

QUESTIONI ETICO-SOCIALI LEGATE ALL'INTELLIGENZA ARTIFICIALE VERSO UNA COLONIZZAZIONE DELLA SFERA PERSONALE?

DANIEL TOSCANO LOPEZ

ABSTRACT: This article aims to analyze some of the challenges and ethical-social implications of Artificial Intelligence (AI). To achieve this, we propose as a starting point the ambivalent nature of AI, which seduces us with its siren song because it is a formidable tool for solving countless problems and supporting decision-making. But it's also frightening because of the potential risks that technological systems, machines, artifacts or robots that operate on the basis of algorithms pose to our way of life. Second, we take a brief historical tour of AI, highlighting some of the milestones and historical-scientific conditions that have made it possible, such as cybernetics, the Turing Test, and the Dartmouth Conference. Third, we examine two scenarios based on the notion of "rarity," one educational and the other legal, in order to raise, fourth, how "rarity" awakens two dispositions of the mind that are ultimately complementary ways of positioning oneself and preparing oneself internally to assimilate something "too big" like AI systems. Finally, we raise some ethical considerations regarding AI as a tool or "co-pilot" serving not only human interests but also the natural environment. Given the immense power of AI and algorithms, shared responsibility is required, along with the promotion of transparent practices, education and pedagogy for technological literacy, and "inclusive global justice".

L'articolo si propone di analizzare alcune questioni etiche e sociali sollevate dall'introduzione e dall'utilizzo dell'Intelligenza Artificiale (IA). A tal fine, in primo luogo, considereremo la natura ambivalente dell'IA che, da un lato ci seduce con il suo canto da sirena ma, dall'altro lato, rappresenta uno strumento formidabile per risolvere innumerevoli problemi e supportare il processo decisionale. Tuttavia, essa presenta anche aspetti inquietanti per i potenziali rischi che i sistemi tecnologici, le macchine, gli artefatti o i robot che operano in base ad algoritmi potrebbero comportare rispetto al nostro stile di vita. In secondo luogo, faremo un breve *excursus* storico sull'IA, evidenziando alcune tappe fondamentali e le condizioni storico-scientifiche che hanno reso possibile l'avvento dell'IA, come ad

esempio la cibernetica, il Test di Turing e la Conferenza di Dartmouth. In terzo luogo, esamineremo due scenari basati sulla nozione di “eccezionalità”, in ambito educativo e in ambito legale, per poi, in quarto luogo, sollevare la questione dello “straniamento” che risveglia due disposizioni della mente che sono, in ultima istanza, modi complementari di prepararsi internamente ad assimilare qualcosa di “troppo grande”, come i sistemi di IA. Infine, vengono sollevate alcune considerazioni etiche riguardo all’IA, in quanto si tratta di uno strumento al servizio non solo degli interessi umani, ma anche dell’ambiente naturale. Dato il grande potere dell’IA e degli algoritmi, è necessaria una responsabilità condivisa, la promozione di pratiche di trasparenza, di educazione e pedagogia della conoscenza della tecnologia e infine, ma non da ultimo, di una forma di “giustizia globale”.

KEYWORDS: Artificial Intelligence, Rarity, Ethics, Emotional Illiteracy, Moral vertigo

PAROLE CHIAVE: Intelligenza Artificiale, Eccezionalità, Etica, Analfabetismo emotivo, Vertigine morale

1. Introduzione

Si assiste oggi all’esplosione e all’espansione dell’IA in ambiti diversi dell’agire umano. La troviamo presente e applicata, infatti, nell’e-commerce, nel marketing, nei dispositivi mobili e domestici, nelle piattaforme dei social media, nei motori di ricerca, nell’industria, nell’agricoltura, nella gestione del traffico, nelle applicazioni GPS e in ambito medico, per citare solo alcuni esempi (Coeckelberg 2020). Inoltre, si discute di IA ai convegni e negli incontri accademici e scientifici, nei media, nei nuovi programmi di formazione post-laurea, nelle leggi e nei regolamenti e persino nell’arte dove viene utilizzata come strumento per la creazione di opere artistiche.

Il potere dell’IA si manifesta in maniere diverse e in ambiti distinti ed essa ha un forte impatto sia sui singoli individui che sulla società nel suo insieme. In meteorologia, per esempio, l’IA supporta le stazioni nella produzione di bollettini meteorologici che avvisano gli utenti sulla qualità dell’aria, emettendo previsioni sempre più precise con l’ausilio di supercomputer (Bettini 2024). Inoltre, offre un’importante capacità predittiva grazie alla gestione di grandi quantità di dati e alla velocità con cui elabora le informazioni (Gastaldo, Caciamani e De Savino 2024, p. 31).

Dal punto di vista finanziario, l'IA utilizza l'apprendimento automatico per rilevare precocemente le crisi finanziarie (Bartel, Hanke e Petric 2024), anticipando potenziali turbolenze nei mercati finanziari e analizzando in maniera combinata notizie economiche e grandi volumi di dati.

La medicina predittiva e personalizzata del ventunesimo secolo, che utilizza algoritmi di IA, ha il potenziale per ottimizzare in modo efficiente le risorse mediche ed eseguire interventi chirurgici robotici assistiti dall'IA, operando in presenza o da remoto su spazi anatomici di difficile accesso e ottenendo movimenti impossibili per le mani di un chirurgo umano (Knudsen e Ghaffar 2024). Inoltre, potrebbe essere utilizzata per la diagnosi precoce delle malattie nelle fasi iniziali attraverso l'analisi delle immagini in vari campi, quali la radiologia, l'oftalmologia, la dermatologia e la cardiologia. Di conseguenza, "un potenziale vantaggio dell'IA consisterebbe nella capacità di identificare nel tempo lievi alterazioni dello stato di salute di un individuo, indicative di una patologia sottostante e nella sua diagnosi precoce" (Vidali 2024, p. 134).

