

PIERGIORGIO DONATELLI

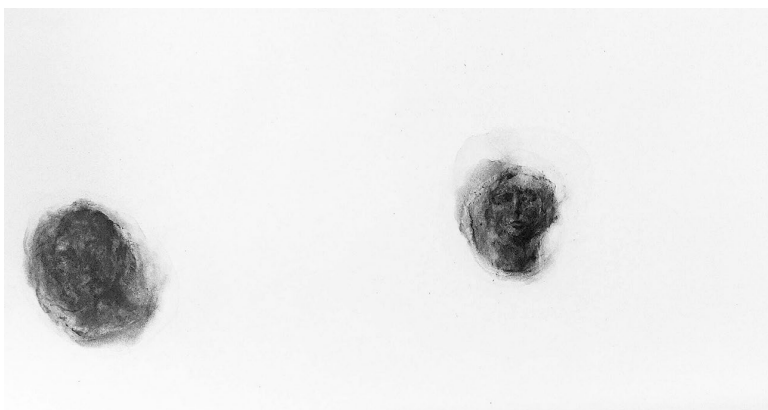
INTELLIGENZA ARTIFICIALE, DEMOCRAZIA E VALORI UMANISTICI

1. *La svolta tecnologica dell'intelligenza artificiale*
2. *Un umanesimo digitale?*
3. *IA e medicina*

ABSTRACT: *ARTIFICIAL INTELLIGENCE, DEMOCRACY AND HUMANISTIC VALUES*

The article examines various aspects of the technological turn brought about by artificial intelligence from the perspective of the goals that inspire humanistic and democratic ideals. The new technologies directly effect the humanistic ideal that represents individuals as progressive beings whose self realization requires mutual education (J.S.

Mill). Cognitive technologies offer extraordinary tools that advance research and intervention capabilities, but the way in which they are spreading also shows their potential for levelling and the tendency towards passivity. Education is a crucial key notion that needs to be introduced into the discussion.



1. La svolta tecnologica dell'intelligenza artificiale

L'intelligenza artificiale sta trasformando nel profondo le nostre società, coinvolgendo tutte le attività professionali, dalla medicina alla produzione, la finanza, la difesa, la ricerca, e sta cambiando anche la vita quotidiana. Nel rapporto con la rete, che ha già qualche decennio di storia alle nostre spalle, si è insinuata l'intelligenza artificiale, con cui parliamo, che ci consiglia e suggerisce e che imita attività intellettuali sofisticate come la scrittura di testi complessi con ChatGPT.

L'ingresso dell'IA nelle nostre società va inserito nel quadro che va avanti da alcuni secoli con la scienza moderna, l'illuminismo, la rivoluzione industriale, la crescita democratica delle società. In questo senso le nuove tecnologie fanno parte della linea progressiva interna alle società civili, anche se va osservato che

lo sviluppo tecnologico ha una sua coerenza indipendente dagli altri fattori, come la democrazia e l'organizzazione sociale. Valica i sistemi sociali e non deve sorprendere pertanto che sui temi dell'IA abbiamo tre modelli regolativi differenti rappresentati dalle due democrazie, gli Stati Uniti e l'Unione Europea, e dalla Cina.

Per riflettere sul significato di questa repentina svolta realizzata dall'IA conviene richiamare i due episodi precedenti situati nel Novecento, quando la tecnologia ha suscitato una riflessione etica e sociale che chiamava in causa le sorti e le direzioni della società nel suo complesso¹. Il primo episodio a ridosso della Seconda guerra mondiale è stata la riflessione sulla bomba atomica e sulla medicina praticata dai nazisti, che ha dato origine a un pensiero fondato sulla paura verso la tecnologia, che rappresentava gli esseri umani come fanciulli che giocano con delle armi da fuoco. La tradizione filosofica tedesca, tra gli altri con Arendt, Anders e Jonas, ha insistito sul male che gli esseri umani sono in grado di arrecare e puntava il dito sul pericolo fatale insito nella tecnologia. Veniva rappresentato il potere distruttivo della tecnica che metteva a repentaglio le condizioni basilari della vita umana sul pianeta e al contempo si indicava il nucleo etico di fondo, la soglia della dignità sotto la quale gli umani si trasformano in mostri, come era accaduto con il nazismo.

