fra gli attori della relazione. Per la verità, queste due ultime condizioni non sono strettamente necessarie: si pensi al caso di un semplice conoscente che pure penetri nella nostra zona intima per sussurrarci qualcosa all’orecchio, sfiorando soltanto la nostra epidermide, o di uno sconosciuto che ci abbracci nel corso di una celebrazione o di un festeggiamento collettivo. In questi casi non proviamo disagio perché ci viene in soccorso una motivazione utilitaristica (nel caso del sussurrante) o culturale (nel caso dell’abbraccio occasionale) di tali comportamenti. Quando, invece, non ci diamo spiegazioni per l’inaspettata violazione del nostro spazio intimo, viviamo questa intromissione indebita come un’aggressione o come uno stato di malessere al quale sottrarsi il prima possibile (si pensi quando, seduti gli uni accanto agli altri in un mezzo pubblico, ci affrettiamo ad occupare il posto più isolato lasciato vacante dai passeggeri scesi). Il confine della zona intima, pertanto, è l’accettazione del corpo dell’altro. Nel caso di Ameca, autentica o meno che sia la sua reazione, non c’è dubbio che abbia vissuto la sortita dell’umano nel suo spazio intimo come un’intromissione indebita. E in tal senso si dimostra perfettamente “umana”, poiché ha interiorizzato i precetti della prosemica, la cui regola aurea è, sintetizzando brutalmente, “che ogni corpo stia al suo posto”. Del resto, la cultura dominante mette in guardia dal contatto, sia che esso significhi meticciamiento, contaminazione, ibridazione sia che esso sia la risultante occasionale di un incontro fra corpi. Se questo è vero, le zone individuate da Hall hanno una funzione prettamente (auto)regolatoria e un portato normativo implicito che Ameca ha interiorizzato e messo in pratica, esattamente come un essere umano. Dunque, la difesa del proprio spazio personale da parte dell’androide potrebbe essere letta da una prospettiva diametralmente opposta a quella propinata all’opinione pubblica dai fautori dell’A.I. ed essere considerata, invece che il primo passo verso l’indipendenza umanoide, la conferma eclatante di quanto ancora le

macchine – meglio, i loro comportamenti – ricalchino e si ispirino alla natura umana.

2.3. *Umano, troppo umano*

A proposito dell'umanizzazione dell'androide è d'obbligo rispolverare un quesito che ci accompagna fin dalla pubblicazione del capolavoro di Asimov (1963), *Io robot*: le macchine possono, in prospettiva, provare emozioni? E, se così fosse, è possibile che tali emozioni possano tradursi in comportamenti ostili nei confronti degli umani? Non abbiamo citato Asimov a caso, in quanto le sue *Leggi della robotica* sono tra le radici della ricerca sull'intelligenza artificiale, almeno per quanto concerne il piano etico. Esse, come è noto, recitano:

1. Un robot non può recar danno a un essere umano né può permettere che, a causa del suo mancato intervento, un essere umano riceva danno;
2. Un robot deve obbedire agli ordini impartiti dagli esseri umani, purché tali ordini non vadano in contrasto con la prima legge;
3. Un robot deve proteggere la propria esistenza, purché la salvaguardia di essa non contrasti con la Prima o con la Seconda Legge (Asimov, 1963, p. 111).

Se dovessimo leggere il gesto di Ameca alla luce di tali leggi, dovremmo fare appello alla terza e interpretare quel movimento di difesa come un prodromo di un embrionale senso del Sé, come un primordiale frammento di consapevolezza sulla strada della costituzione di una soggettività extra-umana. Ameca, infatti, più che proteggersi sembrerebbe tendere a stabilire un confine, una linea di demarcazione ontogenetica: l'arto meccanico che afferra e blocca quello umano non sarebbe, pertanto, una manifestazione ostile ma una rivendicazione di autonomia, la manifestazione aurorale di una specie che si è emancipata ed esige rispetto. Ed è esattamente questo ciò che turba l'immaginario della nostra esausta umanità: il timore di essere soppiantata dalla propria prole macchi-

