Una macchina può essere “intelligente”?



Luca Grion | 30 Dicembre 2023

Quella artificiale non può essere un’autentica intelligenza, ma solo una sua modellizzazione. Qualcosa che semplifica la complessità del vivente per offrirci strumenti di indubbia utilità. IA è un mezzo pensato e realizzato per consentire all’uomo di esprimere il meglio di sè. La via alternativa, quella che riduce l’umano alle sue capacità funzionali e ne misura il valore in termini di performance, rischia invece di alimentare un senso di inferiorità nei confronti dell’IA e di promuovere una meccanizzazione dell’umano, anziché una umanizzazione delle macchine. Una prospettiva, credo, tutt’altro che auspicabile

Dal punto di vista filosofico, uno degli aspetti più dibattuti, quando si riflette sull’*intelligenza artificiale*, riguarda i termini stessi che utilizziamo per articolare tale questione. Infatti, a seconda del significato attribuito al termine “intelligenza”, possiamo ritenere corretto o scorretto il ricorso a tale sostantivo per qualificare le cosiddette “macchine sapienti”. Il che comporta, inevitabilmente, porre la questione su ciò che accomuna e su ciò che distingue l’uomo dagli agenti algoritmici.

Per quanto brevemente, provo a chiarire meglio i termini della questione. Nel farlo prendo le mosse da una breve ricognizione storica. Il *copyright* dell’espressione “intelligenza artificiale” va riconosciuto all’informatico americano Jhon McCarthy, il quale aveva a cuore due cose: innanzi tutto distinguersi da un campo di ricerca per certi versi affine, ovvero la cibernetica (capitanata da un altro statunitense, il matematico Norbert Wiener); dall’altro trovare una definizione capace di attrarre i finanziamenti necessari alla costituzione di un gruppo di lavoro che avrebbe dovuto incontrarsi al *Darmouth College* (nel New Hampshire) nell’estate del 1956. Così, nel 1955, sottoponendo la sua richiesta alla *Rockefeller Foundation*,

McCarthy inventò un sintagma che avrebbe avuto una straordinaria fortuna, soprattutto a motivo delle suggestioni che veicola. McCarthy, però, non se ne rese immediatamente conto e pare che non fosse neppure particolarmente entusiasta della scelta, dato che gli interessava studiare l'intelligenza "umana", non quella "artificiale"; tuttavia, come ebbe a dichiarare lui stesso, «un nome dovevo pur darglielo»[1].

Ai fini del nostro discorso è interessante notare come chi coniò l'espressione "intelligenza artificiale" avesse ben chiara la differenza tra uomo e macchina. McCarthy riteneva però possibile realizzare artefatti in grado di simulare compiti che, ordinariamente, vengono svolti da esseri umani. Non a caso, battezzando l'IA, egli pensava semplicemente di indicare con un nuovo nome ciò che prima veniva chiamato "simulazione computazionale", ovvero il tentativo di far svolgere alle macchine operazioni tipiche dell'uomo. Un "gioco di imitazione" – per dirla con le parole usate da Alan Turing – il cui obiettivo è quello di riprodurre le funzioni tipiche di una mente pensante. Nel saggio pubblicato nel 1950 su "Mind", Turing si chiedeva infatti se una macchina potesse pensare o, detto altrimenti, se potesse esser detta intelligente[2]. Per rispondere in modo affermativo occorreva creare un calcolatore in grado di svolgere compiti solitamente considerati intelligenti (ad esempio rispondere in modo autonomo e credibile a una serie di quesiti), pur senza possedere un cervello o una coscienza umana. Oggigiorno questa prospettiva sembra ripresa da quanti ritengono che il predicato dell'intelligenza non debba essere limitato agli agenti umani. Penso, ad esempio, alla posizione di Nello Cristianini, il quale definisce "intelligente" ogni sistema in grado di agire nel suo ambiente, usando informazioni sensoriali per prendere decisioni, e di comportarsi in modo efficiente in situazioni nuove[3]. Capacità che, a suo avviso, non richiede né un cervello, né una coscienza. Si tratta di una intelligenza "altra", aliena; per certi aspetti più efficace di quella umana, proprio in virtù delle sue peculiarità tecniche.