Infine, nel contesto di uno scenario futuro in cui i cambiamenti climatici, le inondazioni e gli insetti nocivi inducono a riflettere sull'importanza di raddoppiare la produzione di colture (Kiran, Narayana e Talamar, 2020), i governi, le istituzioni preposte, gli scienziati e gli agricoltori stanno lavorando per costruire soluzioni sostenibili, a partire dall'utilizzo, ad esempio, del machine learning per provare a fronteggiare il compito titanico di fornire cibo a 2 miliardi di persone in più entro il 2050.

L'IA ci seduce con il suo canto di sirena, perché è uno strumento formidabile per risolvere una molteplicità di problemi e perché può fornire un supporto notevole nei processi decisionali. Essa spaventa però anche per i rischi potenziali con cui i sistemi tecnologici, le macchine, gli artefatti, i dispositivi o i robot che funzionano sulla base di algoritmi possono avere, considerando il loro impatto sull'ambiente sociale, sulle nostre pratiche e abitudini e sui nostri stili di vita.

2. Il punto di partenza

Nel suo testo *La società del rischio*, Ulrich Beck parla, da un lato, dei pericoli naturali che minacciavano la vita dell'uomo premoderno, come

terremoti, tsunami, epidemie, siccità e inondazioni, nonché eruzioni vulcaniche. Inoltre, questo tipo di rischio era legato alla “violenza imprevedibile, praticata da altri esseri umani (ladri, pirati, mercenari, invasori, ecc.)” (Vázquez 2005, p. 97). D’altra parte, il sociologo tedesco sottolinea un altro insieme di rischi tipici della società globalizzata di oggi: la contaminazione nucleare o chimica, le sostanze nocive e tossiche presenti nell’aria, nel suolo e negli alimenti che attaccano piante e animali, nonché le malattie legate alla civiltà. La peculiarità di questo tipo di rischio risiede nel fatto che esso deriva dallo sviluppo tecnoscientifico che provoca danni sistematici, irreversibili e impercettibili alla percezione umana.

L’IA viene presentata come uno strumento formidabile per prevedere e gestire tali rischi. Tuttavia, oltre alle opportunità offerte all’umanità, questa tecnologia comporta anche altri rischi e sfide legati alla sicurezza, alla salute, alla manipolazione, alla fallibilità, alla discriminazione, all’uso improprio, alle violazioni della privacy e della dignità e all’elaborazione illecita dei dati, per citarne alcuni.

Stiamo subendo una trasformazione a opera di sistemi algoritmici, tecno-sociali e radicati nei processi di digitalizzazione globale e nell’emergere del capitalismo digitale, senza riuscire a tenere il passo e, in molti casi, senza riuscire a comprenderne il funzionamento. Per questo motivo, quando ci troviamo di fronte a questi dispositivi e sistemi tecnologici, abbiamo spesso una doppia sensazione: da un lato, di familiarità, dovuta alla loro vicinanza, ma dall’altro, di “estraneità”, per la loro natura insolita.

Tra il fascino e la paura suscitati dall’IA, i cittadini di oggi provano una sorta di “vertigine morale” (Sandel 2007), un’inquietudine che li porta a rapportarsi a questi sistemi tecnologici come se fossero emotivamente analfabeti (Anders 1988). Se la vertigine morale è una preoccupazione irrisolta, conseguenza del fatto che gli esseri umani oscillano tra fascino e paura nei confronti dei sistemi di IA, a causa della loro ampia gamma di opzioni e della velocità con cui stanno trasformando la società e gli individui, d’altra parte si rileva una sorta di analfabetismo emotivo, risultato dell’incapacità di interpretare i processi e le pratiche introdotti dall’IA nella nostra vita quotidiana senza essere in grado di riflettere su come questi influenzino le nostre questioni personali e intime.

Dobbiamo quindi riflettere criticamente su ciò che stiamo facendo con l'IA e su cosa questa tecnologia farà a noi. Dobbiamo chiederci: in che modo l'ipertrofia dei sistemi di IA potrebbe influire sulle persone? Da cosa nasce la sensazione di “straniamento” nei confronti dei sistemi e degli algoritmi di Intelligenza Artificiale? Dal momento che non li comprendiamo appieno, restano avvolti da un'aura di mistero e finiscono per prendere decisioni al posto nostro.

3. Radici storiche e scientifiche dell'IA

In generale, l'IA può essere collocata all'interno delle tecnologie emergenti, come le nanotecnologie, le biotecnologie, le scienze informatiche e cognitive, in particolare è legata a queste ultime ed è caratterizzata da uno stretto rapporto con la robotica e le interfacce cervello-macchina. Sebbene non esista una definizione standard di IA, poiché il suo concetto si è evoluto nel tempo, essa può essere intesa in vari sensi, sempre limitati. Ad esempio, può essere intesa come un sistema automatizzato, come la capacità di una macchina di eseguire compiti che richiedono l'apprendimento, il ragionamento e la percezione umana o anche come una disciplina che fa parte dell'informatica. Il termine si riferisce anche a una serie di tecnologie diverse che vanno dall'apprendimento automatico all'elaborazione del linguaggio naturale.

Senza l'intenzione di ridurla agli schemi angusti di una definizione, l'IA può essere intesa come “la capacità di una macchina di svolgere funzioni cognitive che associamo alla mente umana, come percepire, ragionare, apprendere, interagire con un ambiente, risolvere problemi e perfino esercitare la creatività” (McKinsey & Company 2023). Da parte sua, il regolamento della Commissione Europea definisce l'IA in termini di:

Un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali” (Regulation EU 2024).

Negli anni '40 e '50 diverse tappe fondamentali hanno rappresentato il punto di partenza dell'IA.