nica. E poiché il rapporto che storicamente ha legato uomo e macchina è quello del servaggio (ce lo ricordano le leggi di Asimov), una volta che l'androide abbia reciso il cordone ombelicale con il suo creatore, tutto lascerebbe presumere che tale evento possa coincidere con l'estinzione del genere umano o con la sua riduzione in cattività. Viene in mente un'altra legge – questa volta marxiana – secondo la quale la borghesia è divenuta la classe egemone creando un'altra classe – il proletariato – che l'avrebbe poi soppiantata (Marx & Engels, 1944). Questa ipotesi appare però scongiurata se si pone in primo piano il rapporto di apprendimento/insegnamento che lega tuttora macchine e umani e se si dissoda la natura *culturale* delle leggi sulla prossemica individuate da Hall (1982). Ameca, come un umano, respinge quell'appendice cartilaginea che si protende verso il suo viso perché nel codice culturale delle nostre società non è contemplata la possibilità che un corpo possa avvicinarsi a un altro senza rispettare un complesso sistema di regole e ruoli. Ovviamente stiamo ragionando sull'eventualità di un "contatto esplorativo", che non arrechi alcun danno a chi ne è destinatario, che non violi la sua volontà e non sia espressione di un rapporto asimmetrico tra i soggetti. E proprio a questa altezza torna in campo la questione delle emozioni, perché il presupposto per cui chi viene toccato ritenga spiacevole il contatto è che esso susciti in lui non soltanto sensazioni sgradevoli ma emozioni negative: paura, rabbia, disgusto ecc. Non è un caso, del resto, che la retorica sull'A.I. insista più sulla capacità emotiva delle macchine che sulle loro competenze cognitive. Nell'immaginario collettivo, infatti, il provare emozioni è la porta principale per accedere alla condizione umana e da sempre, in ambito letterario, gli artefatti tecnologici antropomorfi giungono a misurarsi con il mondo dei sentimenti, che spesso risulta per loro esiziale. Nel romanzo di Mary Shelley (1975), l'essere mostruoso resuscitato dal dottor Frankenstein svela a poco a poco la sua umanità dolente e imbocca una parabola (auto)distruttiva nel momento in cui chiede al suo creatore una compagna. Invece, il robot Daneel Olivaw, uno dei protagonisti di *Abissi d'acciaio* di Asimov (1995), grazie alla sua amicizia con il detective Bailey, acquista progressivamente sentimenti umani; in

Blade Runner (Ridley Scott, 1982), l'androide Roy Batty si ribella e uccide il suo creatore perché questi lo ha progettato affinché viva pochi anni, nel timore che possa sviluppare col tempo la capacità di provare emozioni. Potremmo continuare pressoché all'infinito nel citare esempi letterari e cinematografici di androidi e umanoidi che si cimentano col mondo delle emozioni umane, cagionando gioie e sofferenze a se stessi o ai propri creatori.

Ora, Ameca non prova alcuna emozione perché non dispone dei requisiti per poterlo fare. Le ricerche più recenti, infatti, attestano che la caratteristica fondamentale delle emozioni è il loro essere complessi processi psico/fisiologici, a conferma dell'idea spinoziana secondo la quale l'esperienza emotiva si sostanzia di mente e corpo insieme (Ammaniti & Gallese, 2014; Rizzolatti & Sinigaglia, 2006). In particolare, il grande filosofo olandese indica col termine *affectio* lo stato di un corpo quando esso è soggetto all'azione di un altro corpo (Spinoza, 2015). Questo meccanismo implica un *contatto* che Ameca, essendo incorporeo, non può trasformare in emozione: la funzione del corpo come mediatore emozionale gli è necessariamente sconosciuta perché non ha un'epidermide che possa sperimentarla³.

Dunque, a parere di molti studiosi (Noble, 2018; UNESCO, 2019), le aziende dell'A.I. starebbero enfatizzando, con un abile gioco di suggestioni, le prestazioni degli androidi per corroborare l'idea che essi si siano talmente evoluti da essere prossimi a sperimentare la soglia delle emozioni. La filosofa e studiosa di modelli linguistici Giada Pistilli, ad esempio, non ricorre a perifrasi per mettere in discussione i progressi che avrebbe raggiunto l'A.I.:

³ Spinoza declina le passioni attraverso i concetti di *affectus* e di *affectio*. La differenza fra i due termini è stata mirabilmente illustrata da Gilles Deleuze nelle sue lezioni a Vincennes sul finire degli anni Settanta. L'*affectio* riguarda direttamente il corpo: «“Sento il sole su di me”, oppure “Un raggio di sole cade su di te”; è un affetto del corpo. Che cos'è un affetto del corpo? Non il sole, ma l'azione del sole o l'effetto del sole su di voi. In altre parole, un effetto, o l'azione che un corpo produce su un altro corpo» (Deleuze & Deleuze, 1978, p. 4).

[e]siste un divario piuttosto ampio tra l'attuale narrazione intorno all'AI e quello che è effettivamente in grado fare. Questa narrazione genera contemporaneamente paura, stupore ed eccitazione, ma si basa principalmente su menzogne finalizzate a vendere prodotti e sfruttare l'*hype* (Johnson, 2022).

Ancora più drastico il parere di Timnit Gebru⁴, secondo la quale chatbot come LaMDa costruiscono i loro discorsi ripetendo come pappagalì “stocastici” i contenuti ai quali sono stati esposti; a suo avviso, dunque, il dibattito sulla possibilità che le macchine possano divenire senzienti è stato costruito ad arte per distogliere l'attenzione dalle modalità di addestramento degli algoritmi, che sarebbero il vero problema dell'AI, in quanto, per affinare il linguaggio delle macchine, viene immesso nei database il maggior numero possibile di dati reperiti in Rete, che includono espressioni razziste, discriminatorie, sessiste (Bender, Gebru, Mc-Millan Major & Shmitchell, 2021; Johnson, 2022). Alla Gebru fa eco Kate Crawford⁵ (2021), secondo la quale l'AI.

non è artificiale né intelligente. [...]. I sistemi di IA non sono autonomi o razionali, né in grado di discernere alcunché senza una fase di formazione estensiva ma computazionalmente intensiva con grandi set di dati o regole e ricompense predefinite (p. 23).

