L. Cuocolo, G. Girgenti, A. Granelli
L. Lupária Donati, S. Regazzoni

INTELLIGENZA ARTIFICIALE RISORSA O MINACCIA?

il melangolo *per la* fondazione CARIGE

Andrea Granelli

L'INTELLIGENZA ARTIFICIALE: UN *PHARMAKON* FATTO DI LUCI E OMBRE

> Il pericolo del passato era che gli uomini diventassero schiavi. Il pericolo del futuro è che gli uomini possano diventare robot.
>
> Erich Fromm, *Psicanalisi della società contemporanea*

1. *La tecnologia è un* pharmakon

La cultura greca possedeva nel suo corredo linguistico alcune parole potenti, polisemiche, che definivano con maggiore profondità concetti complessi. Una di queste è *pharmakon*, che in greco significa sia farmaco che veleno. Ed è proprio questa sua duplicità, questa sua apparente contraddizione, che ci illumina sul significato profondo di alcuni concetti. Uno di questi è proprio Intelligenza Artificiale (IA).

Il suo essere *pharmakon* non solo ci ricorda che quanto più la tecnologia è potente tanto più è potenzialmente problematica; ma anche, che la "colpa" non è della tecnologia ma di chi la somministra. Se ben usata, infatti, porta grandissimi benefici. Se è mal somministrata, invece, può causare molti danni e – continuando con l'allegoria medica – arrivare perfino alla morte stessa del paziente.

Vedere l'IA come *pharmakon* ci porta l'attenzione dalla tecnologia alle competenze necessarie per usarla. Una parte rilevante dei danni di queste tecnologie (non consideriamo qui quelli deliberati che richiedono una diversa trattazione) dipende quindi dalle competenze inadeguate sia del somministratore (chi decide) che dei fruitori (chi le utilizza). Il problema principale, dunque, non sta nella tecnologia ma nelle competenze necessarie per utilizzarla in modo corretto.

Il filosofo Paul Virilio, ne *L'incidente del futuro* (2002), fa un passo in più e osserva che "la tecnologia crea innovazione ma – contemporaneamente – anche rischi e catastrofi: inventando la barca, l'uomo ha inventato il naufragio, e scoprendo il fuoco ha assunto il rischio di provocare incendi mortali".

Come dire: è la natura stessa dell'innovazione tecnologica che crea *pharmakon*. Per questo motivo è importante studiare anche le dimensioni problematiche, i lati oscuri del digitale. Ignorando questa dimensione, sarebbe come voler diventare medico studiando anatomia e fisiologia farmacologia… ma non patologia.

Oltretutto, come sanno bene i medici, studiando i malfunzionamenti di un organo si conosce molto di più sull'organo stesso e la sua fisiologia.

Oramai la dimensione problematica non è un fatto ancillare ma è quasi fondativo; Può cioè spostare l'ago della bilancia nel decidere se una certa tecnologia vada utilizzata o meno in maniera massiva. Pensiamo per esempio alla natura energivora dell'IA e al conseguente crescente consumo energetico. Si stima che un grande centro di calcolo (per esempio quelli di Google) consumi quanto prodotto da una piccola centrale nu-

cleare. Inoltre, una parte dell'energia necessaria serve a raffreddare le macchine e richiede quindi anche molta acqua. Già questo fatto basterebbe per guardare con preoccupazione gli sviluppi senza freno dell'IA.

Ma la voracità di questa tecnologia si sta spostando su un'altra risorsa naturale preziosa: le cosiddette terre rare. Questi metalli – per la verità non così rari ma con processi estrattivi costosi e dannosi per l'ambiente – sono fondamentali per la produzione dei chip. Tra l'altro uno dei giacimenti più importanti al mondo è in Ucraina (gli esperti stimano che il valore di questo "tesoro" ammonti a 10/12 trilioni di dollari). E la loro "rarità" è proprio esplosa con la fame di calcolo dell'IA, che sta facendo deflagrare i conti del maggior produttore dei suoi chip: fra aprile e giugno del 2024 Nvidia ha infatti registrato ricavi in crescita del 122% rispetto all'anno precedente e un utile di quasi 17 miliardi di dollari, in crescita del 168%.

Il tema delle terre rare, come pure quello dei vincoli di privacy sui contenuti usati per allenare i sistemi esperti (in questo caso è la Cina lo stato più permissivo) stanno mettendo in luce anche le implicazioni geopolitiche legate allo sviluppo dell'IA.

Ma anche lo smaltimento dei "rifiuti digitali" è molto problematico. Potremmo dire che l'ecosistema digitale è quanto di più lontano si possa concepire dall'economia circolare. Fino a poco tempo fa alcuni produttori di *device* avevano addirittura deciso di non renderli riparabili (non solo dall'utente ma anche dall'azienda stessa) per forzare gli utenti a un ricambio frequente dell'apparato, motivato dall'elevato tasso di innovazione.

Una delle destinazioni mondiali degli scarti dell'e-

lettronica è la discarica di Agbobloshie, un quartiere di Accra, capitale del Ghana. Il posto è apocalittico: Una discarica a cielo aperto, che si perde nell'orizzonte ed è coperta dal fumo prodotto da coloro che cercano di estrarre rame bruciando i cavi; e la soglia di inquinamento e sopra ogni elemento di tollerabilità. Ciononostante, la gente vive ai suoi margini, cercando di guadagnare qualche spicciolo da ciò che i ricchi abbandonano.