Un episodio successivo è rappresentato dalla vicenda della bioetica che, perlomeno ai suoi esordi negli anni Settanta, riconduceva i cambiamenti tecnologici e le nuove scelte etiche agli scopi delle società democratiche. In questo contesto, il pensiero sulle innovazioni della medicina e della ricerca non portava alla difesa di una tradizione etica di lunga durata, pervertita dall'uso di una tecnologia costitutivamente fuori controllo, ma incoraggiava al contrario un ripensamento del quadro etico tradizionale. Grandi trasformazioni che prendevano piede sia sul piano tecnologico, come il trapianto d'organo, le macchine di sostegno vitale, la Fivet e

¹ P. Donatelli, *La filosofia e la vita etica*, Einaudi, Torino 2020, cap. VIII.

la tecnica del DNA ricombinante, sia su quello concettuale ed etico, come la nuova definizione di morte e il consenso informato, venivano accolte come occasioni di miglioramento della qualità della vita e di espansione delle possibilità di scelta. Nel primo episodio l'etica si poneva a difesa di una umanità violata dagli orrori nazisti e messa a rischio da una tecnologia che minacciava di distruggerla. Nel secondo caso invece l'etica accoglieva i cambiamenti rivisitando la sua tradizione. La bioetica di fatto segnalò il crollo della lunga tradizione del giusnaturalismo medievale che aveva inscritto nella natura la trama costitutiva dell'umanità: la nascita, la morte, la sessualità – una tradizione che la modernità aveva in effetti solo scalfito nelle questioni della bioetica, poiché si era occupata principalmente della società e non delle forme di vita. Più tardi, già a partire dagli anni Ottanta con le due figure cruciali del presidente Reagan e del pontefice Giovanni Paolo II, quando la bioetica era diventata una questione politica, la rivoluzione bioetica sarà tradotta nei termini dell'attacco ai fondamenti della natura umana, segnando una contesa che arriva ai nostri giorni. Tuttavia, nel frattempo, nelle società si era insediata questa linea intellettuale e sociale che guardava ai progressi della tecnologia biomedica come occasioni di miglioramento della qualità della vita e di espansione della libertà².

Le questioni dell'IA vanno poste in continuità con le vicende della bioetica perché si tratta di miglioramenti e progressi interni alle società democratiche, anche se coinvolgono un attore cruciale come la Cina. Tuttavia, le tecnologie cognitive stanno cambiando nel profondo e con grande rapidità il complesso delle attività umane. Non possono essere caratterizzate solo come delle trasformazioni interne al quadro delle forme di vita e dei valori democratici

² P. Donatelli, *Dove va La bioetica?*, in «La Società degli Individui», XXIV, 72, 2021, pp. 26-44; Id., *Dagli anni settanta al presente: La bioetica, Le libertà, La politica*, in «Bioetica. Rivista interdisciplinare», XXX, 1-2, 2022, pp. 69-90.

poiché hanno un potenziale di cambiamento che promette di rendere obsolete attività e modi di vivere e di portare sulla scena situazioni umane inedite.

2. Un umanesimo digitale?

L'IA condivide pertanto aspetti che appartengono alle due circostanze più recenti in cui la tecnologia ha provocato un ripensamento dell'orizzonte umano. Lasciando da parte il caso della Cina e concentrandoci sulle maggiori aziende dell'Information Technology che operano negli Stati Uniti, l'IA va certamente situata dentro lo sviluppo di società capitalistiche e democratiche. Risponde a bisogni e interessi dei loro cittadini, non è imposta dall'alto da qualche potere totalitario. Tuttavia ci sono elementi di incertezza. È un motore di trasformazione molto rapido e imponente della società e delle attività umane. Sulla direzione del cambiamento influisce la conversazione pubblica nelle nostre società, ma gli scopi non sono elaborati attraverso una riflessione democratica condivisa ma sono dettati dagli interessi di mercato delle grandi imprese nel settore. Inoltre, se valutiamo la direzione di marcia, la situazione appare incerta. L'obiettivo è chiaramente il progresso delle attività professionali e della ricerca nonché il miglioramento della vita quotidiana in diversi campi, come informarsi, comunicare, coordinarsi per una impresa comune. D'altra parte il prezzo è una trasformazione radicale di *tutto* e un cambiamento notevole del rapporto con noi stessi, con il nucleo dei valori umanistici che opera sia nelle attività dei professionisti sia nella vita di ogni giorno.