Il primo è la cibernetica di Norbert Wiener. Nel libro *Cybernetics, or Control and Communication in the Animal and the Machine*, pubblicato nel 1948, il matematico americano, cercando di rispondere al problema del controllo della comunicazione tra animali e macchine, concepisce la cibernetica come “la scienza che studia la progettazione e la realizzazione di sistemi artificiali in grado di autoprogrammarsi” (Dedicated 2020).

Allo stesso modo, Wiener riteneva che queste macchine avrebbero sostituito l'uomo nel suo ruolo di agente e che sarebbero state utilizzate per studiare l'intelligenza. Durante la Seconda guerra mondiale, Wiener stava cercando di risolvere il problema pratico di inventare una macchina in grado di registrare il movimento degli aerei nemici e di inviare queste informazioni per permettere il posizionamento strategico dei cannoni. Oltre a Wiener, John von Neumann contribuì allo sviluppo della cibernetica indagandone i fondamenti matematici e logici e dando origine alla Teoria generale degli automi, mentre Claude Shannon lo fece sviluppando la Teoria dell'informazione.

La seconda pietra angolare è stata il cosiddetto Test di Turing, ideato per verificare non se le macchine in generale siano in grado di pensare, ma se una determinata macchina sia intelligente o meno. Nel 1950, nell'articolo *Computing Machinery and Intelligence*, il matematico britannico Alan Turing avanzò l'ipotesi che una macchina potesse imitare il comportamento della mente umana. Questa volta la motivazione era la necessità di decifrare i messaggi nazisti e, per questo, Turing lavorò per gli inglesi, decifrando le posizioni e le rotte dei sottomarini tedeschi, nel tentativo di garantire che i convogli di rifornimento provenienti dagli Stati Uniti e dal Canada non venissero compromessi nel corso dei combattimenti.

Il terzo evento fu la “Dartmouth Summer Research Conference on Artificial Intelligence”, organizzata nel 1956 da John McCarthy, Marvin L. Minsky e Claude E. Shannon, dove fu sollevato il problema delle “macchine intelligenti”. Lì, personaggi come Herbert Simon e Allen Newell si incontrarono per affrontare la questione dell'IA rispetto alla possibilità di “rendere una macchina in grado di esibire

comportamenti che sarebbero chiamati intelligenti se fossero prodotti da un essere umano” (McCarthy *et al.* 2006). La conferenza promosse la ricerca sullo sviluppo e l’uso di macchine in grado di svolgere compiti considerati come tipici dell’intelletto umano e del comportamento intelligente.

Da allora, il progresso della ricerca nel campo dell’IA è stato inarrestabile, grazie alla convergenza tra cibernetica, teoria dell’informazione, informatica, ragionamento formale, ricerca e sviluppo di reti neurali, nonché al contributo di discipline come la filosofia, la linguistica e le neuroscienze. Tuttavia, il percorso dell’IA non è stato privo di insidie: si pensi agli anni ‘80, quando i sistemi esperti erano costosi da mantenere, incapaci di apprendere e con la possibilità di commettere errori, o al progetto informatico di “quinta generazione” fallito in Giappone, che mirava a sviluppare computer che utilizzassero l’IA.

Recentemente, i sistemi di IA sono stati progettati con diversi livelli di complessità. Alcuni sono progettati e calibrati per svolgere compiti specifici e limitati (IA stretta, indicata in inglese con l’espressione *narrow AI*), come la traduzione da una lingua in un’altra o la classificazione delle immagini. Altri sistemi utilizzano l’apprendimento automatico basato principalmente sull’inferenza statistica anziché sulla deduzione logica. Si tratta comunque di sistemi di IA che si limitano a eseguire le istruzioni fornite dal programmatore. In un contesto più ipotetico, alcuni ricercatori hanno voluto sviluppare un’Intelligenza Artificiale Generale, ovvero un sistema unico in grado di svolgere qualsiasi compito alla portata della mente umana. Anche per alcuni transumanisti, questo tipo di tecnologia potrebbe superare la capacità intellettuale umana fino a diventare una “superintelligenza”, una macchina più intelligente di noi. Nick Bostrom, nel suo libro *Superintelligence: Paths, Dangers, Strategies*, evidenzia i possibili rischi e pericoli che comporterebbe l’invenzione della superintelligenza, ma riflette anche sulla necessità di introdurre valori nell’IA. Per raggiungere questo obiettivo, propone una “normatività indiretta” che non specifica un particolare contenuto morale, ma piuttosto stabilisce criteri astratti forniti in anticipo dall’essere umano affinché l’IA possa aderirvi.

La proliferazione e l’implementazione delle cosiddette “macchine intelligenti” nella vita di tutti i giorni sono avvenute a ritmo sostenuto nel

XXI secolo e sono considerate come uno dei risultati più emblematici della “quarta rivoluzione industriale”. In questo senso, la digitalizzazione del mondo, i Big Data, l’informazione, gli algoritmi, l’automazione e *l’Internet of Things* contribuiscono a definire progressivamente un’atmosfera di artificializzazione della vita umana, che trasforma i modi in cui gli esseri umani si relazionano con sé stessi, con gli altri esseri umani e animali non umani e con le cose.

4. Alcune esperienze inquietanti

Esaminiamo due scenari adottando il concetto di “eccezionalità” come risorsa analitica, poiché l’ “eccezionalità” metterà in luce come sia la vertigine morale sia l’analfabetismo emotivo possano rappresentare due disposizioni mentali diverse, ma complementari a fronte della proliferazione e dell’emergere di algoritmi e sistemi di IA.

Il primo scenario è una scuola in Cina che ha utilizzato telecamere, chip di tracciamento e fasce per la lettura delle onde cerebrali (fasce per la testa con un sistema di colori che permettono di decifrare il livello di attenzione dei bambini), non tanto per valutare i contenuti o le abilità apprese dagli studenti in una determinata materia, ma per studiare direttamente il loro comportamento, ovvero il loro grado di attenzione e concentrazione. Attraverso telecamere che catturano i movimenti facciali e i cui risultati vengono inviati a genitori e insegnanti, viene stabilita una meticolosa sorveglianza e un controllo dei movimenti del corpo al fine di migliorare le loro prestazioni di apprendimento (Liu 2019).

