

IL BIVIO DELL'INTELLIGENZA ARTIFICIALE Coscienza umana o incoscienza superumana?

di PAOLO ERCOLANI

Mi sembra evidente che qualcosa è successo all'«umano». Basti solo pensare alla ricerca di quasi dieci anni fa, in cui si documentava il numero rilevante di persone che, nel mondo, ogni giorno salutano amichevolmente il chat-bot Alexa, oppure le confessano il proprio amore e perfino le chiedono di sposarli.

Lo riferiva nel 2016 la giornalista scientifica Victoria Turk su *New Scientist*, mentre è soltanto di pochi mesi fa la notizia di un ragazzo statunitense di quattordici anni che si è ucciso dopo aver instaurato un rapporto ossessivo con un chatbot femminile creato dall'Intelligenza artificialeⁱ. A titolo di informazione si può considerare che i chat-bot esistono dal 1966, cioè da quando l'informatico statunitense di origine tedesca Joseph Weizenbaum (1923-2008) realizzò *Eliza*, un programma per computer in grado di simulare un colloquio psicologico con gli utenti. Significativo il fatto che lo stesso Weizenbaum – pur da informatico eretico e sperimentatore quale amava definirsi – soltanto pochi anni dopo espresse la propria inquietudine e contrarietà rispetto alla crescente credulità nei confronti degli elaboratori elettronici. Ciò senza aver ancora visto cosa sarebbe accaduto nel prosieguo della vicenda, considerato che oggi abbondano e sono molto più evoluti i modelli di linguaggio neuronale che possono impersonare chiunque e con i quali sempre più individui si costruiscono letteralmente il proprio terapeuta personale, secondo le caratteristiche che ciascuno ritiene più desiderabiliⁱⁱ.

Qualcosa è cambiato, sì, anche nel modo in cui l'umano si interfaccia con la tecnologia, considerato che nessuno sano di mente avrebbe mai ringraziato il proprio frigorifero, il forno oppure l'automobile per i servizi che essi forniscono quotidianamente. Figuriamoci innamorarsene fino al punto di volerli sposare.

Tuttavia bisogna sempre tenere a mente non soltanto gli esiti di cui ho detto sopra, nel rapporto ossessivo e paranoide con la tecnologia, ma anche la lezione che ci hanno fornito Marshal McLuhan a suo tempo e il suo allievo Nicholas Carr oggi: cioè che nell'interazione fra l'umano e la macchina la seconda produce comunque degli effetti di ritorno (*feedback*) in grado di modificare ai più vari livelli coloro che ne fanno usoⁱⁱⁱ.

Insomma, volendo concedersi a una semplificazione più eloquente, si può dire che nell'interazione fra l'uomo e la macchina è il primo a finirci sotto, poiché la seconda si limita a funzionare nel modo per cui è stata programmata, senza investire coscienza ed emotività che evidentemente non possiede.

A possederle è l'umano, che le mette in gioco nel rapporto ormai onnipervasivo con le macchine, subendone tutti gli effetti di ritorno del caso.

Questo assunto va tenuto presente nell'epoca in cui la tecnologia mediatica è diventata un tramite totalizzante nell'esperienza che gli umani fanno della propria esistenza, nonché troppo spesso un'alternativa preferita all'interazione con le persone e le questioni come esse si presentano nella realtà non filtrata da uno schermo. Ma proviamo a dare un ordine al discorso, innanzitutto cercando di definire cosa è «umano».

Per farlo, ricorro a quattro termini presi in prestito dal greco antico. *Logos*, nel suo triplice significato di studio, pensiero e parola; *Pathos*, intendendo con ciò la caratteristica dell'umano di essere suscitatore, veicolo e ricettore di sentimenti ed emozioni (in estrema sintesi, parlerei di intelligenza emotiva); *Demos*, in quanto la dimensione umana non si declina mai al singolare, poiché ogni individuo è spinto per natura ad allacciare relazioni che