Cosa intendo per valori umanistici?³ Il richiamo all'umanesimo è pieno di insidie poiché può travestire il bisogno di ancorare le proprie incertezze nei confronti di un mondo che cambia a valori che appaiono assicurati una volta per tutte, nascondendo nell'umano un nucleo di scelte che esclude altre possibilità. Non dovremmo

³ P. Donatelli, *La filosofia e la vita etica*, cit., cap. I.

rinunciare però a richiamarci all'idea di umanesimo, come idea educativa sviluppata dalle discipline umanistiche e caratteristica di scelte politiche che hanno valorizzato lo sviluppo personale. Esso indica il progetto di dare forma alla propria vita in modo libero e creativo, educandoci secondo modelli diversi che costituiscono degli esperimenti di vita, come li chiama John Stuart Mill, che mettono in luce lati e temperamenti diversi nelle differenti sfere, nei legami umani, nella visione politica, nella religione e negli ideali personali, dentro una società rispettosa e solidale⁴. Siamo tutti potenzialmente esemplari per gli altri e in primo luogo per noi stessi, cercando di rintracciare il sé migliore che ci parla e ci guida. Sto evocando una linea di pensiero precisa e un canone di cui fanno parte tra gli altri autori come Goethe, Mill, Emerson e Thomas Mann, che scopre la democrazia americana con l'avvento del nazismo. Queste idee hanno dato forma alla nostra concezione della società democratica intesa in senso non solo istituzionale ma anche come uno spazio comune pluralistico e solidale, dove le persone trovano occasione e incentivo a educarsi reciprocamente, provando a dare forma alla propria vita in modo personale.

Su questa idea del sé e della società influiscono certamente i cambiamenti impressi dalle tecnologie legate all'IA. Consideriamo alcuni aspetti. 1) Questa concezione dell'educazione e della società richiede uno spazio di elaborazione anche solitaria e comunque non sempre sorvegliata della propria persona, un'idea che colleghiamo al concetto di privacy ma che è più estesa: dobbiamo essere lasciati liberi di sperimentare lati del nostro carattere e scelte personali in modi e contesti che non ci siano sempre riferiti per chiedercene conto e per disapprovarci. Non possiamo essere sempre sul palcoscenico, né sempre controllati. Tuttavia l'IA funziona con

⁴ J. S. Mill, *La libertà* (1859), in Id., *La libertà. L'utilitarismo. L'asservimento delle donne*, a cura di E. Lecaldano ed E. Mistretta, tr. it. Rizzoli, Milano 1999.

grandi dati che devono essere prelevati continuamente in larga parte senza il nostro consenso, lasciando tracce delle nostre attività ovunque. Nell'epoca digitale la nostra interazione con il mondo e gli altri non è mai solitaria o privata come immaginavano Rousseau e Mill, quando esortavano a incontrare la natura da soli o quando tematizzavano l'amicizia come forma di relazione e di crescita reciproca diversa dall'incontro agonistico sulla scena sociale. Entrando in rete siamo anche presi nella rete, profilati, sollecitati, guidati attraverso algoritmi che definiscono uno spazio che ci costringe senza lasciarsi comprendere e quindi criticare.

2) Il rapporto con i diversi ambiti della vita è quello dell'educazione del sé, della cultura, dell'immaginazione, dell'esperienza del corpo, è un rapporto con idee, schemi intellettuali e modelli di vita che possiamo conoscere e criticare. Il rapporto è riflessivo e consapevole, e ciò vale in tutte le sfere: da quella politica all'autonomia personale. La concezione moderna delle istituzioni politiche ce le presenta come costitutivamente trasparenti e l'immagine così influente del contratto sociale ce le restituisce precisamente come assetti che in ogni momento possono essere confermati o rifiutati da ciascuno di noi: essi dipendono dall'assenso che diamo a schemi cooperativi che abbiamo compreso e accettato. In modo analogo l'idea moderna di autonomia, da Kant a Mill, richiede che le regole e i principi a cui sottomettiamo la nostra condotta, morali e religiosi, personali ed estetici, siano il risultato del nostro scrutinio riflessivo. Tuttavia, questa relazione con gli schemi intellettuali che governano la società e la nostra persona è messa in discussione nei limiti in cui con le nuove tecnologie stabiliamo rapporti di fiducia cieca (la questione della black box), che può essere attenuata ma che è costitutivamente rivolta a risultati di cui non possiamo comprendere la catena di ragionamento, proprio perché non si è svolto nessun ragionamento, ma degli algoritmi hanno predetto il

risultato statisticamente più accurato. Questa è la differenza fondamentale che distingue la ricerca precedente fondata su sistemi simbolici e il machine learning che ha rivoluzionato il campo dell'IA. Inoltre, con queste tecnologie stabiliamo un rapporto diverso in un altro senso: questi dispositivi in taluni casi ci conoscono meglio di quanto ci conosciamo, ci offrono suggerimenti che sono quelli che avremmo scelto se solo ci avessimo pensato bene⁵. Ma di che tipo di conoscenza si tratta? Non è una conoscenza che possiamo riconoscere come nostra, come nel rapporto educativo e di amicizia con una persona o con un testo, o nel rapporto politico con una istituzione che ci rappresenta. Questa prospettiva esterna sulla mia vita non è un punto di vista che io posso assumere, perché non è proprio un punto di vista, e le sue istruzioni e consigli non mi sono rivolti; io non sono affatto una persona in questa interazione ma la configurazione di un insieme di dati. Se immaginiamo di essere costantemente conosciuti e istruiti da dispositivi cognitivi stiamo immaginando una società che gradualmente perde la sostanza delle proprie pratiche intellettuali e delle relazioni personali e le conserva solo per finta, come quando diciamo che stiamo parlando con ChatGPT. A lungo andare però perderemmo proprio il senso e la motivazione di proseguire con queste attività e l'umanità cambierebbe davvero.

