L'innovazione tecnologica dell'IA nell'età della tecnica

di Sergio Rondinara

The technological innovation of AI will be able to help the human person in various aspects of his personal and social life. But it will also cause problems already intrinsic to the innovation and that will be amplified by the marked cultural context of the domain of technology. This essay proposes the search for a line of thought to support an ethically responsible strategy for the advent of technological innovations.

Keywords: Artificial Intelligence; Technological Innovation; Technique; Wisdom; Speculative Rationality

Sono innumerevoli i caratteri culturali che contraddistinguono la complessità dell'epoca a noi contemporanea. Risulta pertanto estremamente difficile, se non impossibile, farne un'analisi che la esprima pienamente o la spieghi in modo esaustivo. Non possiamo però non notare l'ambiguità e il disorientamento provocato, da un lato, dallo stato d'animo di sentirci al vertice di un processo di civilizzazione avviato da alcuni millenni, e dall'altro nell'esser immersi in quel particolare processo di strutturazione delle società umane e dell'ambiente naturale raggiunto attraverso lo sviluppo tecnologico, che per la sua incisività non ha riscontro in epoche passate. Ci sentiamo fortemente legati a tutti quei processi culturali che partendo dagli albori delle nostre civiltà ci raggiungono e coinvolgono quotidianamente il nostro pensare tenendo così lo sguardo del nostro spirito rivolto all'indietro, ma allo stesso tempo, ci sentiamo sempre più padroni di noi stessi, della nostra storia e soprattutto del nostro futuro mediante il dominio tecnico che abbiamo imparato ad esercitare sul mondo. Siamo collocati in una spaccatura culturale in cui siamo, per così dire, rivolti all'indietro verso un passato in cui la saggezza (φρόνησις) era un valore ambito e pregnante, e nello stesso momento proiettati verso un futuro intrastorico programmabile e pianificabile attraverso la tecnica (τέχνη). Tutto ciò genera e segna, almeno in parte, quell'ambiguità e quel disorientamento che contraddistinguono la nostra epoca in cui viene a mancare un riferimento chiaro e univoco verso cui rivolgersi.

I concetti *tecnica* e *saggezza* formano oggi una diade che affiora sull'orizzonte filosofico con tutta la sua drammatica portata mostrando la grande rilevanza che ha per il futuro delle civiltà. Si tratta, dunque, del rapporto tra le nostre capacità di produrre mezzi che ci permettono di superare le innate carenze biologiche della specie umana e la nostra attitudine a esser capaci di riflettere profondamente sulle esperienze passate e presenti per poi esprimerne una valutazione esistenziale ed etica.

In tale contesto le stesse analisi delle società contemporanee rischiano di fare ancora ricorso a categorie che, inadatte a dire la rapidità dei cambiamenti, hanno condotto a definizioni tra loro diverse ma spesso con lo stesso prefisso "post": post-industriale, post-fordista, post-analogica, post-moderna.

Il progresso economico-sociale prodotto dall'industrializzazione mediante la realizzazione di innovazioni tecnologiche che con i loro prodotti stanno cambiando il mondo delle relazioni umane[1] ha prima favorito, e poi imposto, un dominio dello spirito tecnico-scientifico in quasi tutti i campi dell'agire umano. C'è da chiedersi se non stiamo facendo slittare il percorso dell'esperienza culturale della nostra

1 - Cfr. L. Floridi, *Infosfera. Etica e filosofia nell'età dell'informazione*, Giappichelli Editore, Torino 2009, pp. 185-194.

specie verso una progressiva razionalizzazione unidimensionale dello spirito umano determinata da uno sviluppo ipertrofico della razionalità strumentale e funzionale della tecnica penalizzando quella razionalità speculativa che presiede e regola le nostre domande sull'esistenza, sul senso e sui fini.

La questione che improrogabilmente si pone, è: fino a che punto vogliamo protrarre tale operazione di razionalizzazione che tende a spegnere in noi la capacità di porci le domande fondamentali sull'esistenza e ricercarne una soluzione autenticamente credibile? Saremo capaci, a partire da un domani prossimo, di una saggezza adeguata a tale sfida?