Lasciamo per un attimo questi aspetti, che riprenderò nelle battute finali, e torniamo al senso della ricerca promossa da McCarthy. Il suo progetto si basava su una congettura fondamentale, ovvero sulla persuasione che «ogni aspetto dell'apprendimento o qualsiasi altra caratteristica dell'intelligenza è, in linea di principio, descrivibile con precisione tale da poter costruire una macchina capace di simularla»[4]. Come nel caso di

Turing, anche qui torna il tema della simulazione, all'interno della più ampia persuasione circa la differenza essenziale tra "naturale" e "artificiale".

Mantenere ben chiare queste premesse risulta di notevole interesse per la riflessione etica. Osservate da tale prospettiva, per quanto sorprendenti possano essere le prestazioni raggiunte dall'intelligenza artificiale, esse restano mere imitazioni di comportamenti umani. Benché possano apparirci "umane", esse non lo sono. E anche quando sembrano possedere emozioni o un senso del sé – come nel celebre film con Robin Williams, *L'uomo bicentenario* – le macchine non sono soggetti coscienti o senzienti. Ciò non di meno, proprio in forza della loro verosimiglianza mimetica, esse ci possono ingannare, dandoci l'illusione di soggetti in grado di ascoltarci, comprenderci, sostenerci. Noi, infatti, e questo è un aspetto non banale anche dal punto di vista etico, tendiamo a empatizzare con ciò che appare simile a noi, il che comporta, nel nostro rapporto con l'IA, il rischio di confondere cose e persone. Un pericolo di cui Sherry Turkle ha trattato diffusamente[5], soprattutto in riferimento a ciò che si perde quando ci si accontenta di rappresentazioni, anziché sperimentare relazioni autentiche.

La congettura di McCarthy potrebbe essere declinata anche in una diversa direzione. Essa, come detto, presuppone la riproducibilità artificiale di tutte le funzioni ordinariamente svolte da all'uomo. Ora, innestando tale presupposto su una concezione materialista e funzionalista dell'essere umano, giungiamo all'ipotesi che l'intelligenza umana possa essere non soltanto simulata, ma ricreata su un supporto non biologico. Infatti, così come ciò che chiamiamo "mente" emerge dalla complessità dei circuiti cerebrali – ovvero da una fitta rete di enti fisici, nessuno dei quali dotato di coscienza o di pensiero – allo stesso modo un analogo prodotto potrebbe emergere dalla complessità dei circuiti fisici di un calcolatore elettronico, debitamente costruito[6]. Pertanto, sostengono i fautori di tale prospettiva, laddove si realizzassero artefatti capaci di riprodurre alla perfezione i comportamenti umani – come nel caso del replicante a cui Rick Deckard dà la caccia in *Blade Runner* – non avremmo motivo per negare loro la qualifica di persone. A meno di non voler difendere forme di specismo antropocentrico.

Tali questioni, mi permetto di osservare, sono meno fantascientifiche di quanto si possa pensare. Infatti, a mano a mano che deleghiamo compiti deliberativi a sistemi algoritmici dotati di elevata competenza decisionale, ci

ritroviamo a chiederci se non si debba riconoscere loro anche la responsabilità di quelle scelte. In altre parole, poiché si attribuisce alle macchine la capacità di decidere autonomamente, si vorrebbe imputare loro anche la responsabilità morale delle azioni prodotte. Di fatto, dunque, le si pensa come vere e proprie "persone elettroniche", e non è un caso che tale espressione sia entrata prepotentemente nel dibattito giuridico. Dovremmo, però, riflettere con attenzione su questi slittamenti semantici. Ci stiamo abituando, infatti, a considerare normale che nostri simili, qualora incapaci di manifestare in atto alcune funzioni tipiche dell'essere umano, non siano più persone, mentre iniziamo a pensare che lo possano essere alcune realtà artificiali. In questione, per dirla con le parole di Robert Spaemann, c'è la differenza tra "qualcosa" e "qualcuno"[7]. Senza considerare i rischi di deresponsabilizzazione dell'umano che tali prospettive prefigurano.