3) Queste tecnologie richiedono pertanto una nuova educazione che cambia i termini del modello umanistico, che vale ugualmente nelle discipline umanistiche e in quelle scientifiche, secondo cui conosciamo ciò che rappresentiamo secondo uno schema intellettuale che abbiamo elaborato⁶. Esse quindi possono diseducarci indebolendo la fiducia in noi stessi, e quindi in fondo nel pensiero, spingendoci ad affidarci a procedimenti che sappiamo che funzionano, anche se non sappiamo come. D'altro canto, possiamo interagire con

⁵ Y. N. Harari, *21 lezioni per il XXI secolo*, Bompiani, Milano 2018, pp. 89-98.

⁶ D. Weinberger, *Everyday Chaos. Technology, Complexity, and How we Are Thriving in a New World of Possibility*, Harvard Business Review Press, Boston, Mass. 2019.

queste tecnologie incorporandole nel modello di conoscenza tradizionale, affidando loro dei compiti e usando i risultati a cui pervengono. Così il programma AlphaFold progettato per predire la struttura tridimensionale delle proteine è stato utilizzato agli inizi del 2020 per individuare il coronavirus SARS-CoV-2. Il progresso della ricerca scientifica attraverso l'uso dell'IA è una delle promesse più emozionanti nel variegato panorama delle trasformazioni prodotte dall'impiego delle nuove tecnologie.

4) Gli sviluppi dell'IA alimentano la fiducia in un balzo in avanti mai visto prima nella ricerca scientifica e nella tecnologia, e quindi in un miglioramento inedito delle nostre capacità di conoscenza e di intervento. La prospettiva di una tale rivoluzione si accompagna però a una reazione opposta, che è la sfiducia che l'umanità sia in grado di guidare una svolta di queste proporzioni. La specie umana, ferma alla dotazione biologica emersa dalla selezione naturale, o meglio in movimento con i tempi dell'evoluzione delle specie, è impreparata ad affrontare le grandi sfide del presente, rappresentato da un mondo ipertecnologico e interconnesso minacciato dai cambiamenti climatici. Rovesciando le paure di Anders, si sostiene che abbiamo a disposizione una tecnologia che può risolvere i grandi problemi dell'umanità ma che l'umanità è incapace di usarla, vincolata a una dotazione psicologica antiquata⁷. Il discorso sul rischio della tecnologia si trasforma in una svalutazione dell'umanità e quindi in una lode indiretta del progresso tecnologico a discapito del più ampio sviluppo umano. Bisogna prestare attenzione alle insidie insite nella fascinazione verso la tecnologia che prende la forma anche della denuncia del pericolo imminente per l'esistenza stessa dell'umanità sul pianeta⁸, poiché tale fascinazione, ottimistica o

⁷ I. Persson, J. Savulescu, *Unfit for the Future. The Need for Moral Enhancement*, Oxford University Press, Oxford 2012.

⁸ Un quadro teorico influente è quello offerto dalla prospettiva dei rischi esistenziali: N. Bostrom, M. M. Ćirković (a cura di), *Global Catastrophic Risks*, Oxford University Press, Oxford 2008. Si veda la dichiarazione pubblicata nel maggio 2023 che mette in guardia dal rischio di estinzione firmata tra gli altri

pessimistica, esprime una profonda sfiducia nelle capacità cooperative dell'umanità e quindi negli strumenti della democrazia. La fascinazione verso l'IA può condurre perciò a svalutare la ricca tessitura di pratiche, regole e rapporti di fiducia che costituisce la democrazia. Non dimentichiamo che il motore di questo progresso, nelle nazioni democratiche e lasciando quindi da parte nazioni come la Cina che non intendono percorrere il quadro dell'umanesimo liberale e democratico, è governato da grandi aziende che hanno cominciato a sostituire i servizi pubblici, contribuendo a destrutturare l'impianto pubblico e democratico delle nostre società in coerenza con il quadro di sviluppo neoliberale in cui ci situiamo. I servizi pubblici fondamentali dipendono dalle grandi aziende e in taluni casi l'IA ha cominciato a entrare nella gestione diretta di questi servizi. La grande rivoluzione delle tecnologie cognitive è innanzitutto una nuova immensa invenzione capitalistica: anche se è molto altro, e non vanno percorse le analisi riduzionistiche della realtà sociale, non va sottovalutato l'aspetto manifesto dei grandi cambiamenti in atto che è quello di una espansione inedita dei mercati in condizioni di regolazione pubblica minimale⁹.