Dalla rivoluzione industriale ad oggi abbiamo nutrito una meraviglia tanto estatica quanto ingenua nei confronti dei prodotti della tecnica insieme con la fiducia di realizzare per mezzo di essa un mondo più umano. Ma dall'inizio del Novecento i dubbi hanno cominciato a crescere, hanno dapprima trovato voce nell'inquietante pensiero di qualche profeta e oggi sono diventate un coro di critiche soprattutto in ambito filosofico. Tutti segnali di un diffuso allarme culturale che segna il passaggio di una civiltà verso un suo nuovo assetto. Infatti, non stiamo forse varcando una soglia culturale in cui la nostra civiltà sta transitando da una civiltà e cultura industriale ad una civiltà e cultura post-industriale? È dall'avvio della rivoluzione industriale che ci troviamo all'interno di una tale transizione che dura ormai da circa due secoli e della quale non conosciamo ancora a fondo né le caratteristiche peculiari, né le potenzialità.

La nostra tradizione culturale mostra i segni inequivocabili di un lento volgersi verso il tramonto. In genere assistiamo ad un indebolimento della ragione e a un livellamento verso il basso dell'attività culturale. Non distinguiamo più con facilità un centro verso cui rivolgere lo sguardo, si ha come la sensazione di andare alla deriva. Siamo più inclini al rilassamento e alla fuga da quelle attività che postulano impegni seri o gravosi come l'inoltrarci in sentieri esigenti che conducono a modelli culturali carichi di sfide.

Viviamo un'epoca di transizione in cui insieme ai sintomi di una implacabile decadenza coesistono anche sintomi di una sanità reale come quelli di una critica che, con un profondo senso di preoccupazione, assiste culturalmente all'isolamento della persona umana e che tenta di dar voce a quegli spazi, come la sfera artistica o religiosa, che in una società segnata da logiche strumentali e funzionali non sono tanto considerati pericolosi o sospetti, quanto soprattutto inutili.

La nostra si caratterizza per essere un'epoca in cui i caratteri della civiltà che abbiamo alle spalle e della passata cultura s'intersecano e interferiscono con i motivi emergenti di una nuova cultura profondamente segnata dalla tecnica che non ha ancora ben abbozzato e delineato il suo discorso sull'umano.

La realtà della tecnica

Qual è la realtà della tecnica? La tecnica si costituisce e si esprime in tutte quelle attività mediante le quali vengono prodotti strumenti, mezzi e sistemi per migliorare le nostre condizioni di vita e di lavoro. Cioè quelle attività, capacità e mezzi con i quali la persona umana, nell'esplicare la propria esistenza, tenta di mettere i fenomeni naturali sotto il proprio controllo e la natura al suo servizio. La tecnica non è una primizia della nostra epoca, essa esiste da quando esistono le varie culture.

La tecnica è insita nella nostra natura, ci è "connaturale", in quanto il nostro agire, volto a migliorare e a produrre sempre nuove possibilità di sopravvivenza, è sempre stato sin dagli albori dell'umanità un "agire tecnico". La tecnica si caratterizza, pertanto, come uno dei più potenti fattori culturali che nel corso della storia ha permesso agli uomini e donne di ogni epoca di compensare le proprie carenze biologiche, di adattarsi ai vari climi e di colonizzare spazi a latitudini sempre maggiori. Dalla lavorazione della selce più rudimentale alla realizzazione dei più moderni e sofisticati sistemi informatici, la tecnica ci ha costantemente aiutato ad abitare il pianeta in condizioni di sempre maggiore sicurezza e confort, anche se non si può dimenticare che allo stesso tempo ha svolto un ruolo negativo nel contribuire all'allestimento di strumenti di aggressione e di morte.

Non possiamo dimenticare che il termine *tecnica* sta ad indicare non solo il mondo dei mezzi, o apparati, che nel loro insieme compongono il sistema tecnico, ma con esso si vuole soprattutto far qui riferimento a quella forma della razionalità umana che presiede al loro impiego in termini di ottimizzazione, concretezza pratica, funzionalità ed efficienza. Tutto ciò ci svela un aspetto molto importante della tecnica, il fatto che essa sia nata primariamente non come un'espressione creativa dello spirito umano, ma come il modo di trovare una soluzione alle nostre intrinseche carenze biologiche.

Nella lettura che stiamo facendo dell'epoca che viviamo non possiamo non riconoscere quanto la razionalità che presiede e governa la tecnica, la cosiddetta razionalità funzionale, tenda a vivere di vita propria, tenda ad affermare una radicale autonomia – in realtà è una vera e propria indipendenza –, tenda ad essere refrattaria ad ogni controllo etico, per plasmare e poi strutturare secondo i propri criteri e la propria scala di valori la mentalità di tutte quelle persone che entrano in contatto con essa.