2. *La potenza combinatoria dell'IA Generativa*

Sulle potenzialità dell'IA si è scritto molto. Può essere utile, allora, sottolineare l'aspetto che rende potenti e stupefacenti le piattaforme IA generative: si tratta della loro capacità combinatoria e cioè riuscire a generare novità combinando in modo creativo cose conosciute.

Il tema è noto fin dal medioevo; basti pensare alla *Ars magna*, il metodo inventato dal teologo e filosofo catalano Raimondo Lullo tramite il quale si potevano collegare, in una sorta di logica meccanica, concetti fondamentali in modo di acquisire verità in ogni campo del sapere.

Italo Calvino, nelle sue *Lezioni Americane* (Molteplicità), osserva: "Il miracolo di una poetica, apparentemente artificiosa e meccanica, che tuttavia può dare come risultato una libertà e una ricchezza inventiva inesauribile".

E Carlo Rovelli – in un articolo sul "Corriere della Sera" (*L'eremita che immaginò il computer*, 9 ottobre 2016) – spiega da cosa nasce questo nostro stupore che si origina dalla potenza disarmante della combinatoria:

“La nostra intuizione arretra di fronte agli immensi numeri e alla sterminata varietà generati dalle combinazioni. Ci sembra impossibile che combinando cose semplici possano nascere tante cose e tanto complesse”.

Il digitale non è una tecnologia “superficiale” ma trasforma anche il nostro modo di essere, di pensare, di accedere usare i nostri ricordi; ha quindi un impatto molto rilevante sulla nostra vita. Oltretutto è anche difficile parlare di digitale e valutarlo in modo obbiettivo perché – soprattutto con l’Intelligenza Artificiale – viene coinvolto direttamente anche il nostro linguaggio e la produzione e scambio della conoscenza che ci definisce e caratterizza.

Un interessante analisi fatta nel lontano 1982 in *Critical Path* da Buckminster Fuller – noto fra le altre cose per aver inventato la cupola geodetica – mostrava che, fino al 1900, la conoscenza umana codificata (a quei tempi prevalentemente scritta o disegnata) raddoppiava ogni 500 anni. Poi è iniziata un’accelerazione quasi esponenziale, che ha portato nel 2017 il ritmo del raddoppio a 13 mesi. IBM ha continuato la ricerca di Buckminster Fuller (*The Toxic Terabyte*, 2006) arrivando a stimare – ai nostri giorni – il raddoppio della conoscenza ogni 12 ore.

I numeri sono probabilmente iperbolici, ma il tema è sotto gli occhi di tutti noi, basti pensare quanti materiale – testi, immagini, video, foto, messaggi… – si aggiungono ogni minuto alla Rete. In alcune discipline, i ricercatori non riescono neanche a leggere tutti gli articoli scientifici prodotti nel loro ambito di ricerca. Oramai si parla di alluvione informativa o meglio di *information overload*.

Ritornando ancora per un momento ai numerosi lati oscuri che dobbiamo – per amor di senso critico – tenere sottotraccia, uno dei più problematici (e antichi) è legato al timore ancestrale che la tecnologia esca dal nostro controllo. È il mito – iniziato probabilmente con il Golem – che la creatura sfugga di mano al creatore. Frankenstein, l'apprendista stregone e Hal 9000 sono istanze di questo mito che è probabilmente antico quanto l'essere umano.

Alcuni centri di ricerca stanno trasformando questa paura ancestrale in progetti di ricerca. Ad esempio, il MIT di Boston ha lanciato un progetto battezzato *Norman* – nome che fa riferimento al celebre e inquietante protagonista dell'hitchcockiano *Psycho*. L'obiettivo della ricerca è infatti comprendere i meccanismi che consentono di riprodurre il funzionamento di una mente psicopatica – uno dei lati oscuri dell'intelligenza umana. D'altra parte, per rimarcare questa intimità dell'IA con la dimensione psicologica, la parola allucinazione è già oggi usata normalmente per indicare gli output prodotti da una piattaforma generativa che non sono basati sulla realtà o su verità obiettive; ad esempio, una mano con sei dita.

Oltretutto l'attenzione di studio per la mente psicopatica nasce dal fatto che chi la possiede è inquietante non solo per la sua pericolosità ma anche per la capacità di entrare in sintonia con la vittima e usare il linguaggio per ingannare e creare un legame empatico.

3. *Il sofista digitale celato dietro la potenza del linguaggio*

Interessante è richiamare allora quanto Platone, nel *Sofista*, fa dire allo Straniero di Elea (l'attuale Velia, in provincia di Salerno):

> Noi dobbiamo pensare che ci sia anche un'altra arte [*capace di imitare la realtà, come la pittura*], un'arte dei discorsi, per cui sia pure possibile, con i discorsi appunto questa volta, incantare le orecchie dei giovani e di quelli ancora molto lontani dal conoscere la realtà delle cose, mostrando loro immagini parlate [*eidola legomena*] di tutto e in modo da ingenerare in loro l'opinione che chi parla dice la verità e che su tutto è il più sapiente di tutti?

Sembrerebbe una riflessione sull'IA – o meglio l'estratto di una riunione interna di una delle aziende che stanno lanciando sul mercato le nuove piattaforme dell'IA generativa. È invece un'intuizione che risale a circa 2.300 anni fa.