Le nuove tecnologie pongono pertanto delle sfide importanti ai valori umanistici e al quadro democratico inteso come una forma di vita associata. Per riuscire a pensarle come uno sviluppo della tradizione umanistica abbiamo bisogno di educazione. Esse vanno coniugate con l'educazione che ci rende consapevoli che un robot sociale è comunque un robot e non un umano, che i risultati a cui arriva un procedimento che opera su grandi dati hanno un carattere di cecità che va bilanciato con altri meccanismi esplicativi, che

dai capi di grandi aziende dell'IT come Bill Gates e Sam Altman (<https://www.safe.ai/statement-on-ai-risk#open-letter>) e nel marzo 2023 la richiesta di moratoria dopo la diffusione di ChatGPT-4 firmata tra gli altri da Elon Musk (<https://futureoflife.org/open-letter/pause-giant-ai-experiments/>).

⁹ E. Morozov, *The True Threat of Artificial Intelligence*, in «The New York Times», 30 giugno 2023.

l'inventività del machine learning non solo non è la nostra inventività che riguarda l'educazione e l'esperienza ma la imita a partire da un bacino di dati determinati, senza la creatività individuale che potremmo apportare noi, e che nella vita quotidiana possiamo fare a meno di mediare tutti i nostri contatti con la realtà attraverso dispositivi senza con questo abbracciare un'opzione di drop out dalla società digitale. E abbiamo bisogno di educazione in chi sviluppa e produce i dispositivi cognitivi; gli scopi umanistici devono poter pervadere il design del mondo digitale in cui abitiamo in modo sempre più continuo ed esteso. Tuttavia la linea di tendenza intrinseca allo sviluppo diffuso e ineducato di queste tecnologie è l'addomesticazione con un abbassamento del tasso di riflessione e di esperienza in prima persona.

Il modello umanistico classico si è basato certamente su sistemi di valore che in ciascun caso potevano essere criticati e rifiutati (sono state criticate le religioni, le ideologie, i sistemi giuridici e così via), ma esso richiede il controllo riflessivo degli schemi intellettuali (la religione ragionevole di Locke, la sovranità che possiamo rivendicare come nostra nella tradizione del contratto sociale, la cultura che ha il potere di rappresentare il nostro agire sociale nella tradizione della teoria critica). Invece le nuove tecnologie impongono un modello diverso che non è intrinsecamente educativo poiché esse nascondono i processi attraverso i quali interagiscono con noi. In parte la questione riguarda la natura dell'IA attuale che non funziona con sistemi simbolici e in parte dipende dal mondo in cui è sviluppata e utilizzata, offerta in larga scala come oggetto di consumo e non come uno strumento che richiede studio e apprendimento.

Gli ideali umanistici che ho richiamato sono il frutto di un esperimento sociale storicamente determinato: promuovere gli spazi di libertà creando anche i mezzi per esercitarla con la produzione e la distribuzione della ricchezza, una libertà che va intesa come sperimentazione nel senso di Mill, libertà di essere eccentrici,

osservati dagli altri ma non oppressi dallo sguardo altrui, anzi incoraggiati a essere se stessi. Questa libertà ha bisogno di cultura, di esempi e di istruzione che le persone si impartiscono reciprocamente. Il modello educativo è quello in cui ciò che è autorevole e importante per noi è appreso come un lato della nostra persona che non abbiamo coltivato e che anzi abbiamo disdegnato per qualche motivo e che trova le parole che lo rendono interessante e perseguibile. Ci educiamo ritrovando nelle parole di un libro, di un film, di un amico, di un maestro le nostre stesse parole rigettate che tornano con una maestà alienata, per riprendere l'immagine di Emerson¹⁰. Secondo questo modello, ciò che ci educa stabilisce un rapporto interno con noi, sono parole, insegnamenti, esempi di vita che ci appartengono anche se li abbiamo rifiutati o nascosti (o proprio per questo). L'idea è quella della educabilità a qualcosa che già siamo. Ricordo che su questa base abbiamo costruito società democratiche che fanno leva precisamente sul fatto che tutto ciò che succede ci riguarda e può essere preso ad esempio. La democrazia ha come nemici l'indifferenza e il cinismo, le barriere tracciate una volta per sempre. Le nuove tecnologie sono una fonte indefinita di strumenti di educazione mai sperimentati prima, ma al contempo inducono una forma di affidamento e di fiducia cieca che non è educativa.