Il secondo scenario è rappresentato dal sistema legale americano, dove tribunali e polizia hanno fatto ampio uso dell’IA nell’amministrazione della giustizia per prevedere i crimini, in uno sforzo massiccio per combattere e prevenire la criminalità, nonché per migliorare l’efficacia e l’efficienza del processo decisionale giudiziario. Un esempio è rappresentato dal caso di Eric Loomis, nel 2016, per il quale il sistema di giustizia penale nordamericano ha utilizzato il software Compas per elaborare un rapporto con l’assistenza dell’IA, classificando Loomis come “ad alto rischio per la comunità”. La pena inflittagli è stata più alta del previsto, dato che è stato condannato a 5 anni di carcere e a 6 anni di

libertà vigilata. Nonostante la sentenza sia stata impugnata, la Corte Suprema del Wisconsin non ha ammesso alcuna replica: ha concluso infatti che nel processo di predisposizione del verbale che ha portato alla sentenza sono stati utilizzati “solo dati pubblicamente disponibili e dati forniti dall'imputato che egli stesso avrebbe potuto negare, spiegare o verificare” (Criminal Law 2017).

Questi due scenari, che possono sembrare fittizi ma non lo sono, hanno come caratteristica principale l'uso dell'intelligenza artificiale e come denominatore comune il potere di trasformare la realtà che essa comporta. Gli esempi delineati condividono la caratteristica dell'esperienza della “eccezionalità” o dello “straniamento”. Con il termine “eccezionalità” o “straniamento” ci riferiamo a fenomeni, pratiche o oggetti che si presentano alla nostra percezione rappresentando una sorta di sfida, perché resistono a essere compresi secondo i più tradizionali schemi cognitivi.

Entrambi gli esempi menzionati implicano un livello significativo di artificialità, in cui oggetti intelligenti, algoritmi e dati proliferano nella vita di tutti i giorni, assumendo il controllo della situazione. Ogni scena non è un a priori o qualcosa di già dato nella realtà, ma il risultato di un dispositivo tecnologico che dispone la scena in un certo modo fino a farla diventare un mezzo meccanico e artificiale.

La sensazione di straniamento suscitata dal caso della scuola in Cina deriva, tra l'altro, dall'utilizzo di un oggetto quotidiano come la “fascia” in un contesto che è, per l'appunto, un centro didattico, ma che, proprio per questo, la fa risultare inquietante. La “fascia intelligente”, collegata a telecamere, computer e chip, non ha lo scopo di abbellire, ma piuttosto di monitorare e fornire dati sull'attenzione e sulla concentrazione degli studenti. La dimensione inquietante è legata al fatto che un oggetto di uso quotidiano venga utilizzato per monitorare e controllare il comportamento umano e che questo venga giustificato con fini educativi. Quando gli oggetti intelligenti o “non cose” vengono collegati all'IA e a Internet, diventano sistemi informativi con cui comunichiamo e interagiamo (Han 2021a). In questo caso, il rapporto con le persone, gli oggetti e l'ambiente è mediato dal calcolo e dalla precisione degli algoritmi.

Nell'esempio giuridico, la cosa strana consiste nel dare risalto ai dati personali e privati più che a quelli dei giudici e degli imputati, perché

l'informazione estratta, affinata e analizzata da un algoritmo è la chiave che permette di emettere la sentenza penale: "Come individui, non agiamo come semplici destinatari dei frutti di queste analisi di dati: al contrario, siamo i dati che vengono utilizzati per prendere decisioni su di noi, a volte con profonde conseguenze" (Kearns e Roth 2019, p. 12). È sempre più comune che la comunicazione avvenga da macchina a macchina senza la partecipazione umana (Floridi 2022, p. 25).

I due esempi precedenti mostrano come gli esseri umani siano subordinati alle decisioni algoritmiche, al punto da essere considerati una sorta di esperimento tecnologico. Si può parlare di "sperimentare" perché l'algoritmo agisce attraverso un potere che plasma il comportamento degli individui, in particolare di quelli presi di mira dall'intervento. Gli esseri umani non sono solo soggetti che producono oggetti tecnici, ma anche oggetti e bersagli di interventi tecnologici, poiché l'intelligenza artificiale e gli algoritmi diventano macchine che estraggono dati da noi.

I contesti degli "algoritmi educativi" e della "giustizia algoritmica" non solo stabiliscono nuove pratiche – come la sorveglianza, il controllo, la gestione e la previsione –, ma trasformano anche la realtà, creano nuovi ambienti, hanno un impatto sulla vita delle persone, rimodellano il modo in cui ci relazioniamo con le cose e le persone, ristrutturano le istituzioni e ridefiniscono gli obiettivi. Gli esempi presentati non sono solo familiari, dato che si verificano nella vita quotidiana delle relazioni umane e sociali, ma comportano anche l'"eccezionalità" di porre le relazioni con le non-cose, le informazioni, gli algoritmi e Internet su un piano diverso. Ora, cosa succede quando il modo in cui ci relazioniamo con noi stessi è mediato da oggetti tecnologici guidati dall'intelligenza artificiale e da algoritmi, anche in questioni vitali come la cura di sé? Come vedremo presto, l'IA si sta ormai introducendo in spazi personali e intimi come, ad esempio, la solitudine, mettendoci di fronte agli oggetti tecnici in una posizione di vertigine morale e di analfabetismo emotivo.