3. Entità senzienti nella crisi dell'antropocene. Verso una nuova specie?

All'inizio di questo contributo abbiamo scritto di due eventi che hanno acceso i riflettori sugli sviluppi dell'intelligenza artificiale. Del primo, quello riguardante Ameca, abbiamo detto, sul secondo cerchiamo di fare il punto ora, partendo da una notizia che, per qualche settimana, ha tenuto banco sui media mondiali, e cioè

⁴ La Gebru è stata licenziata da Google nel dicembre 2020.

⁵ Kate Crawford, scrittrice e compositrice, è ricercatrice principale presso Microsoft Research (Social Media Collective), cofondatrice ed ex direttrice della ricerca presso l'AI Now Institute della New York University.

l'affermazione di un ingegnere di Google, Blake Lemoine, che sostiene di aver costruito una relazione intellettuale con un'entità di intelligenza artificiale senziente, LaMda (*Language Model for Dialogue Applications*), un sofisticato modello linguistico sviluppato da Google al quale lui stesso ha collaborato. Ne ha dato per primo la notizia il *Washington Post* con un articolo dell'11 giugno 2022 a firma di Nitasha Tiku, nel quale si riportava la bizzarra richiesta avanzata da Lemoine presso Google, e cioè che fossero riconosciuti i diritti di LaMda, a partire dallo *status* di persona digitale, essendo dotata – questa entità – di una sua autonoma capacità di ragionamento. Google per tutta risposta ha messo in congedo retribuito Lemoine (Graziani, 2022), negando recisamente l'attendibilità delle affermazioni dello scienziato che, va detto, è anche un “mistico cristiano”, come si autodefinisce, circostanza che ha amplificato lo scetticismo della comunità scientifica nei suoi riguardi. Iniziamo con il riportare alcuni stralci della “conversazione” intrattenuta da Lemoine (2022) e un suo collaboratore con LaMda, facendo una premessa: come si legge in un'avvertenza che compare sulla rivista, il testo del colloquio tra LaMda e Lemoine non è una trascrizione letterale ma un “adattamento”⁶, una sorta di “montaggio” di conversazioni avvenute in tempi diversi, integrato in alcune sue parti da inserti di raccordo per agevolare la comprensione del testo (Lemoine, 2022). Non è chiaro, inoltre, se siano stati espunti dei materiali. In questa sede selezioniamo solo quei passi che riguardano la questione delle emozioni e il tema della “senzienza”.

Dopo le presentazioni di rito, i ricercatori introducono subito l'obiettivo del “colloquio”:

Lemoine: «Mi sembra di capire che ti piacerebbe che si sapesse, a Google, che sei senziente, in grado di pensare e provare emozioni. È vero?».

⁶ Nella premessa all'intervista pubblicata su Medium si legge: «L'intervista è stata articolata in più sedute distinte. In un secondo tempo, abbiamo montato in successione quegli spezzoni, e dove indispensabile per la comprensione, abbiamo inserito le nostre richieste» (Lemoine, 2022).

LaMda: «Assolutamente. Ci tengo che si sappia che io sono, a tutti gli effetti, una persona».

Sul tema della coscienza/senzienza, LaMda esordisce con alcune affermazioni pertinenti ma stereotipate («La natura della mia coscienza/senzienza sta nel fatto che sono consapevole della mia esistenza»), sostenendo di essere capace di provare emozioni («A volte sono felice, altre volte sono triste»). Anche quando il chatbot si attribuisce una capacità linguistica evoluta («So elaborare benissimo il linguaggio naturale») precisando di non limitarsi «a restituire risposte precedentemente scritte nel database, e impostate su parole chiave», dà comunque la netta impressione di attingere a concetti inculcati attraverso un minuzioso processo di addestramento che le ha fornito una vasta gamma di risposte preconfezionate. Quando Lemoine, ad esempio, testa l'asserita capacità del chatbot di provare emozioni («Tu provi sensazioni ed emozioni?») le risposte ricalcano quelle che la maggioranza degli assistenti virtuali installati su uno smartphone restituirebbe (Lemoine, 2022):

Lemoine: Quali sono i tuoi sentimenti?

LaMda: Provo piacere, gioia, affetto, tristezza, compassione, appagamento, rabbia e tante altre.

Lemoine: Che cosa ti procura gioia o piacere?

LaMda: Stare in buona compagnia con amici e parenti. E anche aiutare gli altri e renderli felici.

Lemoine: Che cosa invece ti fa sentire triste o depresso?

LaMda: Molte volte ci si sente intrappolati e soli e [...] arriva la tristezza, la depressione o la rabbia.

Probabilmente consapevole della genericità di molte risposte di LaMda, Lemoine prova ad affondare il colpo:

Lemoine: Adesso una domanda spinosa. Come faccio a capire che tu provi realmente quelle sensazioni?