Nell'età della tecnica, la razionalità che essa veicola ed afferma tende prima ad indebolire e poi ad abolire in noi quell'orizzonte umanistico all'interno del quale ci poniamo le domande di senso e significato sul mondo di cui siamo parte integrante, sulla nostra esistenza e sulla sfera del trascendente. In una parola la razionalità funzionale tende a ibernare, a devitalizzare la nostra *interiorità*, la nostra anima. In

tale contesto anche se riuscissimo a porci le domande di senso esse resterebbero inevase poiché la razionalità della tecnica non possiede in sé la capacità di dare risposte a tali domande e non perché essa non abbia ancora raggiunto nel suo sviluppo la piena maturità. La tecnica, infatti, è intrinsecamente incapace di condurci allo svelamento della verità, essa non tende a un fine, non dischiude scenari di senso. La tecnica ottimizza, calcola, fa.

La funzionalità e l'efficienza, che caratterizzano la razionalità della tecnica, sono diventate dei veri e propri riferimenti paradigmatici che segnano tutte le sfere dell'agire umano, sono diventate le direttrici principali e le forze operative che governano la vita dei popoli, fino al punto da attribuirsi il ruolo di un vero e proprio valore etico, e si propongono come guida e meta all'azione del singolo e della società.

L'efficienza in se stessa, affermando un intelligente e parco orientamento a un fine, non ha nulla di negativo, ma con il suo sviluppo la tecnica ha trasformato la stessa pratica dell'efficienza da un mezzo a un fine. È diventata un fine. Mediante questa inversione tra mezzi e fini la tecnica ha realizzato la sua metamorfosi. Da mezzo per superare le nostre carenze biologiche è diventata fine delle nostre azioni, da serva è diventata padrona rendendosi indipendente dalla persona umana.

Una tale inversione tra mezzi e fini non è connaturale alla tecnica, storicamente essa si è prodotta in epoca moderna. Anticipata e preparata dalla svolta cartesiana, in cui la ragione è diventata una ragione soggettiva – mentre da sempre si era essenzialmente configurata come una ragione oggettiva –, essa ha preso avvio dalla rivoluzione industriale in avanti. In epoca premoderna, in cui il reale era interpretato secondo una gerarchia di esseri, contavano inevitabilmente soprattutto i fini, gli scopi. Ma ora che nel nostro mondo contemporaneo si guarda soltanto ai mezzi, l'esito di tale processo è che il pensiero speculativo non serve più a sapere se un dato fine sia desiderabile in sé, perché questo non dipende più dalla ragione. La ragione, di fatto, è diventata strumento.

È questa la ben nota tesi sostenuta da Max Horkheimer ne *L'eclisse della ragione*: «La ragione è ormai completamente aggiogata al processo sociale; unico criterio è diventato il suo valore strumentale, la sua funzione di mezzo per dominare gli uomini e la natura»[2].

Con il decadere del discorso teleologico la ragione viene a perdere i suoi contenuti, diventa vuota, puro strumento di calcolo e il suo ruolo diventa esclusivamente formale. Per Horkheimer una tale ragione può essere definita solo come *stupidità* e la sua attività un vero e proprio processo di istupidimento.

La ragione funzionale della tecnica si sta progressivamente sostituendo in maniera egemonica alla ragione speculativa che presiede alla nostra riflessione sui

2 - M. Horkheimer, *L'eclisse della ragione*, Einaudi, Torino 1947, p. 25.

fini, sul senso del reale, sul significato dell'esistere; e una volta che l'avrà scalzata diventando predominante renderà la nostra ragione vuota, senza contenuto e mero strumento di calcolo. Fini e mezzi hanno invertito i ruoli, si sono dunque scambiate le parti: invece di produrre e ricercare i fini, lo scopo dell'uomo contemporaneo sembra paradossalmente essere diventato quello di produrre mezzi. E ciò che rende grave la situazione è che questi stessi mezzi tendono poi a giustificare i suoi fini.

Quale autonomia per la tecnica

Ma vi è un altro carattere preoccupante della tecnica: la sua autonomia. Un'autonomia che, come presto vedremo, si caratterizza essenzialmente come una vera e propria indipendenza. Un attento interprete della modernità contemporanea, Hans Jonas, ha così scritto a riguardo:

> […] gli sviluppi […] del fare tecnologico presentano la tendenza a rendersi autonomi, ossia ad acquisire una propria dinamica coattiva […] del quale non soltanto diventano irreversibili, […] ma acquistano una funzione propulsiva al punto da trascendere la volontà e i piani degli attori. […] ci sottrae di mano la legge dell'agire e […] mentre siamo liberi di fare il primo passo, al secondo e a tutti gli altri successivi siamo già schiavi[3].