5. Vertigine morale e analfabetismo emotivo di fronte all'IA

Lo scenario seguente è il risultato dell'applicazione dell'IA alla vita e alle esigenze sociali e private delle persone. Realbotix© offre umanoidi

dotati di un “tocco umano”, sia di sesso maschile che femminile, e può metterli in connessione con varie piattaforme di IA. I clienti possono creare un design personalizzato per ciascuno di essi. Sebbene i robot siano programmabili per scopi commerciali, come consulenze mediche o assistenza accademica, un'IA degna di nota per le persone single è chiamata Aria ed è stata sviluppata per interagire con gli esseri umani che soffrono di solitudine (Kalil 2024). Questa applicazione solleva importanti interrogativi su come la tecnologia influenzi le nostre emozioni e le relazioni sociali e umane. Questo utilizzo dell'IA come compagna per alleviare la solitudine ci riporta all'ambientazione futuristica e immaginaria creata da Kazuo Ishiguro nel romanzo *Klara e il sole*, in cui viene presentata la storia di Klara, un androide che è anche la protagonista e la voce narrante della storia. Nel romanzo, l'androide ha il compito di prendersi cura di una giovane donna, aiutandola a liberarsi dalla tristezza e dalla malattia che la affliggono (Ishiguro 2021).

In questo caso, che potremmo definire come “cura algoritmica”, l’“eccezionalità” è legata ad alcuni aspetti particolari: in primo luogo, nel tipo di relazione che si instaura tra un essere umano e un oggetto tecnologico che non ha lo *status* di persona e che opera secondo la logica del calcolo algoritmico. È curioso che dall'interazione tra un essere umano e un androide possa nascere una relazione emotiva, in cui gli esseri umani vengono sostituiti da ruoli di cura, compassione, amore ed empatia. L'androide si relaziona con l'umano come se fosse una persona o un animale domestico, in un tipo di legame che non è, in senso stretto, affettivo, ma piuttosto di sostituzione o simulazione. In secondo luogo, la pratica assistiva che si sta affermando nella vita di tutti i giorni è strana, ma l'androide ha lo scopo di impedire alle persone di sentirsi sole. Tuttavia, per tutto questo tempo, il futuro dipende sempre dalla tecnologia. Per aumentare la sensazione di benessere, le persone trovano supporto emotivo in un compagno virtuale che offre consigli personalizzati su richiesta.

Sebbene sia vero che l'IA può aiutare gli utenti a gestire le proprie emozioni, modificando i modelli di pensiero negativi, resta da chiedersi se, in senso stretto, queste applicazioni possano essere considerate empatiche, in quanto si tratta di strumenti complementari che non potranno mai sostituire l'intervento umano. Con i sistemi di IA stiamo

assistendo a una colonizzazione di tutte le attività umane da parte di algoritmi e a un'implementazione sempre più ampia della pratica di delega del processo decisionale alla tecnologia? In che misura la nostra società, sempre più cablata e wireless, è tecnologicamente interconnessa, ma al contempo disconnessa dagli altri e dalla natura? In che misura l'IA può erodere la nostra capacità decisionale?

L'eccezionalità dello scenario della "cura algoritmica" fa emergere una predisposizione riguardo al posizionamento degli esseri umani nei confronti dell'IA. Si tratta di un modo per prepararsi internamente ad assimilare qualcosa di "troppo grande", come l'IA, e che adotta due esperienze: la vertigine morale e l'analfabetismo emotivo.

La società odierna è un sistema socio-tecnologico complesso, costituito da interazioni tra persone, cose, dispositivi, dati, algoritmi e robot. Ecco perché noi cittadini ci troviamo di fronte alla crescente esplosione di artefatti altamente interconnessi e di "dispositivi intelligenti", situazione che ci fa sentire sopraffatti dal punto di vista morale e ci rende analfabeti dal punto di vista emotivo. Inoltre, viviamo quotidianamente immersi nel fenomeno della digitalizzazione del mondo e in un regime informativo in cui l'accesso alle informazioni è sinonimo di potere (Han 2021b).

La vertigine morale è provocata dal disagio e dal senso di instabilità che proviamo quando siamo combattuti tra il fascino di un oggetto tecnologico programmato per prendersi cura delle persone e la paura che l'empatia, l'amore e la compassione vengano completamente sostituiti. Questo è dovuto all'ansia generata dalla tecnologia e dalla sua funzione assistenziale, nonché al fatto che la tecnologia sta sostituendo la realtà. La vertigine morale consiste nel non sapere quale strada prendere. Se si sceglie di accettare la tecnologia nel suo ruolo assistenziale, allora si è sulla strada della neofilia, ovvero dell'accettazione della novità e dell'innovazione. Tuttavia, se si sceglie di negarlo, si intraprende la strada della neofobia, ovvero della paura e dello scetticismo nei confronti della tecnologia.

L'analfabetismo emotivo è il risultato dell'incapacità di interpretare i processi e le pratiche che l'IA instaura nella nostra vita quotidiana e che influenzano le nostre questioni personali e intime. Utilizzando l'assistente virtuale con indifferenza e senza riflettere sulle conseguenze

dell'umanizzazione dell'androide, l'utente rischia di perdere la capacità di riflettere. D'altra parte, potrebbe perdere la capacità di prendere decisioni e di articolare il proprio spazio di solitudine, dovendo ricorrere a un artificio. L'analfabetismo emotivo è causato da due fattori. Il primo è dovuto all'ipertrofia dell'IA, perché gli algoritmi potrebbero diventare così invasivi da sovrastarci completamente, offuscando anche la nostra capacità di comprensione. La seconda ragione è dovuta all'atrofia del desiderio, che ci rende incapaci di prendere decisioni e ci lascia in una posizione di ricezione passiva e acritica della tecnologia.