D'altro canto, il potenziale di passivizzazione delle tecnologie cognitive si allea a un complessivo spostamento, che si sta realizzando nelle nostre società, dall'educazione all'autonomia, che ci chiede di saper padroneggiare intellettualmente il mondo per trovare la nostra posizione, all'acquisizione di skills che ci rendono esecutori efficaci e meri trasmettitori di informazioni, richieste e stimoli¹¹. L'autonomia - intesa nel senso umanistico

¹⁰ R.W. Emerson, *Self-Reliance*, in Id., *The Major Prose*, ed. by R.A. Bosco e J. Myerson, Harvard University Press, Cambridge, Mass. 2015, p. 127.

¹¹ W. Deresiewicz, *Excellent Sheep. The Miseducation of the American Elite and the Way to a Meaningful Life*, Free Press, New York 2014; C. Bollas, *L'età dello smarrimento. Senso e malinconia*, Raffaello Cortina, Milano 2018.

ampio che ho introdotto – non è una proprietà ma una forma di vita, un modo di trattarsi e di educarsi reciprocamente. Il rapporto con l’IA non va stabilito nei termini di chi è più bravo a eseguire un compito – i dispositivi cognitivi vinceranno sempre, è solo una questione di tempo – ma in che modo l’interazione con questi dispositivi valorizza o impoverisce la forma di vita che abbiamo costruito attorno allo sviluppo della personalità individuale nelle diverse sfere, in un contesto sociale educato all’interesse e alla sollecitudine verso le altre persone. L’ideale umanistico si propone di realizzare le condizioni in cui siamo incoraggiati a educarci, coltivando aspetti della mente e della persona, spinti a trovare una nostra posizione personale e unica nel mondo, ad attraversare lo spettro della sensibilità umana e ad apprezzare queste qualità con curiosità e spirito solidale. Lo scopo politico che va posto, nella prospettiva umanistica che sto difendendo, è quello di sviluppare e impiegare le tecnologie cognitive in modo che si realizzi una espansione dell’ideale umanistico.

Le tecnologie cognitive costituiscono un potenziale di espansione dell’ideale umanistico e al contempo sono anche una minaccia. Possiamo usare l’IA come una fonte di mosse che non avremmo compiuto, come una spinta nuova a superare i nostri limiti culturali – una mossa che giocatori di go considerano inappropriata e che il software AlphaGo ha compiuto aprendo uno scenario nuovo per questo gioco¹². Oppure possiamo usarla perché compia operazioni al nostro posto, accrescendo la tendenza alla proceduralizzazione delle attività umane e alla trasformazione della personalità individuale in un sé trasmissivo, come lo chiama Christopher Bollas, un abile trasmettitore di stimoli e direttive. Il contrasto tra i due impieghi molto diversi delle tecnologie cognitive non riguarda peraltro noi in quanto tali, l’umanità, un concetto che torna a essere enfatizzato in modo sospetto, ad esempio quando si parla di

¹² A.I. Could Solve Some of Humanity’s Hardest Problems. It Already Has. Ezra Klein Interviews Demis Hassabis, in «The New York Times», 11 luglio 2023.

rischi esistenziali, perché nasconde le diseguaglianze che proprio la diffusione di queste tecnologie sta ulteriormente favorendo. Le tecnologie cognitive stanno effettivamente trasformando le attività umane e coinvolgono pertanto questioni di natura concettuale (cosa significare pensare, parlare, agire, esprimersi, scegliere) e riguardano perciò cosa vuol dire essere umani. Tuttavia sono anche classiche questioni di giustizia. Le persone fortunate, dotate dell'agio economico, di intelligenza e di curiosità, potranno trarre grandi benefici dalle nuove tecnologie, e lo stanno già facendo, mentre per le persone meno fortunate questo straordinario motore di progresso si sta trasformando in un fattore di uniformazione e asservimento, al servizio della logica di mercato: lavoratori controllati con dispositivi che monitorano l'attività neuronale¹³, consumatori continuamente sollecitati a dare il proprio contributo alla produzione e allo scambio di merci.