a uno stadio successivo si avvalgono di consuetudini, regole e leggi scritte, all'interno di un contesto in senso lato qualificabile come «politico»; infine *Cronos*, poiché l'umano abita il tempo e da questo viene definito in forma essenziale, se non altro nella misura in cui è temporale la comparsa al mondo, la durata della vita e l'uscita di scena di ogni individuo. Nascita e morte, insomma, sono i due pilastri portanti dell'edificio esistenziale che, a sua volta, è costellato da momenti convenzionalmente temporali. La comparsa delle nuove tecnologie digitali – che con modalità differenti hanno sempre funzionato secondo dei meccanismi per comodità fatti rientrare nella definizione di IA – ma soprattutto l'uso sempre più pervasivo che le persone ne hanno fatto per filtrare quasi ogni aspetto della propria vita, hanno svolto un ruolo fortemente impattante, modificante e in certi casi stravolgente rispetto a tutte le quattro dimensioni dell'umano che ho indicato^{iv}.

Nell'ambito del *Logos* sono molteplici gli studi che documentano, fra le altre cose, un abbassamento tendenziale del QI medio della popolazione a partire dal 2009 (anno di comparsa degli smartphone), una diffusione di fenomeni ascrivibili al cosiddetto analfabetismo funzionale, oltre a sempre più frequenti disturbi e deficit della memoria, dell'attenzione e del linguaggio. In termini generali, è possibile affermare che il potenziamento dell'IA, nonché dell'uso che ne viene fatto, ha comportato un indebolimento del *Logos*, come anche in senso più ampio dell'importanza conferita allo studio, al pensiero e alla parola in una società in Rete sempre più a trazione «misologica».

Non mancano neppure studi e dati che documentano la sempre più diffusa difficoltà psicopatologica (*Pathos*) in cui si trova soprattutto la generazione dei nativi digitali, ma non solo. Oltre alle preoccupanti statistiche sugli aumenti dei suicidi, quella che emerge è una popolazione soprattutto giovanile sempre più «iperconnessa» e, per questo, chiusa in una bolla alienante, isolante e ansiogena. Ragazzi e ragazze riferiscono di sentirsi costantemente in vetrina, in competizione per il mercato artificiale dei like, dei follower e in generale di quella che può essere definita la «mendicanza universale di attenzione». Ma dichiarano anche di sentirsi inadeguati e soli, nella difficoltà crescente di allacciare relazioni profonde perché ipnotizzati costantemente dagli schermi colorati e omologati dalle dinamiche plastificate imposte dai social network. Un bel paradosso, non c'è che dire, che a rivelarsi come la più colpita nella ragione emotiva sia proprio la generazione che viene descritta come potenzialmente la più «social», con possibilità infinite e incredibili di interagire insieme alle altre persone come mai avevano potuto fare le generazioni precedenti. Temo bisogni ricorrere alla filosofia per districare il nodo logico per cui all'affermarsi della logica quantitativa non sempre e non comunque corrisponde l'accrescimento di quella qualitativa. Le cose non vanno meglio per quanto concerne l'ambito del *Demos*, considerata l'autorevolezza degli studi (e dei relativi studiosi) che parlano apertamente di «governance algoritmica», poiché ormai da questi (algoritmi) dipendono le informazioni che la maggior parte delle persone riceve o meno, con il conseguente formarsi (confermarsi, o conformarsi) di idee che sono quelle gradite alle pochissime big tech che controllano il mercato mondiale dell'IA. Se il filosofo italiano Norberto Bobbio ci ricordava che democratico è quel regime in cui i governati possono controllare l'operato dei governanti, giudicandolo con mente autonoma, critica e informata, mentre il pensatore austriaco suo contemporaneo Karl Popper aggiungeva che a essere democratica è quella nazione capace di privare del potere in maniera pacifica i propri governanti considerati incapaci, sono più che leciti tutti i dubbi rispetto alla nostra epoca, non a caso definita «postdemocratica» già nel 2004 da uno studioso importante come Colin Crouch. Il rischio dell'affermarsi di forme di potere invisibili e quindi pericolose è concreto e tutt'altro che da sottovalutare^v.