Perché la finalità ultima delle piattaforme d'IA generativa non è tanto la produzione di conoscenza, di sistemi esperti, quanto la realizzazione di un linguaggio potentissimo – detto *language model* – capace di adattarsi ad ogni interlocutore che le interroga e di persuaderlo della correttezza delle loro risposte. Una sorta di "certificato oracolare" che assicura a queste tecnologie di poter persuadere senza dover dimostrare né la coerenza di ciò che dicono né l'attendibilità delle fonti a cui attingono – che infatti rimangono sempre nascoste.

Il cuore del sistema è dunque il *language model* e l'obiettivo dell'allenamento è dare in pasto alla piattaforma miliardi di frammenti di conversazione online

per allinearsi e diventare familiare con i linguaggi della Rete e con ciò che piace (e genera "like") e, in ultima istanza, rafforzare il proprio armamentario linguistico in senso persuasivo per arrivare a costruire ciò che Cicerone chiamerebbe *captatio benevolentiae digitale*.

Che competenze dobbiamo allora sviluppare per usare al meglio questi sistemi senza esserne (ab)usati? Non solo e non tanto le competenze tecniche – che sono condizione necessaria ma non sufficiente – quanto il pensiero critico e le abilità retoriche che ci consentono di smascherare le manipolazioni linguistiche, evidenziare i buchi logici o i ragionamenti fallaci, investigare la qualità delle fonti e i livelli di attendibilità delle risposte.

L'Intelligenza Artificiale, quindi, rimette al centro la cultura umanistica nel senso più autentico, e cioè i saperi attorno all'essere umano, ciò che dobbiamo sapere per conoscerci. E uno degli aspetti critici di questa conoscenza sarà proprio il nesso tra linguaggio e memoria.

Può essere allora utile ai nostri ragionamenti una riflessione sulle memorie artificiali di Boncompagno da Signa, proveniente dal suo *Rhetorica novissima* del 1235 (ancora una volta un testo antico, per ricordarci che la comprensione dell'innovazione richiede anche una conoscenza del passato):

> Che cos'è la memoria. Memoria è un glorioso e ammirevole dono di natura, per mezzo del quale rievochiamo le cose passate, abbracciamo le presenti e contempliamo le future, grazie alla loro somiglianza con le passate. Che cos'è memoria naturale. Memoria naturale viene solamente da dono di natura, senza aiuto di alcuno artificio. Che

> cos'è memoria artificiale. Memoria artificiale è ausiliatrice e assistente di memoria naturale [...] ed è chiamata "artificiale" da "arte", perché è stata scoperta "artificialmente", grazie a finezza d'ingegno.

Pertanto, un utilizzo adeguato dell'IA richiede non solo competenze ma anche abilità di ragionamento, pensiero critico e forza una maggiore comprensione di ciò che sappiamo, che è indissolubilmente legata a ciò che ricordiamo.

Ce lo ricorda Dante nella *Divina Commedia*: comprendere o imparare qualcosa non forma la vera conoscenza se non si tiene poi a mente ciò che si è imparato ("Non fa scienza senza lo ritenere, avere inteso"). Immanuel Kant, nella sua *Antropologia dal punto di vista pragmatico*, estende la riflessione richiamando la centralità della memoria: "Sappiamo solo ciò che conserviamo nella memoria". E perfeziona il tutto in modo poetico ma perentorio Giuseppe Ungaretti durante una conferenza su Giambattista Vico: "Tutto, tutto, tutto, è memoria".

4. *Tre ulteriori punti di attenzione*

Il digitale rende queste paure sulla incapacità di controllare la conoscenza e di perdere la memoria non solo attuali ma anche concrete e realistiche. E comunque la potenza di queste tecnologie aumenta il rischio che cadano in mani sbagliate. La scena finale de Il dottor *Stranamore*, in cui il Maggiore T.J. "King" Kong sgancia l'atomica cavalcandola come un cowboy nel rodeo, è sempre nei nostri occhi.

Tre ulteriori aspetti sono subdoli e potenzialmente perniciosi e vanno pertanto presi in considerazione con particolare attenzione:

1. il depotenziamento cognitivo;
2. gli utenti "*wrongdoer*";
3. i temi (algor)etici.

Partiamo dal depotenziamento cognitivo. Il rischio in agguato, soprattutto per i giovani, è ciò che potremmo chiamare la sindrome di Aladino. Richiamiamo la storia: Aladino è un ragazzo fannullone che vive in una città della Cina. Suo padre Mustafà è morto di crepacuore per la preoccupazione che gli dava la sua indolenza, motivo anche della disperazione della madre. Dopo varie avventure e vicissitudini trova una vecchia lampada dalla quale, se lucidata, esce uno spirito *jinn*, che si offre di servire i portatori della lampada ed esaudire tutti i loro desideri.

Dopo molte avventure e disavventure, tutti i suoi nemici vengono uccisi e lui e la sua sposa vivranno felici e contenti, senza che mai Aladino si sia posto il problema di rafforzare le proprie competenze e abilità … tanto ad ogni bisogno c'era la lampada.

Ed è proprio questo che incominciano a fare i giovani: a ogni bisogno o richiesta, necessità di trovare materiale, di fare delle sintesi o riassunti, di produrre proposte, la ricetta è sempre la stessa: lo chiedo all'IA Generativa.