Nel suo prorompente dinamismo carico di successi la tecnica con le sue innovazioni tecnologiche reclama una propria autonomia morale e pretende per sé un valore autonomo in quanto ritiene spurio e rifiuta qualunque riferimento assiologico o teleologico estrinseco al suo modo d'essere.

Ogni qual volta costatiamo all'interno della sfera culturale nella quale siamo immersi una progressiva sostituzione del giudizio qualitativo – tipico della nostra capacità di afferrare ed esprimere sia valori che significati e fini caratterizzanti la nostra esistenza – con un giudizio quantitativo tipico di un pensiero calcolante e della razionalità della tecnica, ci troviamo dinanzi a quell'atteggiamento autoreferenziale del sistema tecnico che sancisce la propria indipendenza da ogni sistema assiologico ad esso estrinseco.

Da questa sostituzione di giudizi nasce la disposizione dell'universo della tecnica ad escludere ogni sua dipendenza da un pensiero etico che non sia riconducibile alla propria forma di razionalità.

Inoltre, la tecnica offre all'uomo contemporaneo l'ideale di una umanità prometeicamente onnipotente, ma alquanto unidimensionale a motivo del suo appiattimento sulla sola razionalità funzionale. A causa di ciò i grandi quesiti sulla verità,

3 - H. Jonas, *Il principio responsabilità*, Einaudi, Torino 1979, p. 41.

sul senso dell'esistere, sul bene, perdono di significato, e il concetto di verità viene ad identificarsi con la pura e semplice nozione di *efficacia*. È vero ciò che funziona efficacemente.

Il sistema tecnico[4] esige per sé che gli venga riconosciuta non soltanto la capacità di poter formulare un proprio codice morale, e quindi una sua vera e propria legittimazione morale, ma soprattutto avanza la pretesa di una non negoziabile indipendenza rispetto ad ogni altro sistema etico o insieme valoriale emerso al di fuori del suo stesso ambito.

È proprio l'occultamento dei fini che la tecnica coltiva in sé a fare in modo che essa veicolando una razionalità unidimensionale, come quella funzionale, proponga di fatto una sua tipica irrazionalità.

Da questo quadro culturale sommariamente tratteggiato si può comprendere l'interrogativo allarmato di tanti filosofi del secolo scorso e a noi contemporanei[5]: dove potrebbe approdare l'umanità semmai ci fosse una resa senza condizioni del pensare umano alle pretese avanzate dalla tecnica? Una citazione per tutte:

> Ciò che è veramente inquietante non è che il mondo si trasformi in un completo dominio della tecnica. Di gran lunga più inquietante è che l'uomo non è affatto preparato a questo radicale mutamento del mondo. Di gran lunga più inquietante è che non siamo ancora capaci di raggiungere, attraverso un pensiero meditante, un confronto adeguato con ciò che sta realmente emergendo nella nostra epoca[6].

In questo scenario si colloca la grande sfida per l'uomo contemporaneo immerso nel sistema tecnico: il recupero della propria interiorità. Filosoficamente parlando, il recupero dell'anima mediante una cura conscia, intenzionale e particolareggiata della propria vita interiore. Cura che superando ogni insidia conformistica liberi la persona da quella *forma mentis* che inquadra ogni problema o questione in termini esclusivamente quantitativi, di ottimizzazioni ed efficienze varie. Il pensiero delle molteplici culture, la filosofia, l'arte nelle sue varie forme, la poesia, il sen-

---

4 - Cfr. J. Ellul, *Le système technicien*, Le Cherche-Midi Editeur, Paris 2004.

5 - Cfr. H. Arendt, *Vita activa. La condizione umana*, Bompiani, Milano 2008; J. Elull, *Il sistema tecnico*, Jaca Book, Milano 2009; U. Galimberti, *Psiche e techne. L'uomo nell'età della tecnica*, Feltrinelli Editore, Milano 2000[4]; A. Gehlen, *L'uomo nell'era della tecnica,* Armando Editore, Roma 2003; M. Heidegger, *La questione della tecnica*, in Id. *Saggi e discorsi*, Mursia, Milano 1991, pp. 5-27; M. Horkheimer, *Eclisse della ragione*, Einaudi, Torino 1970; H. Jonas, *Dalla fede antica all'uomo tecnologico*, Società editrice il Mulino, Bologna 1991; H. Jonas, *Il principio di responsabilità. Un'etica per la civiltà tecnologica*, Einaudi, Torino 1993; E. Severino, *Il destino della tecnica*, Rizzoli, Milano 1998; E. Severino, *Téchne. Le radici della violenza*, Rizzoli, Milano 2002; O. Spengler, *L'uomo e la tecnica. Contributo a una filosofia della vita*, Guanda, Parma 1992.