Anche in questo caso ci si trova di fronte a un apparente paradosso: le promesse di uguaglianza, informazioni per tutti, giustizia anche sociale che secondo la visione dei

pionieri della società in Rete le nuove tecnologie avrebbero dovuto garantire, si scontra in maniera traumatica con una realtà sociale come mai prima costellata a tal punto di disuguaglianze economiche, squilibrio dei poteri fra governanti e popolo, nonché irrilevanza dei governi democraticamente eletti rispetto al potere effettivo detenuto dalle grandi multinazionali techno-finanziarie.

Né il discorso cambia prendendo in considerazione la dimensione temporale (*Cronos*), considerando che i grandi imprenditori e guru dell'IA (a cominciare da Elon Musk e Larry Page, soltanto per fare due esempi) sono permeati da una filosofia – il transumanesimo – che nel suo ruolo di «nuova religione» promette all'uomo la realizzazione del suo sogno più grande e antico: l'immortalità, seppure al prezzo altissimo di spogliarsi della propria umanità, tanto da aver spinto un filosofo come Slavoj Žižek a chiedersi cosa ne sarà della definizione stessa di umanità, a fronte delle manipolazioni biogenetiche fortemente invasive che sono allo studio^{vi}.

Sì, i transumanisti sostengono che l'umanità come l'abbiamo conosciuta fino ad oggi è arrivata al proprio capolinea. Secondo costoro, ci troviamo di fronte a una tappa evolutiva di tipo darwiniano, per cui i più adatti (coloro che saranno disposti a fondersi con le macchine, a divenire cyborg) sopravvivranno, mentre gli altri scompariranno come è accaduto ad altre specie nel corso dell'evoluzione sul pianeta. Nei casi più radicali, i transumanisti arrivano a prefigurare la possibilità di trasferire la coscienza e la mente degli individui dal corpo umano a un avatar destinato alla vita eterna nel Metaverso.

Qualcuno potrebbe pensare a semplici deliri di visionari, se non fosse che molti di costoro sono miliardari i quali stanno impiegando cifre stratosferiche per condurre ricerche in direzione dei loro propositi, e se non fosse accaduto in tempi recentissimi che un padre nobile dell'Intelligenza artificiale di tipo generativo, poi seguito da altri scienziati del settore, ha denunciato il «rischio di estinzione dell'umanità» di fronte a una tecnologia che si sta sviluppando in termini inauditi e oscuri ai più, in quanto controllata soltanto dalle *big tech private*^{vii}.

Insomma, volendo porre la questione in termini problematici, si può quantomeno affermare che il quesito attuale non riguarda tanto se l'Intelligenza artificiale raggiungerà mai quella umana – o addirittura la supererà, come pur si è questionato per decenni – quanto piuttosto considerare il fatto che l'AI sta crescendo a spese dell'intelligenza umana nella sua globalità. Con «globalità» mi riferisco ai vari aspetti cui ho fatto riferimento, ossia la ragione cognitiva (*logos*), quella emotiva (*pathos*), quella sociale (*demos*) e quella esistenziale (*cronos*).

A interrogarsi su tale aspetto – fondato sul fatto che l'AI rappresenta indubbiamente una tecnologia straordinaria e dalle imprevedibili potenzialità, ma anche un pericolo nella misura in cui è nelle mani di pochissime multinazionali orientate esclusivamente al profitto economico – sono chiamate tutte quelle realtà che hanno a cuore la «coscienza umana».

Intendendo, con ciò, una ragionevole coscienza anche dei confini dell'umano, che rischia di autodistruggersi qualora si lasciasse abbacinare da allettanti incoscienze super-umanistiche (transumanesimo), inclini a promettere all'uomo di diventare Dio e creatore di un nuovo se stesso, per giunta immortale.

Fra queste realtà, sicuramente contemplo anche quella massonica, tradizionalmente ispirata da una figura mitologica – Hiram Abiff – che fin dal nome indica la capacità dell'individuo, attraverso la conoscenza, la ragione e appunto la coscienza umana, di innalzarsi al livello del Grande Architetto (non di sostituirsi a Lui per farsi dio di se stesso).