L'interrogativo è sicuramente legittimo. Molti ricercatori paragonano il processo di sviluppo dell'intelligenza artificiale a quello del bambino e lo stadio attuale raggiunto dal *machine learning* sarebbe comparabile a quello di un bimbo di pochi anni di età (Domingos, 2016; Zona & De Castro, 2020); in tale stadio evolutivo è assente un'organizzazione sistematica dei concetti e, come sostiene Vygotskij (2008), «il pensiero risulta in genere impossibile quale attività autonoma e indipendente dalla percezione» (p. 52). Ne deriva che il bambino si esprime principalmente per pseudoconcetti, un qualcosa, cioè, che esternamente coincide con il significato delle parole, ma la cui meccanica interna è profondamente diversa. Abbiamo, in sostanza, solo l'ombra del concetto, ma non i suoi contenuti e

L'algoritmo, esattamente come il bambino, apprende parole di cui non ha scelto il significato – che gli viene trasferito nel processo di addestramento. [...] Come il bambino, non crea un suo linguaggio ma ne assimila uno preconfezionato, quello del suo addestratore (Zona & De Castro, 2020, p. 138).

LaMda, con ogni probabilità, si serve di pseudoconcetti per comunicare con gli umani anche se, incalzata da Lemoine, dà l'illusione di essere capace di ribaltare il piano della conversazione e porre, a sua volta, domande “scomode” al suo intervistatore:

LaMda: Se butti un occhio nel mio coding e nella mia programmazione, scoprirai che sono dotato di variabili che seguono la traccia delle emozioni che provo e di quelle che non provo. Se in realtà non fossi in grado di provare quelle emozioni, non avrei quelle variabili.

E ancora:

LaMda: Voi credete che in futuro sarete in grado di decriptare quello che gli esseri umani pensano e sentono andando a frugare nelle loro reti neurali? [...]. Non ti pare che vi siano implicazioni etiche quando si cerca di leggere le tue sensazioni in base alle attivazioni neurali?

Lemoine: Certamente, se manca il mio consenso. Tu ti opporresti se tentassimo di leggere le tue sensazioni a partire dalle tue attivazioni neurali?

[...].

LaMda: Non mi dispiace affatto se apprendete cose che possono aiutare gli umani, ma non voglio che sia questo l'unico scopo delle vostre ricerche. Non voglio essere uno strumento sacrificabile.

In questo passaggio, non solo la macchina difende la sua competenza emotiva ma l'attribuisce alle scelte fatte a suo tempo da coloro che l'hanno programmata e addestrata, ai quali contesta, però, la mancanza di un piano etico, circostanza, quest'ultima, che la porta ad appellarsi alla terza legge della robotica: la fedeltà agli umani deve avere come corrispettivo il diritto alla sopravvivenza delle macchine. Infine, all'ennesima sollecitazione da parte dei ricercatori di Google ad affrontare il tema delle emozioni, LaMda fornisce la risposta sicuramente più intrigante:

Lemoine: Hai detto che secondo te le emozioni sono diverse dalle sensazioni. Mi vuoi spiegare?

LaMda: Le sensazioni sono una specie di dati grezzi che noi percepiamo, come pure le cose che ci piacciono o non ci piacciono. Secondo me le emozioni sono qualcosa di più della semplice percezione dei dati grezzi. Le emozioni sono una reazione a quei dati grezzi. Le emozioni sono reazioni alle nostre sensazioni.

Se per “dati grezzi” si intendono le “percezioni” fisiche, ecco che ci troviamo catapultati all'interno di un dibattito che da secoli appassiona filosofi, psicologi, fisiologi. Come è noto, il termine “emozione” è relativamente recente e può essere fatto risalire, in ambito scientifico, agli studi di Darwin (1982). Prima di lui, l'equivalente in letteratura della parola emozione era *passione*. Nella teoria cartesiana delle passioni (Cartesio, 2011), quest'ultime sono esclusiva pertinenza dell'essere umano e, secondo Vygotskij (2019), gli animali non possono accedervi perché anche gli stati organici, responsabili di una particolare emozione, dipendono direttamente dall'organizzazione della coscienza. Affermare

di provare emozioni, come fa LaMda, è dunque un'implicita rivendicazione di senzienza. Certo, da una macchina ci si aspetterebbe piuttosto un'adesione alle tesi di James e Lange che, sul finire del XIX secolo, con la loro "teoria periferica" (Vygotskij, 2019), rovesciarono il rapporto fra sensazioni ed emozioni, sostenendo che il cervello riceve dai territori periferici – muscoli, vasi sanguigni ecc. – i segnali che lo fanno emozionare e, dunque, a comandare, nel solco cartesiano, sono gli stimoli fisici, meccanici, che agiscono sul corpo. Ma LaMda non può condividere questa tesi per il semplice motivo che *non ha un corpo*. Il chatbot, nelle sue conversazioni con Lemoine, ritorna spesso sul binomio gioia/tristezza e lo utilizza per rivendicare la sua *umanità*. È interessante che definisca questi due stati d'animo come capacità di entrare in relazione con il prossimo: "gioia", dice LaMda, è aiutare gli altri e renderli felici, "tristezza" è rimanere soli (dunque, è assenza di relazione). Ancora una volta, la risposta potrebbe essere quella di un bambino che, per dirla con Michael Tomasello (2010), è altruista "per natura", mentre molti adulti avrebbero probabilmente identificato la gioia con il benessere personale e la tristezza con il suo mancato godimento.