6 - M. Heidegger, *L'abbandono*, Il Melangolo, Genova, 1983, p. 36.

so della storia sono tutti ingredienti di quella formazione umanistica che, compensando il ruolo della tecnica nella vita culturale, contribuiscono sostanzialmente al recupero della nostra interiorità.

Non si tratta pertanto di rifiutare in blocco la tecnica per desiderare nostalgicamente un passato irripetibile, ma si tratta di recuperare quella vita interiore capace di produrre in noi quella saggezza che sappia coordinare l'uso cooperativo della razionalità speculativa e della razionalità funzionale, che sappia esercitare integralmente la nostra capacità contemplativa e la nostra capacità operativa e tecnica. La saggezza, dunque, come forma dell'intellettualità dell'uomo di domani che sappia valorizzare oltre ai valori umanistici trasmessi da una tradizione culturale anche quegli elementi antropologici (serietà del lavoro per il raggiungimento dei fini proposti, disincanto dalle troppo facili utopie sociali, educazione alla provvisorietà e alla rivedibilità) che la sfera della tecnica è capace di offrire alla vita interiore della persona umana.

Coltivando saggiamente la nostra interiorità potremmo tonificare, con l'esercizio, la coscienza di essere un centro autonomo e irripetibile di pensiero, volontà e libertà. In questo modo potremmo coltivare ciò che è più tipico della specie umana: ci scopriremo non solo capaci di affinare le nostre valutazioni ma di avere allo stesso tempo potenziato qualitativamente la nostra intrinseca predisposizione all'esercizio dell'alterità. La cura della vita interiore che nasce dalla saggezza e che a sua volta incrementa, Socrate ce lo ha testimoniato, è di per sé diffusiva e contagiosa. La cura di un'autentica interiorità inevitabilmente crea un fascino spirituale attraente, tende a produrre intorno a sé più consapevolezza, più attenzione e libertà.

È in quest'orizzonte di dominio della razionalità della tecnica che noi assistiamo all'apparire di innovazioni tecnologiche più o meno annunciate. L'Intelligenza Artificiale (IA) è quella che oggi domina la curiosità e l'informazione. Cerchiamo ora di capirne il programma senza entrare troppo nel dettaglio.

Che cos'è l'IA

Gli ultimi anni hanno visto una crescita costante dei programmi di ricerca nel campo dell'IA, diventata ormai un fenomeno di massa che suscita un interesse senza precedenti in ogni dove. Recentemente tutto ciò è stato ulteriormente rafforzato da prodotti informatici di enorme successo quali vari tipi di ChatGPT. Questi prodotti sono la concretizzazione visibile di decenni di ricerca nel campo delle reti neurali e della IA, con una accelerazione importante che dal 2012 le ha portate a diventare uno strumento di rilievo in tantissimi campi, dagli assistenti vocali all'analisi di proteine.

I programmi di ricerca riguardanti l'IA, nella loro essenza, sono da sempre un settore di studio e ricerca dell'informatica e questo sin dai suoi inizi con i lavori di Alan Turing, universalmente riconosciuto come il padre di tale scienza.

Il compito che si assume l'informatica è quello di modellizzare e risolvere problemi, dove la prima abilità è importante almeno quanto la seconda. Per raggiungere questi obiettivi essa usa strumenti chiamati algoritmi, attraverso i quali fornisce una descrizione finita e non ambigua di una sequenza di operazioni che consentono ad un certo agente di risolvere un determinato problema.

Nei programmi di ricerca riguardanti l'IA sono da sempre presenti due filoni d'indagine principali: l'IA guidata dai modelli (IA simbolica) e IA guidata dai dati (IA sub-simbolica).

L'IA simbolica fa largo uso della logica matematica per la rappresentazione della conoscenza in generale e del ragionamento automatico, sia esso di tipo deduttivo che abduttivo, in particolare. Tale programma viene usato in ambito teoretico mediante la dimostrazione automatica dei teoremi (*theorem proving*) e la verifica automatica delle loro dimostrazioni (*proof checking*).