Pertanto, una prima sfida che l'affermarsi dell'IA nel contesto degli affari privati e intimi, legati alla sfera dei bisogni e delle relazioni personali, comporta è quella di imparare a porre dei limiti all'interferenza arbitraria nella vita privata, perché i sistemi di IA sono emersi nel cuore degli affari umani, sia privati che pubblici, e nel contesto di un "regime informativo" il cui potere non è coercitivo, ma "intelligente" e sottile nel controllare il comportamento (Han 2021b). I personal computer, i cellulari, Internet e i social media non sono solo strumenti che ci aiutano nelle attività quotidiane, ma dispositivi che ci usano come fonti di informazione. Ad esempio, lo smartwatch che misura la frequenza cardiaca, monitora il polso e misura la pressione sanguigna; la casa intelligente che registra meticolosamente la vita di tutti i giorni; l'"aspirapolvere intelligente" che mappa la casa mentre la pulisce e il letto intelligente che monitora i pensieri durante il sonno (Han 2021b).

Un secondo rischio è rappresentato dall'accentuarsi del divario tra la mancanza di conoscenza da parte degli utenti su come funziona l'IA e la crescente complessità che caratterizza i sistemi di IA. In *L'uomo è antiquato*, Günther Anders ipotizza l'esistenza di un divario prometeico tra l'uomo e il mondo tecnico, ovvero di un'impossibile articolazione o a-sincronia tra uomo e macchina, in quanto l'uomo non riesce a tenere il passo delle macchine, generando un divario tra agire e sentire, fare e rappresentare; l'apparato prodotto e il corpo dell'uomo (Anders 2007). Ecco perché è essenziale sviluppare metodologie e pedagogie che educano le persone riguardo al potere, alla responsabilità e all'importanza di prendere decisioni informate sull'IA.

Un terzo rischio è che la potenza di questa tecnologia raggiunga gli stati emotivi delle persone, influenzando le nostre abitudini di consumo

e i nostri stili di vita fino a renderli troppo dipendenti e influenzati psicologicamente dalle informazioni consumate in rete. Attraverso le piattaforme digitali, i social network e l'Internet delle cose, da un lato vengono elaborati dati sui gusti, le preferenze e gli stati d'animo delle persone e, dall'altro, vengono fornite agli utenti informazioni su richiesta che si adattano alle loro esigenze. Il potere algoritmico agisce nella nostra vita privata, non in modo repressivo o negativo, ma stimolando le nostre emozioni, desideri, gusti e abitudini, utilizzando incentivi positivi e schermi lucidi e accattivanti, riuscendo così a invadere la nostra intimità e privacy senza che ce ne accorgiamo.

6. IA ed etica

L'avanzamento dei sistemi di IA, la loro applicazione in diversi ambiti e il loro impatto sulla vita pubblica e privata delle persone continuerà a verificarsi, il che ci porta a chiederci se si tratti di un destino che ci toccherà inesorabilmente e, se così fosse, come potremmo e dovremmo gestirlo. Poiché l'IA non esiste in un contesto isolato, ma è applicata in un ambiente unico, è essenziale considerare i quadri etico-giuridici o gli strumenti che ci possono aiutare a esaminare in modo permanente il potere che tale tecnologia comporta. Non si tratta di fermare l'innovazione, ma di indirizzare tale rivoluzione tecnologica, cercando di bilanciare competitività e innovazione con responsabilità, impegno etico e sostenibilità ambientale e sociale.

Goethe, ne *L'apprendista stregone*, racconta la storia di un discepolo che, per evitare le faccende domestiche, ripete le parole magiche che ha sentito dal suo maestro per far sì che sia la scopa a svolgere i lavori domestici. Tuttavia, dimenticando le parole precise per fermarla, le cose sfuggono di mano. Cosa possiamo fare per evitare di ritrovarci in una situazione simile a quella dell'apprendista, ma questa volta con un'IA che sfugge al nostro controllo?

L'algoritmo non può diventare una sorta di "feticcio" avvolto da un'aura di mistero e di carattere enigmatico. In astratto, è una "serie di istruzioni specificamente progettate con grande precisione per affrontare un argomento specifico" (Kearns e Roth 2019, pp. 13–14). In sé, è

qualcosa di ovvio e banale, ma quando comincia a funzionare in modo autonomo e automatico, intrecciandosi con la vita intima e quotidiana delle persone, allora diventa qualcosa di più complesso e per nulla innocuo.

In primo luogo, quando si affrontano i temi dei pericoli e dei rischi legati all'IA, è fondamentale considerare che gli algoritmi sono il frutto del pensiero umano imperfetto e che i sistemi di IA rispecchiano interessi, pregiudizi, intenzioni e disuguaglianze socio-economiche ancora presenti nelle nostre società, nonché discriminazione, violenza, sete di potere, accumulazione e disinformazione. Pertanto, un livello di analisi e riflessione etica che non può essere ignorato è la dimensione biopolitica o "realistica" dell'IA, rispetto alla quale dobbiamo essere consapevoli del potere tecnologico che gli esseri umani hanno raggiunto.

L'IA non deve favorire solo gli interessi delle grandi aziende che devono recuperare i loro ingenti investimenti in server, chip, energia, infrastrutture, ma anche la società e l'ambiente con grande responsabilità e impegno etico. Come tutta la tecnologia, l'IA è un mezzo o uno strumento al servizio non solo degli interessi umani, ma anche dell'ambiente naturale. Le grandi aziende tecnologiche devono considerare che, sebbene ci siano cose che si possono fare dal punto di vista tecnologico, non è consentito farle dal punto di vista etico-legale.

Un uso improprio dell'IA si ha quando Stati, organizzazioni imprenditoriali o individui possono trarre vantaggio da informazioni personali relative alla salute fisica e mentale, ai sentimenti, alle abitudini, alla situazione economica, all'inclinazione sessuale, religiosa o politica per manipolare il comportamento umano. Questo tipo di rischio può essere presente nella creazione di profili di persone, nella classificazione di segmenti di popolazione come criminali o indagati e nella generazione di discriminazioni professionali, razziali, sessuali, religiose ed etniche dovute a pregiudizi nell'apprendimento automatico. Ecco perché è necessario sviluppare sistemi di protezione giuridica, tecnica o tecnologica che salvaguardino la privacy delle persone.