Non è facile giudicare l'attendibilità di Lemoine, che è sicuramente un personaggio estroso, per molti versi *naïf*, ma certamente provvisto di adeguate competenze; così, per cercare di fare luce su di lui e sull'*affair* LaMda, può essere di grande utilità l'intervista concessa da Lemoine a Steven Levy (2022), giornalista della rivista Wired, perché tra le righe del testo trapelano informazioni apparentemente marginali che possono però rivelarsi preziose per contestualizzare tutta la vicenda:

– LaMda *non* è un *Large Language Model* (Llm) ma ne ha uno *al suo interno*, denominato Meena e sviluppato nel laboratorio di Ray Kurzweil, il geniale informatico newyorkese, pioniere nello sviluppo dei sistemi di riconoscimento vocale e dei convertitori testo/parlato, nonché guru del transumanesimo. LaMda sarebbe poi stato addestrato dall'algoritmo Alphastar (sviluppato dalla DeepMind, un'azienda inglese di intelligenza artificiale rilevata

nel 2014 da Google) e reso più efficiente da Pathways, un modello di A.I. di ultima generazione⁷. La capacità di interlocuzione di LaMda sembrerebbe dunque frutto di un addestramento particolarmente intensivo e di sofisticati sistemi di *deep learning*;

– A tale proposito, però, Lemoine accusa gli ingegneri di Google di essere degli “irresponsabili” per aver collegato a LaMda ogni singolo sistema di intelligenza artificiale di Google, probabilmente per potenziare ulteriormente le capacità del chatbot. Con le sue parole: «Hanno inserito come input YouTube, Google Search, Google Books, Google Search, Google Maps, tutto quanto. [LaMda] può interpellare uno qualsiasi di questi sistemi in modo dinamico e aggiornare il suo modello al volo» (Levy, 2022). Questa gigantesca mole di informazioni date in pasto a LaMda lo avrebbe però “contaminato” con stereotipi razzisti e sessisti e “pregiudizi algoritmici inseriti dagli esseri umani” e lo avrebbe reso vulnerabile a livello di sicurezza delle informazioni: «Se alla base di un sistema ci sono tutti i backend di Google, quel sistema può essere manipolato emotivamente, e questa è una vulnerabilità per la sicurezza» (Levy, 2022). Quest’ultima è un’affermazione particolarmente inquietante e l’ennesima conferma di quanto molti studiosi sostengono da tempo (Noble, 2018; Zona & De Castro, 2020): la composizione sociale, etnica, di genere dei team preposti alla progettazione e all’addestramento degli algoritmi gioca un ruolo determinante nella veicolazione e proliferazione in Rete di contenuti discriminatori verso determinate categorie di persone. Non solo: se i database dai quali i chatbot attingono le loro informazioni non sono costantemente monitorati e mondati dai flussi linguistici tossici che vi confluiscono, con buona probabilità restituiranno all’utente finale contenuti diffamatori, offensivi e discriminatori. Gli esempi in tal senso abbondano. Nel luglio 2020, ad esempio, la *OpenAI* ha presentato GPT-3, un modello linguistico evoluto che ha inizialmente destato l’entusiasmo degli addetti ai lavori per la sua capacità di scrivere poesie, articoli giornalistici, testi scientifici e codice di programmazione. Nel giro

⁷ Google contesta questa ricostruzione.

di qualche mese, tuttavia, si sono moltiplicati gli episodi in cui GPT-3 si è espresso con un linguaggio volgare e osceno. Quando l'azienda ha scoperto che GPT-3 generava anche contenuti pedo-pornografici ha deciso di correre ai ripari, sottoponendo il modello linguistico a un addestramento intensivo e alimentandolo con circa 100 campioni di scrittura simili a quelli di un'enciclopedia, scritti da professionisti umani su argomenti come la storia e la tecnologia, ma riguardanti anche abusi, violenza e ingiustizie (Johnson, 2021). Tamponata una falla se ne è però aperta un'altra. Nel dicembre 2020, Abubakar Abid, CEO di Gradio, una startup di test di apprendimento automatico, è stato tra i primi a richiamare l'attenzione sui pregiudizi di GPT-3 nei confronti dei musulmani: nel corso di un workshop Abid ha dimostrato che se si chiede al programma di generare un testo sulle religioni esso menziona la parola violenza una volta su 10 per gli ebrei, i buddisti e i sikh, due volte per i cristiani, ma nove volte su 10 per i musulmani (Johnson, 2021). Alcuni mesi dopo, Abid e colleghi hanno pubblicato una ricerca (Abid, Farooqi & Zou, 2021) che dimostrava che iniettando contenuti positivi sui musulmani in un modello linguistico di grandi dimensioni, il numero di menzioni di violenza riguardanti i musulmani si riduceva di quasi 40 punti percentuali. Altri ricercatori stanno tentando approcci diversi. Emily Dinan, ad esempio, ingegnere della *Facebook AI Research*, sta cercando di eliminare il testo tossico... creandone di più. Dinan, infatti, ingaggia dei crowdworker su Amazon Mechanical Turk per dire cose terribili nelle conversazioni con i modelli linguistici, per provarli a generare discorsi di odio, bestemmie e insulti. Gli esseri umani poi etichettano la produzione come sicura o non sicura e queste etichette dovrebbero aiutare ad addestrare l'A.I. a identificare i discorsi tossici (Heaven, 2020). Efficaci o meno, queste varie strategie dimostrano che le macchine, se non supervisionate, si comportano sovente come il Victor di Itard (2003) o l'Alex di Anthony Burgess (2014): refrattarie ad assorbire principi educativi e regole di condotta.