Ma le facoltà intellettuali – il saper ricordare, sintetizzare, collegare elementi diversi, cercare nuove prospettive, criticare … – sono come un muscolo: se non si esercitano, si atrofizzano. Ma c'è quasi un pregiudizio ad allenare la mente: si tende subito a pensare ai mentalisti, ai manipolatori, ai *mental coach*…

Mentre a nessuno verrebbe in mente di dire a uno che fa jogging "perché vai a piedi quando c'è la macchina?", nel caso dell'allenamento mentale, il meccanismo è proprio questo: "perché devo sforzarmi a riassumere questo testo quando ChatGPT lo fa in un attimo?".

Ma i danni di questo tipo di comportamento sono già noti agli psicologi da molto tempo: uno dei più frequenti si chiama Google Effect: chiamato anche amnesia digitale, è la tendenza a dimenticare informazioni che si possono trovare facilmente online tramite i motori di ricerca, invece di ricordarle ("perché devo fare la fatica di memorizzarle quando le trovo subito sulla Rete appena mi servono").

Ma la memoria funziona non accedendo a dati che diventano deperibili con il passare del tempo, ma piuttosto ricostruendo ogni volta un ricordo sensato tramite l'assemblaggio di frammenti mnestici. È il processo, quindi, che va allenato: non basta la tutela del dato. Osserva Paul Ricoeur nel suo *Passato, memoria, oblio* (2003): "La memoria in quanto facoltà attiva non ritrova le cose 'come erano', ma le ricostruisce, le modifica anche […] e l'aspetto ricostruttivo della memoria assume caratteri paradossali dal momento che i nostri pretesi ricordi sono in realtà spesso presi a prestito da racconti ricavati da altri".

Il secondo punto è l'"esistenza dei malfattori" come utenti di innovazione; aspetto di cui spesso siamo inconsapevoli e di cui ci sfugge sia l'entità che l'impatto. Il tema fu posto dalla rivista *The Economist* in un articolo del 10 giugno 2017 (*Tech Firms could do more to help stop the jihadists*) – mia la traduzione che segue: "Internet, con miliardi di utenti e una potenza di ela-

borazione illimitata, è la rete più potente di tutte. Era inevitabile che diventasse il bersaglio dei malfattori".

In passato, le aziende che operano su Internet hanno avuto la tendenza a "costruire prima, e a capire le regole poi". Tuttavia, le argomentazioni relative al terrorismo e ai contenuti estremisti ci ricordano che l'era senza legge e a ruota libera dei primi anni di Internet è finita. Le aziende tecnologiche potrebbero trovare difficile accettarlo. Ma devono accettarlo, come parte della responsabilità che deriva dal loro nuovo potere e come parte del prezzo del loro successo"

La scelta dell'espressione malfattore – *wrongdoer* – non è casuale ed è molto più incisiva della parola criminale. Ci ricorda infatti che anche l'occasione può far l'uomo ladro e che quindi se una tecnologia rende facile anche comportamenti illeciti (o non etici), spesso gli utenti le utilizzano.

Questa riflessione a prima vista sembra banale e quasi ingenua. Ma nei fatti non lo è, e fa riferimento al modo con cui le aziende sviluppano i *business plan* relativi alle nuove tecnologie. In questi tipi di progetto la sfida non sta solo nello sviluppare una nuova tecnologia ma anche nell'identificare il target a cui si pensa questa soluzione possa essere utile. Ed è qui che la riflessione dell'*Economist* è pertinente e suggestiva.

I Business Plan sono infatti sempre "buonisti" e ipotizzano consumatorii ideali, sempre etici e consapevolmente bisognosi di quel tipo di prestazione. Ma la tecnologia è sempre una spada a doppio taglio: controllando l'energia atomica si possono curare i tumori ma anche costruire la bomba atomica. Ed è ciò che capita sempre più frequentemente con l'innovazione digitale,

che i malfattori usano – spesso per primi – rileggendole e cogliendone altri possibili utilizzi. Ciò è per esempio capitato con le criptovalute, con la crittazione dei dati e con altre tecnologie di frontiera.

Il terzo punto da studiare con attenzione è il tema dell'etica, o meglio dell'(algor)etica secondo una felice espressione di Paolo Benanti, teologo e presbitero del Terzo ordine regolare di San Francesco ma soprattutto presidente della Commissione sull'Intelligenza Artificiale per l'informazione, organo della Presidenza del Consiglio dei Ministri.

Il collegamento tra etica e tecnologia non è episodico ma strutturale. Ce lo ricorda Paul Goodman nel suo *New Reformation* del 1971: "Dipenda o no dalla nuova ricerca scientifica, la tecnologia è un ramo della filosofia morale, non della scienza", perché attiene agli impatti, più o meno consapevoli, delle sue azioni. E il tema diventa esplosivo con il digitale per la sua infinita versatilità che lo rende capace potenzialmente di quasi qualsiasi cosa.

A ciò si somma il suo potere seduttivo, che rende il digitale potenzialmente pericoloso anche per gli utenti inconsapevoli. Osserva Sherry Turkle, psicanalista e studiosa dei rapporti fra digitale e psiche: "La tecnologia è seducente quando i suoi richiami incontrano la nostra umana vulnerabilità" (*Alone together*).