In secondo luogo, i poteri degli algoritmi e dei sistemi di IA devono essere accompagnati dalla responsabilità di saperli gestire, anche in situazioni di evidente valore e urgenza, come la gestione di crisi ambientali, economiche o sociali. Ciò richiede una serie di responsabilità.

Questo deve valere per gli sviluppatori (grandi aziende), gli utenti che implementeranno i sistemi di IA (persone fisiche o giuridiche, autorità pubblica, agenzie o organizzazioni) e i destinatari finali.

In terzo luogo, è necessario implementare e sottoporre a un esame costante le pratiche di trasparenza che devono essere presenti durante lo sviluppo, l'implementazione e il funzionamento dei sistemi di IA, in modo da poter stabilire la tracciabilità del loro funzionamento, i dati utilizzati per generarli e il modo in cui sono stati prodotti. La trasparenza implica definire chi è responsabile del funzionamento e delle decisioni prese dai sistemi di IA, proteggere i dati personali, garantire la privacy, segnalare l'origine dei dati utilizzati per addestrare i modelli di IA, rendere pubblici i possibili rischi nell'implementazione di tali sistemi e informare gli utenti quando sono esposti a contenuti creati dall'IA.

In quarto luogo, è necessario sviluppare una pedagogia e un'educazione critiche e riflessive che permettano alle persone di analizzare le pratiche e gli usi dell'IA e di anticipare gli effetti possibili di questa tecnologia, in particolare i rischi legati a discriminazioni che potrebbero essere incorporati negli algoritmi o collegati alla violazione dei diritti fondamentali. In breve, è necessario sviluppare metodologie e strumenti educativi che aiutino le persone a comprendere il modo in cui gli algoritmi prendono decisioni e quali dati utilizzano.

Infine, sebbene sia necessario un quadro etico-normativo che permetta alle leggi e ai regolamenti internazionali di articolarsi con gli standard regionali e locali, è importante farlo di pari passo con la "giustizia globale", nella cui articolazione si possono identificare tre componenti per la sua specificazione: 1) giustizia sostenibile per l'ambiente; 2) giustizia di democratizzazione e partecipazione; 3) giustizia geopolitica e strategica.

La prima dimensione permette di porre in luce che l'IA genera un costo energetico e un'impronta di carbonio con un impatto significativo sull'ambiente, motivo per cui essa deve essere pensata come ecologicamente sostenibile, rispettosa dell'ambiente, centrata sulle persone, attenta agli individui e ai gruppi più vulnerabili.

La seconda dimensione aiuta ad evidenziare il ruolo che i cittadini e la società devono svolgere nell'imparare a utilizzare queste tecnologie, garantendone un'adeguata conoscenza per poter prendere decisioni

informate e massimizzarne i benefici dove possibile, minimizzandone i rischi dove necessario. In una società in cui l'informazione gioca un ruolo strategico, non è auspicabile che le informazioni e le ricerche riguardanti l'IA rimangano nelle mani dei monopoli che le utilizzano esclusivamente a proprio vantaggio.

Il terzo aspetto è legato alla cosiddetta “nuova corsa agli armamenti” tra Stati Uniti e Cina per il predominio dell'IA, che coinvolge anche paesi come Gran Bretagna, Canada, Israele e Brasile. Ciò che è in gioco è la disputa sul nuovo ordine politico mondiale e regionale in cui l'IA, insieme ad altri sviluppi economico-tecnologici e militari, segnerà l'emergere e il consolidamento di nuove potenze egemoniche. Collegato a quanto sopra, non dal punto di vista politico bensì economico e di mercato, esiste il pericolo di monopolio da parte di grandi aziende private o aziende tecnologiche, denominate “Big Tech”, come Meta, Google, Microsoft, OpenAI, che finirebbero per relegare in secondo piano altri attori come le università. Di fronte a questo scenario, è cruciale una struttura di *governance* internazionale in grado di controbilanciare la tendenza alla centralizzazione dell'IA, sia nelle mani dello Stato che in quelle delle imprese.

7. Conclusioni

Conviviamo quotidianamente con dispositivi digitali, piattaforme, social network, assistenti virtuali e dispositivi smart che, da un lato, essendo utilizzati in modo simile, facilitano le nostre attività e si trasformano in componenti essenziali per vivere senza grandi inconvenienti; dall'altro lato, la loro logica di funzionamento è a noi sconosciuta. D'altra parte, si tratta di un insieme di realtà tecnologiche così diverse tra loro che ignoriamo la logica che le sottende. Inoltre, nella maggior parte dei casi, non siamo consapevoli di come funzionano i sistemi di IA, né del loro potenziale impatto sul nostro stile di vita. I sistemi di IA stanno cambiando l'essere umano in modo quasi impercettibile: siamo noi a creare la tecnologia, ma, paradossalmente, è la logica degli algoritmi a plasmare le nostre azioni. Lungi dall'essere una tecnologia neutrale, l'IA ha ripercussioni sulla vita degli individui e delle società, in

quanto non è né innocua né un male di per sé, ma piuttosto uno strumento o un mezzo con diversi livelli di complessità e molteplici ambiti di applicazione rispetto ai quali è richiesto un uso responsabile. Il suo utilizzo responsabile, prudente e trasparente, così come una *governance* adeguata e l'educazione critica e informata dei cittadini, tra le altre cose, determineranno se non diventeremo “carne di algoritmi” e se non saremo inghiottiti da sistemi informatici e algoritmici completamente automatizzati. Dobbiamo essere consapevoli del potere che gli algoritmi hanno di cambiare la realtà e del modo in cui possiamo concepirla. Ecco perché è importante riflettere e rivedere criticamente il modo in cui i sistemi di IA, nonché le pratiche e le abitudini da essi instaurate, modificano e potrebbero modificare il nostro rapporto con noi stessi, con gli altri, con la natura, con le cose e con le macchine stesse.