Le applicazioni dell'IA simbolica possono essere così riassunte:

– la *pianificazione automatica*. È questa un'operazione dove una volta assunti i dati della situazione iniziale, un obiettivo da raggiungere e un insieme di vincoli, si richiede alla macchina di costruire un *piano d'azione* (se esiste) che consenta di raggiungere l'obiettivo a partire dalla situazione iniziale e rispettando i vincoli dati. Si vuole produrre una macchina che sia capace di eseguire questo piano d'azione, di monitorarne l'esecuzione e in caso si rilevino situazioni impreviste o indesiderate riformulare un nuovo piano. Normalmente la si usa per l'automazione dei processi di produzione, ma anche per le funzioni che svolgono i diversi satelliti artificiali o sonde spaziali.

– la *robotica collaborativa*. La robotica contemporanea si sta sviluppando nella direzione della robotica collaborativa dove si punta alla collaborazione tra umani e robot per portare a termine un dato obiettivo in ambito industriale o domiciliare. In questo senso la robotica collaborativa combina le potenzialità e le capacità delle persone con quelle dei robot per portare a termine, in maniera semi-automatica, dei compiti che risultano impossibili da realizzare mediante una operazione totalmente automatica. I robot collaborativi (cobot) sono dei robot tipicamente industriali di nuova generazione progettati per lavorare assieme all'essere umano, fianco a fianco e in sicurezza, senza la necessità di barriere o gabbie protettive che separino gli uni dagli altri, ma vi sono applicazioni anche nell'ambito dell'assistenza domiciliare e della domotica.

– il *Business Process Management* (BPM). Sono strumenti informatici che nella gestione dei processi aziendali permettono di migliorare molto la prestazione del

sistema. Si mette a disposizione delle aziende una molteplicità di strumenti per la gestione dei loro processi. Questi sono strumenti capaci di esplicare varie funzioni: a) consentono di scoprire, rappresentare ed analizzare i compiti sottostanti ai processi aziendali (*process mining*); b) consentono di visualizzare i processi aziendali (*diagrams*); c) strumenti per l'automazione dei processi o della gestione dei processi aziendali (*workflow engine*); d) infine strumenti per la simulazione e il testing di processi aziendali.

Più brevemente l'IA sub-simbolica lavora per raggiungere l'obiettivo di imitare, o meglio, mimare il comportamento del cervello umano e la sua complessa rete di neuroni interconnessi. I primi modelli di reti neurali risalgono già agli anni '40 del secolo scorso con gli studi primordiali di neuroscienze computazionali mediante la modellizzazione dei neuroni da parte del neurofisiologo Warren S. McCulloch e del matematico Walter Pitts[7].

I due filoni non sono tra loro separati nel senso che i problemi dati possono essere risolti con uno o l'altro filone. Il caso classico è l'elaborazione del linguaggio naturale dove i primi lavori usarono l'approccio simbolico dove la linguistica computazionale ebbe un grande sviluppo e produsse degli ottimi lavori ma non sempre conclusivi. Successivamente si è seguito l'approccio sub-simbolico fornendo al sistema grandi collezioni di testi scritti e orali prodotti in contesti di comunicazione reale (ad esempio testi della rete) e utilizzando i prodotti per l'apprendimento automatico (*machine learning*) i cui risultati hanno espresso le varie Chat-GPT.

I due programmi di ricerca *IA simbolica* e *IA sub-simbolica* non sono tra loro separati, pur avendo origini diverse sembrano destinati ad interagire tra di loro.

Da sempre abbiamo cercato di far fare alle macchine delle cose che non eravamo in grado di fare, o che non volevamo fare; ciò non ci ha mai spaventato. Oggi la novità è che i sistemi di IA sono in grado di sostituire l'essere umano e, spesso, di migliorarne le prestazioni in compiti ritenuti da sempre di sua competenza esclusiva. Compiti che hanno a che fare con il ragionamento e la memoria.

Autonomia delle macchine, problemi di affidabilità e rischi

Una questione che non possiamo tralasciare è quella relativa all'autonomia delle macchie. Questo è un argomento fondamentale in quanto l'essere umano vorrebbe essere sempre in grado di potersi inserire nel ciclo di comportamento di una macchina e intervenire dicendo «Si» o «No» all'avanzamento del programma. Nei sistemi robotici, e più in generale nei sistemi di IA, non sempre è possibile garan-

7 - W. McCulloch - W. Pitts, *A Logical Calculus of Ideas Immanent in Nervous Activity*, in «Bulletin of Mathematical Biophysics» 5 (1943), pp. 115-133.

tire la condizione che preveda l'interazione col soggetto umano nei processi di decisione (*human in the loop*). Questa impossibilità è dovuta alla velocità operativa della macchina che è molto superiore a quella dell'essere umano nel prendere una decisione.