E infine la segretezza degli algoritmi complessifica ulteriormente la questione. Il tema è vasto; vediamone però un esempio concreto. Nel 1967 la filosofa Philippa Foot formula il cosiddetto *Trolley Problem*:

> Un tram ferroviario ha perduto il controllo. Il guidatore non può frenare ma può solo azionare lo scambio tra

> i binari. A un certo punto si trova di fronte a un bivio: seguendo il percorso previsto, ci sono cinque persone sul binario; mentre sull'altro binario – che può far percorrere solo decidendo di azionare lo scambio – ce n'è solo una. In entrambi i casi, le persone moriranno nell'impatto. Cosa deve fare il guidatore? Subire passivamente quanto programmato e assistere alla morte di cinque persone o azionare deliberatamente lo scambio e ucciderne "solo" una?

Sembrava un esercizio del pensiero destinato a rimanere nell'ambito delle riflessioni filosofiche. La sua finalità era infatti porre l'attenzione sul fatto che non basta la comparazione numerica del male minore; vi è anche la differenza tra assistere e determinare una morte. Ma il recente dibattito sugli algoritmi che guidano la "macchina senza guidatore" richiama questa riflessione.

L'opinione comune ritiene che quando una macchina guida autonoma intercetti qualcuno che attraversa la sua strada debba decidere "scegliendo il male minore". Ha infatti generalmente due opzioni: investire chi si trova sulla sua strada, oppure uscire di strada col rischio di uccidere il passeggero (che con tutta probabilità è anche proprietario della stessa auto).

Un sondaggio fatto dalla rivista del MIT sull'argomento mostrava che tutti erano d'accordo per il male minore, ma molti aggiungevano che non avrebbero mai accettato di comprare un'auto che decideva deliberatamente di ucciderli…

Inoltre, vi è un aspetto poco discusso ma potenzialmente problematico: il fatto cioè che l'algoritmo utilizzato possa essere parametrico, cioè in grado di "contare" diversamente le persone a seconda del loro profilo (di colore, di opinioni politiche diverse rispetto a chi

governa, immigrato non regolarizzato, con precedenti penali …). Oggi, infatti, i sistemi di riconoscimento facciale – grazie all'esigenza di allenare le piattaforme generative di IA (e di evidenziare eventuali terroristi in zone pubbliche) – sono incredibilmente efficienti e agiscono in tempo reale. Ma il software che implementa l'"algoritmo di morte" è protetto da segreto industriale e mai sapremo le logiche del suo funzionamento.

In questa scia altre applicazioni ancora – soprattutto quelle in ambito militare come i droni a guida autonoma (o *unmanned weapons* altresì detti "robot killer") – hanno creato talmente tanti timori da far nascere un'iniziativa dal nome emblematico – *Future Of Life Institute* – il cui motto è "*Technology has given life the opportunity to flourish like never before... or to self destruct*" e a cui hanno aderito scienziati e imprenditori del calibro di Steve Wozniak e il recentemente scomparso Stephen *Hawking. Un'azione importante di questo gruppo è* stata una lettera aperta firmata da centinaia di ricercatori ed esperti di Intelligenza Artificiale *e* presentata il 27 luglio 2015 durante la *International Joint Conference on Artificial Intelligence* di Buenos Aires, che chiede alle nazioni del mondo di bandire l'utilizzo bellico dell'Intelligenza Artificiale.

Un altro aspetto del nesso multiforme fra digitale ed etica è legato all'articolo 3 della nostra Costituzione, che recita: "Tutti i cittadini hanno pari dignità sociale e sono eguali davanti alla legge, senza distinzione di sesso, di razza, di lingua, di religione, di opinioni politiche, di condizioni personali e sociali".

Nel frattempo, però, il cittadino tende ad "articolarsi" grazie all'emergenza di nuovi "diritti/bisogni", che

a loro volta definiscono e/o dipendono da nuove forme di "normalità": donna, bambino, anziano, portatore di handicap… Inoltre, i progressi delle tecnologie (soprattutto biologia, genomica e digitale) "dilatano" i confini ontologici del cittadino, creando tre nuove categorie:

i *quasi-umani*: animali antropomorfi e con un sentire "simil-umano";

gli *oltre-umani*: esseri umani "aumentati" grazie alla tecnologia (steroidi, psicofarmaci, protesi, esoscheletri…);

i *simil-umani*: la crescente schiera di robot e androidi più o meno antropomorfi.

Questa suddivisione, apparentemente teorica, ha già implicazioni estremamente pratiche.

Nel primo caso, ad esempio, si può citare il caso dell'orango Sandra, una scimmia che è stata liberata dallo Zoo di Buenos Aires dopo un lungo processo intentato dagli animalisti argentini. La causa della liberazione è stata dimostrare che l'animale provava sentimenti "umani" ed era trattenuto ingiustamente in cattività senza un *Habeas Corpus*.

Considerando gli "oltre-umani" è utile richiamare il caso di Oscar Pistorius, un atleta diversamente abile, che si è ritrovato al centro di un dibattito quando la Federazione Internazionale di atletica leggera gli ha impedito di partecipare a una gara con i normodotati e di conseguenza di partecipare alle Olimpiadi di Pechino (gennaio 2008) perché, con le sue protesi, riusciva a correre più velocemente degli altri.