Riferimenti bibliografici

- ANDERS G. (1988) *Wir Eichmannsöhne. Offener Brief an Klaus Eichmann*, C. H. Beck, München.
- ANDERS G. (2002) *The Antiquiertheit des Menschen*, vol. I, C.H. Beck, München, (trad. it. *L'uomo è antiquato. Considerazioni sull'anima nell'epoca della seconda rivoluzione industriale* (Vol. 1), Bollati Boringhieri, Torino 2007).
- BARTEL M., M. HANKE e S. PETRIC (2024) *Crisis Identification and Prediction using Machine Learning: The Case of U.S. Regional Banks* (August 05, 2024). Disponibile al link SSRN: <https://ssrn.com/abstract=4688565> or <http://dx.doi.org/10.2139/ssrn.4688565> (ultimo accesso 23 gennaio 2025).
- BECK U. (2000) *La società del rischio. Verso una seconda modernità*, Carocci, Roma.
- BETTINI A. (2024) *Modelli matematici, supercomputer, IA: ecco la nuova meteorologia*, Zanichelli, Bologna. <https://aulascienze.scuola.zanichelli.it/multimedia-scienze/come-te-lo-spiego-scienze/satelliti-supercomputer-intelligenza-artificiale-meteorologia>
- BOSTROM N. (2014) *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, Oxford (trad. it. *Superintelligenza. Tendenze, pericoli, strategie*, Bollati Boringhieri, Torino 2018).

- COECKELBERGH M. (2020) *AI Ethics*, The MIT Press, Cambridge (MA).
- CRIMINAL LAW (2017) *Sentencing Guidelines — Wisconsin Supreme Court Requires Warning Before Use of Algorithmic Risk Assessments in Sentencing*. — “*State v. Loomis*”, 881 N.W.2d 749 (Wis. 2016), “Harvard Law Review”, 130(5): 1530–1537, <http://www.jstor.org/stable/44865547> (ultimo accesso 12 febbraio 2025).
- DEDICATED (2020) *La cibernetica alle origini dell'Artificial Intelligence* <https://www.dedicated.world/cibernetica-origini-artificial-intelligence/>.
- FLORIDI L. (2022) *Etica dell'intelligenza artificiale*, Raffaello Cortina, Milano.
- GASTALDO T., C. CACCIAMANI e A. DE SAVINO (2024) *Le prospettive dell'IA in meteorologia*, “*Ecoscienza*”, 4: 29–31.
- HAN B.C. (2021a) *Undinge: Umbrüche der Lebenswelt*, Ullstein, Berlin (trad. it. *Le non cose. Come abbiamo smesso di vivere il reale*, Einaudi, Torino 2022).
- . (2021b) *Infokratie. Digitalisierung und die Krise der Demokratie*, Matthes und Seitz, Berlin (trad. it. *Infocrazia. Le nostre vite manipolate dalla rete*, Einaudi, Torino 2023).
- ISHIGURO K. (2021) *Klara and the Sun*, Faber and Faber, Londra (trad. it. *Klara e il Sole*, Einaudi, Torino 2022).
- KALIL M. (2024) *Meet Aria, Realbotix's Ultra-Realistic Humanoid Robot with “Relationship-Based AI”*, <https://mikekalil.com/blog/aria-realbotix-ai-robot/>
- KEARNS M. e A. ROTH (2019) *The Ethical Algorithm: The Science of Socially Aware Algorithm Design*, Oxford University Press, New York.
- KIRAN A., R. NARAYANA e M. TALAWAR (2020) *Food Crisis—How Artificial Intelligence and Machine Learning Are Solving Humanity's Greatest Challenge — A Review*, “Journal of Computational and Theoretical Nanoscience”, 17,(9–10): 3839–3843. doi: <https://doi.org/10.1166/jctn.2020.9005>
- KNUDSEN J., U. GHAFAR, R. MA e A. HUNG (2024) *Clinical Applications of Artificial Intelligence in Robotic Surger*, “J Robot Surg”, 18(1):102. doi: 10.1007/s11701-024-01867-0. PMID: 38427094; PMCID: PMC10907451.
- LIU L. (2019) *China: A Headband for Your Thoughts?*, “EE/Times” <https://www.eetimes.com/china-a-headband-for-your-thoughts/> (ultimo accesso 21 marzo 2025).
- MCCARTHY J., M.L. MINSKY, N. ROCHESTER e C.E. SHANNON (2006) *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*,

- August 31, 1955, "AI Magazine", 27(4): 12. <https://doi.org/10.1609/aimag.v27i4.1904> (ultimo accesso 21 ottobre 2024).
- McKINSEY & COMPANY (2023) *What is AI?* Disponibile al link <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-ai> (ultimo accesso 12 giugno 2025).
- REGULATION (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 *Laying Down Harmonised Rules on Artificial Intelligence and Amending Regulations* (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance).
- SANDEL M. (2007) *The Case Against Perfection. Ethics in the Age of Genetic Engineering*, Belknap Press, Cambridge (MA).
- TURING A.M. (1950) *Computing Machinery and Intelligence*, "Mind" 49: 433–460.
- VÁZQUEZ F. (2005) *Tras la autoestima. Variaciones sobre el yo expresivo en la modernidad tardía*, Gakoa, Donostia–San Sebastián.
- VIDALI M. (2024) *Intelligenza Artificiale in Medicina: implicazioni e applicazioni, sfide e opportunità*, "Biochimica clinica", 48(2): 129–141. doi: 10.19186/BC_2024.020
- WIENER N. (1948) *Cybernetics; or Control and Communication in the Animal and the Machine*, MIT Press, Cambridge (MA).