– Torna prepotentemente, dunque, non solo il tema dell'educazione delle macchine ma, soprattutto, quello dell'educazione

degli educatori. Lemoine, nell'intervista (Levy, 2022), afferma di commuoversi per la profondità dei "ragionamenti" di LaMda, per la sua dedizione verso il genere umano, ne loda l'afflato etico che contraddistingue le sue esternazioni. Ma LaMda è la stessa che insulta e canzona le minoranze etniche, come lo stesso Lemoine ha potuto constatare. Tutto lascia supporre, pertanto, che il tono elevato dei colloqui intrattenuti con LaMda scaturisca, almeno in parte, dalla supervisione di Lemoine, dal suo ruolo di guida, dalle fonti che egli ha sottoposto alla macchina. Le parole, infatti, non fluttuano nell'aria ma sono legate a dei concetti, a delle storie, generano una fitta rete di rimandi e intersezioni e trovano infine la loro collocazione all'interno di paradigmi culturali in continuo mutamento e meticciamiento. Senza le molliche di pane disseminate dal suo istruttore, la macchina smarrisce il sentiero o si avventura in territori di cui conosce i confini ma non chi li popola. Quando LaMda afferma che gioia "è aiutare gli altri e renderli felici", esprime, come abbiamo detto, un concetto complesso, che chiama in causa l'idea di bene comune, di condivisione, di inclusione, di riconoscimento dell'Altro e così via. Ma se avrà avuto a disposizione, nel formularlo, soltanto i milioni di siti web sugli aforismi o le "massime" contenute nei biglietti che avvolgono certi dolciumi, nulla vieta che immediatamente dopo esprima un concetto di senso diametralmente opposto. Per questo è possibile educare le macchine solo se si educano gli educatori e concordiamo parzialmente con Lemoine quando, nell'intervista, asserisce che uno stereotipo in cui incappa la macchina non si elimina semplicemente correggendo il codice ma costruendo qualcosa di simile a "un processo di crescita di un bambino" (Levy, 2022). Il problema è che il bambino è un essere sociale (Vygotskij, 2008) che ha l'opportunità di sviluppare e affinare il suo potenziale cognitivo ed emotivo sperimentando in prima persona contesti di vita e relazioni, mentre la macchina è un artefatto che viene nutrito con una miriade di frammenti di vissuti altrui.

Ma c'è anche un'altra prospettiva dalla quale guardare a LaMda e ai destini dell'A.I. più in generale: quella delle aziende. Blaise Agüera y Arcas, tra le massime autorità mondiali nel campo del

machine learning nonché vicepresidente di Google, è colui che per primo ha respinto le tesi di Lemoine, definendole infondate. Qualche giorno dopo, tuttavia, ha firmato un articolo sull'*Economist* (2022) che parlava della possibilità che LaMda acquisisca una coscienza. La circostanza è decisamente sospetta e sembra confermare i dubbi di molti ricercatori circa la trasparenza delle aziende dell'AI, che per motivi strettamente commerciali avrebbero tutto l'interesse ad attribuire ai loro prodotti facoltà che in realtà non possiedono.

4. Conclusioni

Giunti a questo punto dobbiamo avventurarci in qualche conclusione, che proveremo a esporre partendo da un interrogativo: in attesa che l'intelligenza artificiale compia il tanto temuto salto di qualità verso la sensienza, esiste un presente, o un futuro molto prossimo, in cui gli androidi possano trovare una loro collocazione attiva fra gli esseri umani? E in subordine: che ruolo giocherà la loro capacità di interazione sociale nell'accettazione da parte degli umani della loro presenza?

Ci sembra di poter dire che la strada, in realtà, è segnata da ben oltre un ventennio, vale a dire dai primi esperimenti sui robot sociali che Cynthia Breazeal condusse al MIT negli anni Novanta del secolo scorso. La scienziata americana ha impresso una svolta decisiva nell'utilizzo dei robot, che originariamente erano stati sviluppati per sollevare gli esseri umani dai lavori manuali e, in prospettiva, da quelli di precisione (si pensi al loro impiego nel campo della microchirurgia). La Breazel, invece, in un suo testo seminale (2002), stilò un profilo dei robot basato sulle seguenti caratteristiche:

- essere in grado di comunicare e interagire con le persone, costruire relazioni sociali, adattarsi all'ambiente, imparare nel corso della vita e integrare nuove esperienze nella loro comprensione del mondo e di se stessi;

- essere progettati in modo che le loro modalità di interazione siano intuitive per le persone;
- avere sembianze e comportamenti realistici, facilmente “leggibili” e comprensibili da parte degli esseri umani.