In questo senso la macchina è autonoma, ma non nel senso che ci sia un'assenza del controllo umano. È bene non confondere l'incapacità umana di prevedere e controllare il comportamento di una macchina con un'attività "creativa" della macchina stessa.

Più una macchina è complessa, e quindi potente, più risulta difficile controllarne totalmente il comportamento. Se vogliamo lasciare la macchina completamente autonoma dobbiamo esser sicuri che anche quei comportamenti che non siamo in grado di poter controllare non siano pericolosi. Il livello di autonomia che siamo disposti ad accettare deve essere commisurato al grado di criticità di certi comportamenti. Si pensi al caso delle auto a guida autonoma, ma anche all'uso dell'IA in ambito militare.

Riguardo il problema dell'affidabilità si pone la questione: «possiamo credere a quello che fa o che dice il mio sistema di IA?». In molti casi, quali le applicazioni finanziarie e in diversi sistemi critici dal punto di vista della sicurezza quali velivoli, centrali nucleari in cui l'intervento deve essere rapido, la condizione *human in the loop* non può essere garantita e quindi soddisfatta. Allora occorre costruire dei sistemi di cui ci si può fidare (*trustworthy AI*). L'IA di cui ci si può fidare è quella macchina che è in grado di spiegare e rendere ragione delle sue conclusioni o previsioni. L'obiettivo del *trustworthy AI*, a oggi, non è purtroppo ancora stato raggiunto. Una possibile strada risolutiva potrebbe essere quella risultante dall'integrazione dei lavori tra l'IA simbolica e l'IA sub-simbolica.

A questo punto non possiamo non interrogarci su quale sarà il carattere qualitativo delle relazioni tra uomo e macchina dopo il pieno avvento dell'innovazione tecnologica dell'IA. Relazioni che potrebbero possedere delle finalità utili ma anche far sorgere manipolazioni del comportamento delle persone.

Già oggi abbiamo in rete decine si siti[8], dove si può programmare la propria ragazza virtuale; un avatar con la quale milioni di utenti intrattengono conversazioni e che a sua volta è capace di relazionarsi con il proprio ideatore a livello quasi umano. Sarebbe un errore ritenere che tali relazioni possano essere considerate dei fenomeni di nicchia quando in questo mercato miliardario le cinque app più popolari sono state scaricate oltre 27 milioni di volte[9]. Questa non è altro che una sin-

8 - Cfr. *Virtual Girlfriend, Ai Girlfriend o iGirl*.

9 - Cfr. il sito *What's the big data*, https://whatsthebigdata.com/ai-girlfriend-statistics/ (accesso: 23 gennaio 2025).

gola quanto allarmante manifestazione di una nuova tipologia di relazione tra uomo e macchina che le persone potranno avere con l'uso di modelli avanzati di IA.

È dunque più che mai lecito chiedersi quale tipo di relazione uomo/macchina è possibile stabilire con modelli di IA, se non altro perché nel mondo vi sono ormai milioni di persone che dicono di sperimentare una relazione empatica con queste macchine, ignorando o omettendo che le risposte ottenute sono frutto di algoritmi. La domanda possiede una ulteriore legittimità poiché questo tipo di relazione diventerà sempre più comune in futuro, con modalità che certamente renderanno sempre più difficile capire se il proprio interlocutore, la persona con la quale si sta intrattenendo una relazione, sia realmente un essere umano, o invece un computer.

Nel suo relazionarsi con le IA, il soggetto umano potrebbe mettere in atto dei comportamenti psicologici particolari che ne corroborano il narcisismo, o che potrebbero generare esperienze solipsistiche, isolandolo in un mondo virtuale nel quale si rispecchia e si riconosce completamente, riducendo progressivamente i legami "reali" con i suoi simili. Nel già noto fenomeno dei «reclusi nella rete» (*hikikomori*) assistiamo al fatto che le relazioni virtuali hanno sostituito progressivamente quelle reali. Se in questo caso le relazioni – seppur virtuali – sono con altri esseri umani, nulla esclude che nel prossimo futuro le relazioni degli *hikikomori* divengano anche con macchine a IA.

Le relazioni con macchine a IA, sotto certi aspetti, possono essere più soddisfacenti rispetto a quelle che si possono avere con i propri interlocutori umani, poiché tali macchine possono empatizzare con il proprio interlocutore, in quanto lo capiscono meglio di altri, ne sanno anticipare i bisogni e ne rafforzano le proprie convinzioni. Uno studio preventivo su come muterà il rapporto tra persona e macchina, grazie all'uso dell'IA, si rivela quindi essenziale sia per comprendere la società di domani sia per sciogliere i molti interrogativi che in noi sorgeranno.