Un altro esempio è particolarmente interessante per le sue implicazioni etiche e giuridiche. Nel 2011, il fotografo David Slater stava facendo un reportage sui maca-

chi neri crestati dell'isola di Sulawesi, quando una femmina adulta si avvicina alla sua macchina fotografica – montata su un cavalletto – e vede incuriosita il proprio riflesso nella lente. Allora inizia a toccare la macchina, facendo scattare casualmente l'otturatore e producendo inconsapevolmente una serie di autoritratti. Questi selfie sono stati pubblicati su molte copertine e sono diventati virali nella Rete, tanto che Wikipedia ha deciso di utilizzarne una per illustrare la voce «Macaco nero crestato». Slater si è sentito allora danneggiato da questa riproduzione non autorizzata – dichiarando di aver perso almeno 10.000 sterline di guadagno e che la sua attività di fotografo naturalista era stata compromessa – e ne ha chiesto la rimozione. Wikipedia ha sostenuto che si trattava di una foto fatta da un animale e che quindi era di dominio pubblico. La controversia è finita in tribunale e, nell'agosto del 2014, lo *US Copyright Office* ha emesso un'ampia sentenza di oltre mille pagine nella quale ha dichiarato che le opere prive di paternità umana, come "una fotografia scattata da una scimmia", non possono essere registrate presso l'Ufficio del copyright degli Stati Uniti. Slate ha dunque perso la causa.

Il tema non è ovviamente arrestare il progresso, ma non dobbiamo neanche fare gli struzzi e mettere la testa sottoterra per non vedere. Dobbiamo esercitare con maggiore frequenza il principio di responsabilità enunciato dal filosofo Hans Jonas e alla base del suo principio legislativo di precauzione – principio sancito nella Conferenza delle Nazioni Unite sull'Ambiente e Sviluppo di Rio de Janeiro nel 1992: "Non si deve mai fare dell'esistenza o dell'essenza dell'uomo globalmente inteso una posta in gioco nelle scommesse dell'agire".

L'uomo deve dunque agire in modo che le conseguenze della sua azione siano "compatibili con la sopravvivenza della vita umana sulla terra". L'imperativo fondato sul principio di responsabilità, che nella sua forma semplice presuppone una tendenziale ipotizzabilità delle conseguenze, deve essere pertanto adattato al fatto che questa responsabilità si estende ora all'ignoto. Questo principio assume pertanto anche una sorta di "responsabilità intergenerazionale", poiché vuole rappresentare anche gli interessi delle generazioni future, che oggi non hanno rappresentanza (formale e autorevole) – e quindi voce in capitolo –su ciò che condizionerà il loro presente.

Si tratta dunque di riapplicare in modo sistematico il pensiero critico anche all'innovazione tecnologica. D'altra parte, il digitale non è solo una tecnologia ma un vero e proprio "ambiente" e quindi ben si presta a un'applicazione del principio di precauzione.

Dobbiamo quindi usare in modo sistematico il pensiero critico non per creare alibi al non fare, ma per costruire su fondamenta solide. Lo diceva Cartesio nelle sue *Meditationes de prima philosophia* – "Il dubbio è l'origine della saggezza" – e lo riprende in modo icastico Bertrand Russell (*Storia della filosofia occidentale*): "Il problema dell'umanità è che gli sciocchi e i fanatici sono estremamente sicuri di loro stessi, mentre le persone più sagge sono piene di dubbi".

5. *Dalla muscolarità tecnologica all'arte di arrangiarsi*

"*Less is more*" aveva lapidariamente affermato il grande architetto Ludwig Mies van der Rohe. "Weniger

aber besser" (*Less but better*) gli aveva risposto Dieter Rams, perfezionando la visione di Mies. Chi avrebbe pensato che questa massima del minimalismo architettonico ed espressione di una specifica visione estetica avrebbe influito in modo così rilevante nel mondo del digitale.

Innanzitutto, partendo proprio da Dieter Rams. Il grande designer degli anni '50 della Braun, ha dato forma a prodotti iconici come giradischi, bilance, sveglie e soprattutto radio. Non è un segreto che Rams sia stato la musa ispiratrice di Jonathan Yves, colui che ha plasmato l'estetica dei prodotti Apple: di ciò, infatti, non ha mai fatto mistero, tanto che si è addirittura parlato delle radici radiofoniche del design Apple.

In tempi più recenti vi è il caso di *DeepSeek*, che segna una svolta sull'approccio *mainstream* all'innovazione digitale oggi guidato dai colossi statunitensi del *bigtech*, e ci fa capire che d'ora in poi cambierà molto, e rapidamente, la lettura dei fenomeni tecnologici, mettendo in luce anche la loro filigrana geopolitica.

La nascita di *DeepSeek* sembra una risposta al modello *Maga – Make America Great Again* – sposato fin dalla prim'ora dalle principali aziende dell'HiTech digitale americano. La loro aderenza al modello non dipende solo dall'essere diventati tecno-vassalli (uso l'efficace espressione coniata da Lucio Caracciolo su Limes) come la foto dell'*Inauguration Day* per la nomina di Trump ha mostrato a tutto il mondo. Dipende – in maniera molto più profonda – dall'uso spregiudicato e sprecone delle risorse: finanziarie, energetiche e naturali (le terre rare). Il modello statunitense punta a creare una leadership protetta non dalla qualità dell'offerta ma dalla quantità di risorse necessarie per contrastarla.