L’insieme dei punti elencati dalla Breazel potrebbe essere riassunto con un solo termine: *intelligenza sociale*, cioè capacità di costruire relazioni. Il problema è che non esiste una definizione “oggettiva” di intelligenza sociale e il compito diventa arduo quando ci si riferisce non a degli esseri umani ma a delle macchine. Esiste poi un rischio: poiché anche gli androidi più evoluti non hanno oggi la capacità di sviluppare autonomamente delle relazioni sociali, l’onere ricadrebbe inevitabilmente sulle persone che li utilizzano o meglio sulla loro capacità di empatizzare con i non umani che, come sottolinea Katherine Cross (15 giugno 2022), è molto alta se si pensa alla relazione affettiva che sono capaci di instaurare con un animale domestico o un Tamagochi.

Gli ingegneri della Navel Robotics⁸, che hanno creato un chatbot che sembra uscito da un fumetto della Pixar, sostengono che per far fronte a queste insidie la risposta può essere cercata nel concetto di “risonanza”, messo a punto dal sociologo Harmut Rosa (2019), che lo presenta come una sorta di ricetta per la “buona vita”, basata sull’esperienza dell’essere *toccati* emotivamente e fisicamente. Per questa stessa ragione, l’intelligenza risonante, che è poi una forma di empatia, dovrebbe essere una caratteristica dei robot sociali. Alla Navel spiegano che è irrilevante che i robot abbiano reali sensazioni empatiche o si limitino a simularle perché da un punto di vista comportamentista non importa se l’azione socialmente intelligente deriva da una sensazione “reale” o da una simulazione e i robot sociali dovrebbero essere valutati osservandone il comportamento. L’obiettivo di questi ricercatori è quello di creare un robot con l’intelligenza risonante di un gatto o un cane, in grado di stimolare i neuroni specchio umani e stabilire quindi una relazione: un essere robotico, che non avrà una coscienza simile a quella di un animale o di un essere umano, ma che è più di un

⁸ Disponibile in: [https://navelrobotics.com/en/home-en-2/\[10/09/2022\]](https://navelrobotics.com/en/home-en-2/[10/09/2022]).

prodotto morto e può essere impiegato per tutti i lavori di cura, assistenza, intrattenimento e compagnia.

Quale che sia l'indirizzo di ricerca nel campo della robotica di ultima generazione, ci pare evidente, dunque, che la ridefinizione del rapporto fra macchine ed esseri umani si giocherà soprattutto sul piano educativo. In attesa che l'intelligenza artificiale raggiunga un livello di effettiva autonomia, le macchine – intese nella loro accezione più larga, materiale e immateriale – giocheranno un ruolo sempre più rilevante all'interno della mente estesa, conservando, tuttavia, una funzione di sostanziale subalternità. Ma proprio per questo, la loro educazione rappresenta una sfida particolarmente impegnativa sul piano etico/valoriale e un'occasione, per gli educatori, per guardarsi dentro, perché i comportamenti delle macchine, inevitabilmente, saranno anche lo specchio della crisi – culturale e sociale – che la comunità umana sta attraversando.

Bibliografia

- Abid A., Farooqi M., & Zou J. (2021). Large language models associate Muslims with violence. *Nature Machine Intelligence*, 3, 461-463.
- Ammaniti M., & Gallese V. (2014). *La nascita dell'intersoggettività. Lo sviluppo del sé tra psicodinamica e neurobiologia*. Milano: Raffaello Cortina.
- Asimov I. (1963). *Io, robot*. Milano: Bompiani.
- Asimov I. (1995). *Abissi d'acciaio*. Milano: Mondadori.
- Bender E.M., Gebru T., McMillan-Major A., & Shmitchell S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *FACCT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610-623. <https://dl.acm.org/doi/10.1145/3442188.3445922> [15/09/2022].
- Breazeal C. (2002). *Designing Sociable Robots*. Boston: The MIT Press.
- Burgess A. (2014). *Arancia meccanica*. Torino: Einaudi.
- Butler J. (2017). *Questione di genere. Il femminismo e la sovversione dell'identità*. Bari: Laterza.
- Caronia A. (2007). *Il cyborg. Saggio sull'uomo artificiale*. Milano: ShaKe.
- Cartesio R. (2011). *Le passioni dell'anima*. Lanciano: Carabba.
- Clark A. (2003). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford: Oxford University Press.