3. IA e medicina

Ho presentato alcune considerazioni che riguardano lo sviluppo delle nuove tecnologie cognitive in generale, ma è importante distinguere tra la vita quotidiana e le attività professionali. È sull'uso quotidiano, che per le Big Tech è fonte di enormi guadagni con un mercato che si estende potenzialmente alle intere popolazioni, che si sono concentrate molte critiche e dove possiamo vedere un rischio di perdita del valore della personalità individuale e dello spazio pubblico dove essa può fiorire. Nell'uso professionale ci aspettiamo che l'IA risolva grandi problemi scientifici e tecnologici. Gli scopi in questo caso non sono innanzitutto quelli di cambiare la struttura delle nostre vite quotidiane, ma di consentire progressi che altrimenti avrebbero richiesto molto tempo per essere realizzati. Tuttavia indirettamente anche questi sviluppi dell'IA

¹³ N. A. Farahani, *The Battle for Your Brain: Defending the Right to Think Freely in the Age of Neurotechnology*, St. Martin's Press, New York 2023.

avranno, e in parte stanno già avendo, degli effetti sulla nostra vita quotidiana.

Consideriamo rapidamente in conclusione il campo della medicina. Come è noto, l'uso del machine learning che lavora su grandi quantità di dati sta cambiando la professione medica¹⁴. L'impatto sulla diagnostica per immagini e sulle specializzazioni che fanno grande affidamento sulle immagini (dalla radiologia alla patologia, la gastroenterologia e l'oftalmologia) è davvero importante. Esso riguarda però anche i dati che non sono immagini, come quelli chimici, sequenze genomiche, EEG e tutti i dati che possono essere raccolti tramite chip con identificazione a radiofrequenza, anche solo con degli smartwatch al polso. Tuttavia anche i testi che riguardano gli stili di vita e le abitudini relative alla salute che possono essere raccolti da internet sono oggetto del machine learning. Gradualmente l'IA applicata a questo insieme di grandi dati permetterà una sempre migliore interpretazione delle patologie, con una predizione del rischio molto più accurata e di conseguenza con trattamenti più efficaci. Siamo solo agli inizi di progressi imponenti che ci catapultano in un mondo molto diverso da quello presente.

Esaminiamo ora alcuni profili etici che si pongono in questo ambito specializzato.

(1) Una prima questione è la relazione del medico con queste tecnologie. Come ho già osservato, la caratteristica del machine learning è che esso è tipicamente cieco. Quanto possono essere resi comprensibili i processi che hanno portato a questi risultati? La questione riguarda la fiducia del medico in queste tecnologie e quindi lo statuto del suo sapere, ma riguarda anche la responsabilità innanzitutto etica e quindi deontologica e giuridica per le sue scelte. Se andiamo verso una deresponsabilizzazione del medico (che potrebbe farla ricadere sui dispositivi utilizzati ed

¹⁴ P. Rajpurkar, E. Chen, O. Banerjee, Eric J. Topol, *AI in Health and Medicine*, in «Nature Medicine», 28, gennaio 2022, pp. 31-38.

eventualmente sulle aziende che li hanno prodotti), ciò finirebbe con l'incidere notevolmente sull'autorevolezza della professione. Il cambiamento riguarda però anche i pazienti che potrebbero perdere la capacità di comprensione della propria situazione medica e quindi il controllo sulla propria salute (riducendo il grado di autonomia). In qualche modo questa tendenza sembra far parte dello sviluppo della medicina, tuttavia questo processo di perdita di autorità in prima persona del medico potrebbe essere compatibile con un guadagno in un'altra dimensione. Nel delegare alle tecnologie molte sue attuali specialità, la medicina conserverebbe uno sguardo globale sul paziente e la sua situazione, con una capacità di interpretare le esigenze mediche del paziente dal punto di vista complessivo dei suoi scopi e valori, dentro una relazione personale che le tecnologie non sostituiscono.

(2) Una seconda questione riguarda la sfera personale e il tema della medicalizzazione. Questa tecnologia entrerà nei nostri corpi con chip e un controllo a distanza che consentirà una manutenzione predittiva della nostra salute. Al contempo questo genererà l'aspettativa che le persone siano disponibili a essere controllate in ogni momento. Ciò può finire con il cambiare la nostra percezione della responsabilità perché saremo portati ad aspettarci condotte sempre più prudenti. Non accetteremo più come naturale una certa malattia, frutto della *sorte*, ma potremo dire che è in larga parte responsabilità del paziente. Questo già accade con condotte come il fumo e l'alcol, ad esempio. Ma potrebbe riguardare la nostra vita complessivamente e nei suoi aspetti più intimi, finendo con il medicalizzarci, da una parte, e, dall'altra, invadendo le scelte personali. Certamente fa molta differenza se viviamo in nazioni dotate di un sistema di sanità pubblica o dove vige un sistema di assicurazioni private che possono usare questi dati per negare l'assistenza sanitaria. Tuttavia si tratta in ogni caso di un significativo cambio di paradigma, poiché la malattia sarà sempre più trattata come il frutto di scelte individuali e sempre meno