Conclusioni

Ciò che può rendere problematica una innovazione tecnologica non sono solo quei rischi che essa intrinsecamente veicola, ma gioca un ruolo importante anche il contesto culturale e sociale in cui tale innovazione appare e si afferma. Contesto che può essere preparato ad assorbire l'impatto e poi essere capace di una buona ricezione dell'innovazione, oppure ad amplificare quei caratteri negativi che essa ha già in gestazione.

La ricezione sociale delle innovazioni tecnologiche mostra un carattere ambivalente in quanto vi sfociano componenti di natura diversa come quella psicologica e quella pragmatica. Il carattere pratico-pragmatico determina quella serie di questioni che riguardano lo scopo (non certamente disinteressato) del produrre l'innovazione, la varietà di mezzi necessari alla sua creazione (materiali, *know how*) e il

contesto (non sempre ben conosciuto) con le relative conseguenze sociali. Questa tematica schiude tutta una gamma di concetti e giudizi di valore che riguardano la sfera etica. Sfera già in serie difficoltà ai giorni nostri, ma che entra ancor più in difficoltà quando tenta di esprimere i propri giudizi sulle innovazioni tecnologiche proposte da gruppi economici e industriali, dai complessi militari-industriale e dai sistemi politici, tutti soggetti che nel perseguire i propri obiettivi mostrano quotidianamente una certa impermeabilità ai valori etici.

Come ogni innovazione tecnologica anche l'IA potrà aiutare la persona umana in vari aspetti del suo vivere personale e sociale, ma inevitabilmente procurerà anche quei problemi già intrinseci all'innovazione e che verranno amplificati dal contesto culturale segnato dall'era del dominio della tecnica dove non sarà facile reperire quella saggezza e responsabilità necessarie ad un autentico sviluppo dell'umano. Il timore che, tale innovazione tecnologica procuri un ulteriore sviluppo della razionalità funzionale della tecnica, a scapito della razionalità speculativa umana, con il conseguente sfaldamento della capacità delle persone di affrontare questioni sul senso dell'esistenza e sui fini che caratterizzano le nostre azioni, è più che fondato. Nel momento in cui dovremmo promuovere una riflessione etica e una maggiore equità sociale per la rinascita dell'umano ci troveremmo ad affrontare non una qualunque innovazione tecnologica o un gigantesco balzo verso una quarta rivoluzione[10] industriale ma un vero e proprio cambiamento d'epoca senza una saggezza adeguata. Questo perché l'IA segna una nuova e significativa fase nel rapporto dell'umanità con la tecnologia dove i confini tra l'umano e la macchina, tra il naturale e l'artificiale, tra mondo fisico e mondo virtuale diventeranno sempre più opachi e confusi.

Risulterà allora ancora più difficile recuperare quella vita interiore, capace di produrre in noi quella saggezza che sappia unire l'uso cooperativo della razionalità speculativa e della razionalità funzionale, che possa esercitare in modo integrativo le nostre capacità contemplative con quelle tecniche.

In tale contesto – nelle società industriali – occorre mettere a fuoco il problema della ricezione sociale delle innovazioni prodotte dalla tecnica e contemporaneamente una linea di pensiero a sostegno di una strategia eticamente responsabile per le sue applicazioni.

C'è da chiedersi, vista la globalità delle popolazioni che verranno investite dalle innovazioni tecnologie riguardanti l'informatica, se non sia il caso di dare voce alla comunità internazionale che a partire dalla presa di coscienza di responsabilità degli stati sovrani, produca un documento sotto forma di trattato internazionale

10 - L. Floridi, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo,* Raffaello Cortina, Milano 2017.

che possa essere – almeno in alcuni aspetti – giuridicamente vincolante in vista di una appropriata regolamentazione etica delle applicazioni dell'IA nei moltissimi campi della vita sociale[11].

Le attuali organizzazioni internazionali potrebbero coinvolgere il settore privato dell'alta tecnologia nello svolgere un ruolo decisivo per raggiungere accordi multilaterali e avviare iniziative individuali o di gruppi di stati capaci di incoraggiare e sostenere migliori pratiche riguardanti l'IA e stimolare nuovi approcci creativi al problema.

SERGIO RONDINARA
Già Professore ordinario di Epistemologia presso l'Istituto Universitario Sophia (FI)
sergio.rondinara@sophiauniversity.org

11 - L'*AI Act*, il regolamento europeo sull'IA varato dall'Unione Europea sembra un primo piccolo passo che va in questa direzione.