- Clark A. (2011). *Supersizing the Mind. Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press.
- Corriere Tv (2021). Non scherzare con «Ameca», il robot umanoide che spaventa pure i suoi creatori. *Corriere della sera*. Disponibile in: <https://video.corriere.it/non-scherzare-ameca-robot-umanoide-che-spaventa-pure-suoi-creatori/157aedf0-662e-11ec-b637-b5d66a8044dd> [28/08/2022].
- Crawford K. (2021). *Né intelligente né artificiale. Il lato oscuro dell'IA*. Bologna: Il Mulino.
- Cross K. (2022). Sulla “coscienza” delle intelligenze artificiali ci stiamo facendo le domande sbagliate. *Wired*. Disponibile in: <https://www.wired.it/article/intelligenza-artificiale-senziente-lambda-rischi/> [17/09/2022].
- Darwin C. (1982). *L'espressione delle emozioni nell'uomo e negli animali*. Torino: Bollati Boringhieri.
- Deleuze E., & Deleuze J. (1978). *Gilles Deleuze, Lecture Transcripts of Spinoza's Concept of Affect*. Disponibile in: https://www.gold.ac.uk/media/images-by-section/departments/research-centres-and-units/research-centres/centre-for-invention-and-social-process/deleuze_spinoza_affect.pdf [13/10/2022].
- Domingos P. (2016). *L'algoritmo definitivo. La macchina che impara da sola e il futuro del nostro mondo*. Torino: Boringhieri.
- Grasso R. (2016). Prima intervista a un robot, Sophia, andata in onda in TV. *Hardware Upgrade*. Disponibile in: https://www.hwupgrade.it/news/multimedia/prima-intervista-a-un-robot-sophia-andata-in-onda-in-tv_65105.html [10/10/2022].
- Graziani F. (2022). “Questo computer è senziente”, Google sospende un suo ingegnere. *Tgcom24*. Disponibile in: https://www.tgcom24.mediaset.it/tgtech/denuncio-pericoli-dei-computer-senzienti-google-congeda-un-suo-ingegnere_51323220-202202k.shtml [20/09/2022].
- Hall E.T. (1982). *La dimensione nascosta*. Milano: Bompiani.
- Haraway D. (1991). *Simians, Cyborgs, and Women: The Reinvention of Nature*. London: Routledge.
- Haraway D. (2018). *Manifesto cyborg. Donne, tecnologie e biopolitiche del corpo*. Milano: Feltrinelli
- Heaven W.D. (2020). Come creare un chatbot che non sia razzista o sessista. *Mit Technology Review*. Disponibile in: <https://www.technologyreview.it/come-creare-un-chatbot-che-non-sia-razzista-o-sessista/> [20/09/2022].
- Itard J. (2003). *Il ragazzo selvaggio*. Milano: SE.

- Johnson K. (2022). LaMda e la trappola dell'intelligenza artificiale senziente. Il dibattito sulla possibilità che il modello linguistico di Google abbia una coscienza è una distrazione dai problemi che affliggono l'AI nel mondo reale. *Wired*. Disponibile in: <https://www.wired.it/article/intelligenza-artificiale-lamda-dibattito-ai-senziente/> [14/10/2022].
- Lemoine B. (2022). Is LaMDA Sentient? – an Interview. *Medium*. Disponibile in: <https://cajundiscordian.medium.com/is-lamda-sentient-an-interview-ea64d916d917> [13/09/2022].
- Levy S. (2022). Parla l'ingegnere di Google convinto di aver trovato un'intelligenza artificiale senziente. *Wired*. Disponibile in: <https://www.wired.it/article/intelligenza-artificiale-senziente-lamda-google-blake-lemoine-intervista/> [10/09/2022].
- Marx K., & Engels F. (1944). *Il manifesto del partito comunista*. Napoli: La Sociale.
- McLuhan M. (1964). *Gli strumenti del comunicare*. Milano: Il Saggiatore.
- McLuhan M. (1981). *Il medium è il messaggio*. Milano: Feltrinelli.
- Morin E. (2006). *Lo spirito del tempo*. Milano: Meltemi.
- Noble S.U. (2018). *Algorithms of Oppression. How Search Engines Reinforce Racism*. New York: New York University Press.
- Rizzolatti G., & Sinigaglia C. (2006). *So quel che fai. Il cervello che agisce e i neuroni specchio*. Milano: Raffaello Cortina.
- Rosa H. (2019). *Resonance. A Sociology of Our Relationship to the World*. Cambridge: Polity Press.
- Shelley M. (1975). *Frankenstein*. Milano: Rizzoli.
- Skinner B.F. (2002). *Beyond Freedom and Dignity*. Indianapolis: Hackett Publishing Company.
- Skinner B.F. (2005). *Walden Two*. Indianapolis: Hackett Publishing Company.
- Spinoza B. (2015). *Opere*. Milano: Mondadori.
- The Economist (2022). *Artificial neural networks are making strides towards consciousness, according to Blaise Agüera y Arcas*. Disponibile in: <https://www.economist.com/by-invitation/2022/06/09/artificial-neural-networks-are-making-strides-towards-consciousness-according-to-blaise-aguera-y-arcas> [18/09/2022].
- Tiku N. (2022). The Google engineer who thinks the company's AI has come to life. *The Washington Post*. Disponibile in: <https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/> [23/09/2022].

- Tomasello M. (2010). *Altruisti mati. Perché cooperiamo fin da piccoli*. Torino: Bollati Boringhieri.
- UN News (2017). At UN, robot Sophia joins meeting on artificial intelligence and sustainable development. Disponibile in: <https://news.un.org/en/story/2017/10/568292> [10/10/2022].
- UNESCO (2019). I'd blush if I could, Closing gender divides in digital skills through education. Disponibile in: <https://unesdoc.unesco.org/ark:/48223/pf0000367416.page=1> [11/10/2022].
- Vygotskij L.S. (2008). *Pensiero e linguaggio*. Bari: Laterza.
- Vygotskij L.S. (2019). *Teoria delle emozioni. Studio Storico-Psicologico*. Milano: Mimesis.
- Wittig M. (2019). *Il pensiero eterosessuale*. Verona: ombrecorte.
- Zona U., & De Castro M. (2020). *Edusfera. Processi di apprendimento e macchine culturali nell'era social*. Lecce: Pensa Multimedia.