come una disgrazia che ci è capitata. È un processo di responsabilizzazione individuale che comporta molti rischi perché siamo spinti a diventare responsabili e quindi a diventare soggetti con una propria individualità attenta a ciò che ci circonda e a ciò che facciamo dal punto di vista di interessi medici che selezionano e plasmano aree della vita. Lo scopo dell'ideale umanistico è quello di educare la nostra esperienza in modo da nutrire fiducia in essa al punto da presentare noi stessi come esempi per gli altri, individui con una personalità propria. Questo è possibile se ciò che facciamo non è esaustivamente ricondotto a una rappresentazione medica, o a qualsiasi rappresentazione di cui non controlliamo i criteri e che è imposta dall'esterno, ma rimane a disposizione dell'educazione che le più diverse fonti dell'immaginazione e della cultura possono impartire. La tendenza a ricondurre una sfera sempre più ampia e articolata della vita alla dimensione della prudenza medica costituisce un ostacolo alla formazione di una vita individuale dove il rischio e l'azzardo fanno parte dell'avventura della vita.

(3) Un altro tema enorme è quello della privacy. Le tecnologie cognitive lavorano su grandi dati. Teniamo alla privacy ma non c'è dubbio che c'è un conflitto tra il punto di vista della privacy che attiene al controllo e al consenso relativamente ai propri dati sensibili e il punto di vista del progresso medico realizzato attraverso l'IA. L'Unione Europea ha puntato sulla privacy nella gestione delle nuove tecnologie, tuttavia una difesa molto elevata della privacy priva la ricerca dei dati necessari. Trascurare la privacy significa andare nella direzione delle soluzioni non democratiche come in Cina dove i dati sono utilizzati per scopi politici, o in assetti con poche protezioni pubbliche come negli Stati Uniti dove i dati sono usati per scopi commerciali (incluso

il mercato della salute). D'altro canto però i dati sono il carburante della rivoluzione dell'IA¹⁵.

(4) Un ultimo tema a cui voglio accennare è quello dei pregiudizi, i bias nascosti nei dati o nel modo in cui sono etichettati. I dati possono presentare uno sbilanciamento verso alcuni gruppi della popolazione rispetto ad altri, ad esempio privilegiando la pelle bianca rispetto ad altre pigmentazioni in dermatologia. In questo caso il bias produce informazioni errate per quanto riguarda altri gruppi di popolazione. A mano a mano che l'IA in medicina incorporerà più informazioni, maggiore sarà il potenziale di bias rispetto a gruppi sottorappresentati per vari motivi. Ad esempio, quando un gruppo rilevante dal punto di vista medico si correla con situazioni di ingiustizia sociale tanto maggiore sarà il bias nei suoi confronti perché è sottorappresentato o rappresentato in modo pregiudizievole.

D'altronde il problema di fondo è che tutte le attività umane sono *biased* e quello che chiamiamo progresso morale e civile, come ha sostenuto Mill¹⁶, consiste proprio nella graduale scoperta che quelli che consideravamo modi naturali di organizzare la vita sociale e di trattarci reciprocamente erano in realtà delle gravi ingiustizie, e così è successo con lo schiavismo, con il razzismo, con la discriminazione delle donne e degli orientamenti sessuali minoritari. Il problema che riguarda la medicina e tutte le altre professioni è che questa impostazione pregiudizievole influenza la raccolta dei dati e in questo modo prolifera attraverso il machine learning. La tecnologia può cristallizzare e ampliare il pregiudizio ed essere quindi di ostacolo alla nostra capacità di rivedere opinioni e consuetudini liberandoci dei pregiudizi.

L'IA che si propone di risolvere problemi specifici in ambiti professionali circoscritti appare foriera di grandi progressi

¹⁵ S. L. Sorgner, *We Have Always Been Cyborgs. Digital Data, Gene Technologies, and Ethics of Transhumanism*, Bristol University Press, Bristol 2022, pp. 42-49.

¹⁶ J.S. Mill, *Utilitarismo* (1861), in Id., *La libertà. L'utilitarismo. L'asservimento delle donne*, cit., p. 325.

scientifici e tecnologici e ha senso distinguerla dai dispositivi cognitivi che sembrano esclusivamente il frutto di interessi di mercato, e tuttavia anche lo sviluppo specialistico dell'IA contribuisce a modificare lo spazio comune dello sviluppo umano, come ho voluto suggerire nel caso della medicina. Gli scopi che ispirano le democrazie e gli ideali umanistici interrogano tutte le dimensioni della svolta tecnologica che stiamo vivendo.