Ed è qui che entra il suggerimento del grande designer tedesco. Il modello *DeepSeek* è una sorta di versione cinese della *Yougad innovation*, l'innovazione frugale, teorizzata dieci anni fa dagli autori indiani Navi Radjou, Jaideep Prabhu, Simone Aùehuja e appartenente allo stesso filone dell'arte di arrangiarsi partenopea.

È un'abilità – molto sostenibile e a km zero – legata al riuscire a inventare proprio perché si ha poco, una reazione intelligente a limiti e vincoli (naturali o anche imposti, come i dazi o la brevettazione); una sorta di *design thinking* a base zero che nasce da un ragionamento: non abbiamo i soldi per crescere di potenza computazionale (e forse è anche inutile e dispendioso farlo) né possediamo fonti naturali di energia e quindi dobbiamo necessariamente inventare un sistema capace di usare al meglio quello di cui già disponiamo.

Ricorda molto l'intelligenza metica greca, quella arguzia – legata al pensiero laterale, all'intuizione femminile, all'abilità mimetica di alcuni animali – chiamata dagli antichi *metis* e contrapposta al *logos*, al pensiero razionale e calcolante, alla forza pura delle armi. Un'intelligenza pragmatica e adattiva di cui era molto dotato anche Ulisse, per questo motivo chiamato spesso da Omero *polumetis*, e cioè colui che possiede molte *metis*. E fu proprio questa sua straordinaria abilità, che gli consentì di vincere la guerra di Troia non con la forza ma con l'inganno seduttivo del cavallo di legno, e di sopravvivere al lunghissimo e periglioso viaggio di ritorno, superando situazioni complicatissime come la fuga dalla caverna di Polifemo.

Questo uso arguto del digitale per orientare l'innovazione è sempre più necessario. Non possiamo infatti

più permetterci di sprecare risorse. Per fare un esempio, il programma Word, cardine della suite Office di Microsoft, viene usato mediamente al 10% delle sue potenzialità ma tutti lo devono comprare pagando il prezzo pieno che copre lo sviluppo di tutte le funzionalità – anche di quel 90% non utilizzato. Oltretutto la sua crescente sofisticazione richiede macchine sempre più potenti e sofisticate. Per questo motivo ogni qualvolta cambiano le release dei software operativi, si devono comprare computer più performanti. Mentre – come dimostrano molti word processor in open source – ciò che serve è semplice, compatto e richiede poche risorse di calcolo.

Ecco perché *DeepSeek* ha spiazzato, soprattutto negli Stati Uniti. È come se il mondo si stesse dividendo – sul concetto di innovazione digitale (ma non solo …) – in due grandi blocchi: Nord-occidentale e Sud-orientale. Chi ha meno soldi, può comprare meno chip e dipende energeticamente da altri Paesi, deve inventare qualcosa di diverso e molto più sostenibile economicamente, che esca dal *main stream* del modello nato in Silicon Valley e strettamente legato al mondo della finanza, del Venture Capital e della borsa. Una parte rilevante di questa innovazione non è infatti finanziata da utili e dividendi ma dalla vendita di azioni il cui valore contiene la crescita che promettono. Incomincia però a serpeggiare il dubbio che questo "schema Ponzi" possa davvero saltare.

6. *Un'ultima riflessione sul sempre più necessario pensiero critico*

Come abbiamo visto, la competenza forse più importante nell'era dell'IA Generativa è il pensiero criti-

co. Un pensiero critico abile ad analizzare non solo la qualità dei contenuti che ci vengono forniti dalle piattaforme generative, ma anche il modo in cui vengono rappresentati, la logica e le argomentazioni sottese. Il tema non è tanto la falsità delle affermazioni quanto il possibile errore che si cela nel come anche fatti reali vengono composti insieme per costruire un ragionamento sensato.

Come direbbe Cicerone, dobbiamo porre l'attenzione non solo sull'*inventio* (la qualità e veridicità dei singoli elementi) ma anche sulla loro *dispositio* che, svela l'ordine logico, la connessione argomentata fra le parti. Ci ricorda infatti Quintiliano: "Come a chi innalza una costruzione non basta accumulare pietre e materiali ed altri arnesi utili all'edilizia, se non vi aggiunge abilità nel disporli e nel collocarli, così nell'eloquenza l'abbondanza di argomenti si ridurrà a un cumulo informe, se la *dispositio* non li legherà in un tutto armonioso e organico" (*Institutio Oratoria, VII, 1*). È infatti proprio nella *dispositio*, che si comprende quando il meccanismo combinatorio è una semplice giustapposizione di affermazioni singolarmente pertinenti ma senza logica d'insieme.

Questo è il cuore dell'arte indiziaria, di quel metodo inventato da Ippocrate e perfezionato da Peirce – passando naturalmente da Sherlock Holmes – che permette di sviluppare ipotesi interpretative e far emergere il senso collegando fra di loro in modo abduttivo (questo è il nome usato da Peirce) elementi apparentemente slegati. Metodo potentissimo ma fragile e cuore della creatività umana.

È quindi necessario sviluppare questa capacità, sen-

za cadere nei trabocchetti dei ragionamenti fallaci. Per questo motivo la competenza critica che ci permetterà di utilizzare correttamente le piattaforme generative è meno tecnica e più logica: è l'inossidabile arte retorica, quella *techne retorike* – insieme di tecnica e arte – che si è sviluppata nel Mediterraneo, partendo dalla Grecia ed arrivando in Italia, dove è stata perfezionata e sistematizzata.