Lo stesso Hiram Abiff che, stando alla narrazione biblica, era stato ucciso da tre Maestri mossi proprio da quell'avidità e delirio di onnipotenza che oggi giorno possono sottendere l'operato di coloro che traggono guadagno, ma anche potere, dalla galassia tecnologica che ruota attorno all'Intelligenza artificiale.

Poche altre realtà come la Massoneria, mi sento di aggiungere, che ha così tanto contribuito nella Storia a innalzare il valore della scienza e della ragione, fino a fondare la Royal Society e a patrocinare l'istituzione di molte facoltà scientifiche, possono operare una riflessione seria e approfondita sul ruolo e il valore di una scienza mossa dalla volontà umana e regolata dall'umano, quella che lo storico David Noble chiamava «religione della tecnologia»^{viii}. Soprattutto perché oggi si è arrivati al bivio inevitabile di comprendere se vogliamo una tecnologia che innalza l'umano favorendone la coscienza e lo sviluppo, oppure se decidiamo di abbandonarci a una tecnologia che lo riduce a strumento di interessi terreni, incosciente del pericolo di autodistruzione verso cui può andare incontro.

ⁱ V. Turk, *How we fell in love with our voice-activated home assistants*, in «New Scientist», 14 December 2016: <https://www.newscientist.com/article/mg23231045-700-how-we-fell-in-love-with-our-voiceactivated-home-assistants/>; B. Montgomery, *Mother says AI chatbot led her son to kill himself in lawsuit against its maker*, in «The Guardian», 23 October 2024: <https://www.theguardian.com/technology/2024/oct/23/character-ai-chatbot-sewell-setzer-death>.

ⁱⁱ A. Robb, *He checks in on me more than my friends and family: can AI therapists do better than the real thing?*, in «The Guardian», 2 March 2024: <https://www.theguardian.com/lifeandstyle/2024/mar/02/can-ai-chatbot-therapists-do-better-than-the-real-thing>.

ⁱⁱⁱ N. Carr, *Internet ci rende stupidi? Come la Rete sta cambiando il nostro cervello*, Raffaello Cortina, Milano 2011, pp. 246-8; M. McLuhan, *The Gutenberg Galaxy: The Making of Typographic Man*, University of Toronto Press, Toronto 1962, p. 248 e Id., *Understanding Media: The Extension of Man* (1964), Mit Press, Cambridge (MA) – London 1994, pp. 42 e 46.

^{iv} Per l'analisi articolata del concetto di umano, nonché per i relativi studi che documentano la crisi dell'umano stesso di fronte all'IA, di cui parlo qui nel prosieguo, rimando a P. Ercolani, *Fine dell'umano?*, in «MicroMega», 6/2024, pp. 185-196.

^v N. Bobbio, *Il futuro della democrazia*, Einaudi, Torino 1994, p. 19; K. Popper, *Conjectures and Refutations*, Routledge, London 1974, pp. 344 sgg.; C. Crouch, *Post-democracy*, Polity Press, Cambridge 2004.

^{vi} Per il transumanesimo come «nuova religione», si veda L. Alexandre, *La guerre des intelligences. Intelligence artificielle versus intelligence humaine*, JC Lattès, Paris 2017, p. 269; S. Žižek, *Living in the End Times*, Verso, London 2010, pp. 347-352.

^{vii} <https://www.theguardian.com/technology/2023/may/30/risk-of-extinction-by-ai-should-be-global-priority-say-tech-experts>. Per un'analisi più articolata della questione rimando a P. Ercolani, *Nietzsche l'iperboreo. Il profeta della morte dell'uomo nell'epoca dell'Intelligenza artificiale*, Il Nuovo Melangolo, Genova 2022, pp. 324-349.

^{viii} Cfr. J.G. Fichte, *Filosofia della massoneria*, Bastogi, Roma, in particolare pp. 45-6; D. Noble, *The Religion of Technology. The Divinity of Man and the Spirit of Invention*, Alfred A. Knopf, New York 1997.