Esiste un ulteriore modo di considerare il rapporto tra uomo e macchina? Secondo l'approccio anti-riduzionista – tipico, ad esempio, della tradizione aristotelico-tomista – vi è una differenza essenziale tra naturale e artificiale o, meglio, tra vivente e inanimato. Da questa prospettiva, com'era solito osservare anche Enrico Berti[8], espressioni quali "intelligenza artificiale", "cervelli elettronici" o "macchine pensanti" rappresenterebbero vere e proprie contraddizioni in termini, in quanto predicano capacità specifiche dei viventi a realtà inanimate quali le macchine. Dotati di vita, infatti, sono solo gli enti naturali, che hanno in loro stessi il principio del proprio movimento. Per essi, quanto vi è di più tipico è proprio il vivere, ovvero l'esercitare autonomamente le proprie funzioni specifiche; a differenza degli artefatti che, certo, possono svolgere determinate funzioni (anche complesse), ma solo se messi in movimento da altro (solitamente un uomo). Anche l'algoritmo più sofisticato, se ci pensiamo, necessita del contributo determinante dell'essere umano per essere messo in condizione di svolgere i compiti che gli vengono affidati.

Ora, ciò che chiamiamo "pensiero" rappresenta la funzione propria dell'intelligenza, ossia un'attività tipica di quel vivente che chiamiamo uomo. Il pensare, dunque, si predica propriamente di quel vivente che è l'uomo e si esplica nelle forme di un'intelligenza viva, consapevole, capace di farsi intenzionalmente una con la realtà che conosce. Il pensare, pertanto, non è riducibile al computare informazioni, ma comporta la presenza intenzionale dei dati di coscienza al soggetto conoscente. L'esperienza in prima persona – l'esperienza soggettiva dei cosiddetti *qualia* – è irriducibile a semplici

informazioni elaborabili da una macchina, proprio perché quest'ultima non è un vivente. Fa, ma non sa di fare[9].

Quella artificiale, dunque, non può essere un'autentica intelligenza, ma solo una sua modellizzazione. Qualcosa che semplifica la complessità del vivente – di per sé irriducibile a mero meccanismo – per offrirci strumenti di indubbia utilità. Ed è proprio su questa differenza tra cose e persone, su questa riaffermazione della natura strumentale della macchina, che l'uomo può edificare un autentico "umanesimo digitale", ovvero pensare a una *partnership* con le macchine che gli consenta di realizzarsi in pienezza. IA, dunque, come mezzo pensato e realizzato per consentire all'uomo di esprimere il meglio di sè. La via alternativa, quella che riduce l'umano alle sue capacità funzionali e ne misura il valore in termini di performance, rischia invece di alimentare un senso di inferiorità nei confronti dell'IA e di promuovere una meccanizzazione dell'umano, anziché una umanizzazione delle macchine. Una prospettiva, credo, tutt'altro che auspicabile.

Note

1 Ne parala M. Mitchell nel suo *L'intelligenza artificiale. Una guida per esseri umani pensanti* (2019), tr. it. Einaudi, Torino 2022, p. 6).

2 A. Turing, *Computing Machinery and Intelligence*, in "Mind", 49, 1950, pp. 433-460.

3 N. Cristianini, *La scorciatoia. Come le macchine sono diventate intelligenti senza pensare in modo umano*, il Mulino, Bologna 2023, p. 13.

4 D. R. Hofstadter, *Gödel, Escher, Bach: un'Eterna Ghirlanda Brillante* (1979), tr. it. Adelphi, Milano, 1984, p. 733.

5 S. Turkle, *Insieme ma soli. Perché ci aspettiamo sempre più dalla tecnologia è sempre meno dagli altri* (2011), tr. it. Einaudi, Torino 2019.

6 Cfr. P.M. Churchland e P.S.Churchland, *Could a Machine Think?*, in "Scientific American", 262, 1, 1990, pp. 32-39.

7 R. *Spaemann, Persone. Sulla differenza tra «qualcosa» e «qualcuno»* (2001), tr. it. Laterza, Roma-Bari 2007.

8 Cfr. E. Berti, Il *pensiero come forma di vita. A proposito dell'«intelligenza artificiale»*, in Id. *Nuovi studi aristotelici*, Morcelliana, Brescia 2005, vol. II, pp. 217-232.

9 Nonostante le obiezioni che gli sono state mosse, ritengo sempre di grande suggestione l'esperimento mentale ideato da Jhon Searle e noto come la "stanza cinese". Cfr. J.R. Searle, *Is the Brain's Mind a Computer Program?*, in "Scientific American", vol. 262, n. 1, gennaio 1990.

Tags: filosofia intelligenza artificiale intelligenza umana