Ma come possiamo sviluppare queste necessarie capacità retoriche visto che – come ha osservato il sociologo francese Gérald Bronner – "viviamo nella società dei creduloni superinformati"? (*La démocratie des crédules*). La conoscenza codificata è sempre più ricca, diffusa e accessibile, eppure sappiamo sempre di meno. L'ignoranza si diffonde e il pensiero critico sembra rarefatto e spesso si trasforma in sospettosità complottista e paranoica.

Nel lontano 1922 Giuseppe Prezzolini aveva già colto con lucidità i primi segnali di questa tendenza e rischio, tanto che pubblicò sulla rivista La *Rivoluzione liberale* una proposta: "Noi potremmo chiamarci la Congregazione degli Apoti, di "coloro che non la bevono", tanto non solo l'abitudine ma la generale volontà di berle è evidente e manifesta ovunque". Nonostante questo prezioso avvertimento, tendiamo oramai a berci di tutto… e siamo solo all'inizio dell'era dell'Intelligenza Artificiale.

C'è però una novità interessante anche se poco reclamizzata. Un manipolo di studiosi appartenenti a un gruppo di prestigiose università statunitensi ha ritenuto utile affrontare questo tema di petto e in modo sistematico e ha addirittura fondato una nuova disciplina:

l'*agnotology*, lo studio del farsi dell'ignoranza, e quindi del perduto e del dimenticato. La parola – derivata dal greco – vuole complementare l'*epistemology*: una studia come si forma la scienza e l'altra l'ignoranza.

L'ipotesi di partenza è potente: viviamo in un'epoca di ignoranza, ed è allora importante capire come è successo e perché. L'obiettivo è dunque esplorare come viene prodotta o mantenuta l'ignoranza nei contesti più diversi. I meccanismi sono molteplici; in particolare la negligenza deliberata o involontaria, la segretezza che sconfina nella soppressione, la distruzione sistematica di documenti, il riferimento a una tradizione indiscutibile che toglie ogni apertura al diverso e una miriade di forme di selettività culturale.

La disciplina risponde a un'accorata richiesta fatta dal grande scrittore statunitense Thomas Pynchon nel 1984: "Spesso non siamo consapevoli della portata e della struttura della nostra ignoranza. L'ignoranza non è solo uno spazio vuoto sulla mappa mentale di una persona. Ha contorni e coerenza, e per quanto ne so anche le regole di funzionamento. Quindi, come corollario alla scrittura di ciò che sappiamo, forse dovremmo aggiungere la familiarizzazione con la nostra ignoranza".

I saggi raccolti nel libro *Agnotology. The Making and Unmaking of Ignorance* (2008) mostrano che l'ignoranza è spesso più di una semplice assenza di conoscenza; può anche essere il risultato di lotte culturali e politiche. L'ignoranza ha infatti sia una storia che una geografia politica. Sono in azione molte forze, tra cui quelle economiche (ad es. il segreto industriale), quelle giuridiche (la *privacy*), quelle geopolitiche e legate alla sicurezza nazionale (il segreto militare) e, naturalmen-

te, anche i pregiudizi: "non approfondisco, così non mi faccio influenzare".

Interessante in particolare la riflessione fatta da Robert N. Proctor – docente di storia della scienza presso l'Università di Stanford – sulle varie tecniche adottate per una costruzione sistematica del dubbio e dell'ignoranza nell'industria del tabacco. È uno degli esempi più potenti e micidiali di "agnogenesi" dove i lobbisti sono riusciti per molto tempo e in modo sistematico e pervasivo a produrre dubbi sui rischi del fumo. Osserva in modo disincantato Proctor che non è ancora noto "quale genio del male abbia escogitato lo schema per associare la continua produzione di sigarette alla prudenza, usando la richiesta di "più ricerca" per rallentare la minaccia della regolamentazione, ma va considerato come uno dei più grandi trionfi della connivenza aziendale americana." … che i produttori di IA abbiano studiato il caso?

Una curiosità: il libro è introvabile! Risulta infatti fuori catalogo nonostante sia interessante, attuale e pubblicato dalla prestigiosa Stanford University Press nel 2008; io l'ho trovato, con fatica, su un sito di libri usati. Leggendolo ho forse capito perché è introvabile. È un libro scomodo perché svela molti dei meccanismi usati dal potere – non solo politico ma anche economico – per manipolare, contraffare e spostare l'attenzione dalle verità scomode.

7. *In sintesi*

Volendo provare a tracciare un'impossibile sintesi su un fenomeno di cui sappiamo ancora troppo poco in quanto è appena iniziato, la enuncerei così: la rivoluzio-

ne digitale – di cui le piattaforme di IA generativa sono una parte – è un'onda potente e inarrestabile. Certamente porterà grandi benefici ma va conosciuta in modo intimo e approfondito, per coglierne anche le dimensioni problematiche, vista la sua natura di *pharmakon*. Ciò richiede molte competenze – non solo tecnologiche, ma anche umanistiche, sociologiche, psicologiche, economiche e giuridiche – con uno sguardo guidato dalla curiosità e dalla voglia di innovare ma complementato da un solido e costruttivo pensiero critico. Il fine ultimo dovrà essere, quindi, un uso efficace, equo e non dis-umano, trattando queste tecnologie più come allenatori che non come sostituti delle attività che caratterizzano l'essere umano.