



ἘΠΕΚΕΙΝΑ

International Journal of Ontology
History and Critics

CLAUDIA NAVARRA

Gerd Gigerenzer, *Perché l'intelligenza umana batte ancora gli algoritmi*, Raffaello Cortina, Milano 2023

EPEKEINA, vol. 16, nn. 1-2 (2023), pp. 1-7

Book Reviews

ISSN: 2281-3209

DOI: 10.7408/epkn.

Published on-line by:

CRF – CENTRO INTERNAZIONALE PER LA RICERCA FILOSOFICA

PALERMO (ITALY)

www.ricercafilosofica.it/epekeina



This work is licensed under a Creative Commons

Attribution-NonCommercial-NoDerivs 3.0 Unported License.

Gerd Gigerenzer, *Perché l'intelligenza umana batte ancora gli algoritmi*, Raffaello Cortina, Milano 2023

Claudia Navarra

Gerd Gigerenzer, psicologo cognitivo e direttore del Max Planck Institute for Human Development di Berlino, nel suo ultimo libro intitolato *Perché l'intelligenza umana batte ancora gli algoritmi*, edito nel febbraio 2023 da Raffaello Cortina Editore, si pone come obiettivo quello di salvare il lettore dalla narrazione da un lato catastrofica e dall'altro "messianica" riferita all'Intelligenza Artificiale.

Quella contro cui Gigerenzer concentra la sua attenzione è l'idea di una IA come *Superintelligenza* capace di sostituire l'essere umano in ogni campo, dagli scacchi al Go, dalla guida alla finanza, dall'educazione all'amore. Continuamente lo spettro di una IA che prende coscienza e riduce in schiavitù l'umanità assilla fior di studiosi – facendo la fortuna di altrettanti registi –, mentre altri accademici con la stessa intensità e CEOs di grandi aziende e istituti finanziari auspicano, invece, un mondo in cui ogni attività umana sarà svolta da una Intelligenza non-umana.

L'IA batte gli umani, è questo l'argomento dominante:

«l'IA ha sconfitto i migliori umani negli scacchi e nel Go.

La capacità di elaborazione raddoppia ogni paio d'anni.

Quindi, le macchine ben presto faranno ogni cosa meglio degli umani» (Gigerenzer 2023, 2).

Ma è davvero così? Siamo, in quanto specie, meno intelligenti delle macchine che abbiamo costruito? Esiste un effettivo pericolo di estinzione per la specie umana? Sarà una Superintelligenza a governarci tutti? L'erronea idea generale sembra essere quella per cui sostituire le persone con un software possa rendere il mondo un luogo migliore a patto, però, di rinunciare a mantenere il controllo sulla propria vita e ad avere una fiducia incondizionata nell'algoritmo. È la fede nella complessità e nell'opacità a garantire la fiducia nell'IA, cioè la credenza in quello che Gigerenzer chiama «dilemma accuratezza-trasparenza: più l'algoritmo è accurato, meno sarà trasparente» (2023, 140). L'uso, negli Stati Uniti, di algoritmi nelle aule di tribunale che aiutino il giudice a

decidere, l'utilizzo di IA da parte della polizia che predicono il luogo di future scene del crimine o quali persone hanno più possibilità di commettere o essere vittima di reati (Sappiamo chi sei!) e ancora le lunghe ed esaustive spiegazioni sul consenso al trattamento dei dati personali rendono il mondo un luogo più sicuro.

L'IA garantisce trasparenza e previene le situazioni di pericolo. Secondo Gigerenzer le cose stanno diversamente. Anzi, lungi dall'essere onnisciente e super intelligente, l'IA non ha speranze di raggiungere le prestazioni umane, almeno quando ci si muove nel mondo vero che è un mondo dominato dall'incertezza la cui causa sono proprio gli esseri umani.

L'analisi pungente, e a tratti provocatoria, dello psicologo tedesco tocca quelli che potremmo chiamare i "cavalli di battaglia" dell'IA: l'amore, la privacy, la salute, il denaro, la giustizia, senza dimenticare uno dei settori di cui ultimamente ci si è interessati di più: la guida automatica. In ognuno di questi settori l'IA ha avuto una evoluzione straordinaria e ha messo in circolazione, più o meno virtuosamente, milioni di dollari promettendo altissime prestazioni e garantendo la capacità di prendere la decisione migliore possibile, nonché di riuscire a prevedere eventi, soprattutto di natura economica, che possano permettere alle aziende di aumentare i loro profitti con sicurezza.

Gigerenzer, in poco meno di 400 pagine e 11 capitoli divisi in due parti, ci mostra come non esista nessuna garanzia e nessuna prova che, in effetti, l'IA, con i suoi algoritmi complessi, riesca a fare quello che promette né a battere l'essere umano quando si tratta di giocare sul serio, quando cioè le situazioni non sono stabili ma cariche di incertezza. L'algoritmo, insiste Gigerenzer, funziona nelle condizioni in cui è possibile applicare quello che egli definisce come il *Principio del mondo stabile*: «gli algoritmi complessi funzionano al meglio in situazioni ben definite, stabili, in cui sono disponibili grandi quantità di dati. L'intelligenza umana, invece, si è evoluta in modo da gestire l'incertezza a prescindere dalla quantità di dati disponibili» (2023, 22). Se è vero che l'IA batte gli umani in ambienti stabili e quando ha a disposizione una enorme quantità di dati (i *big data*), come quando gioca a scacchi, questo non significa che possa farlo in ogni situazione. Lo scopo di Gigerenzer con questo libro è quello di fornire al lettore strategie e metodi per mantenere il controllo della propria vita e, contemporaneamente, quello di comprendere le potenzialità delle nuove

tecnologie. Dunque, nessuna visione catastrofica o “messianica”, ma un’analisi chiara di ciò che l’IA può fare e dei vantaggi, così come dei pericoli, che ne può comportare l’utilizzo.

L’appello di Gigerenzer è di *Restare Intelligenti*. Significa comprendere vantaggi e rischi delle nuove tecnologie e mantenere il controllo. Le nuove tecnologie hanno, e avranno, un impatto profondissimo sugli esseri umani, poiché cambiano il modo in cui sono percepiti il bene e il male. La colpa di questo cambiamento non è dell’algoritmo in sé, ma è delle aziende che progettano l’algoritmo. L’autore definisce le aziende come delle macchine per predire e modificare il comportamento degli esseri umani: propongono soluzioni tecnologiche a problemi umani utilizzando un modello di business “paga con i tuoi dati”; alle aziende non interessa risolvere il problema, interessa il tempo e l’attenzione degli utenti. Infatti, spesso le aziende possono perseguire degli obiettivi che risultano in contrasto con quelli che i loro software dovrebbero riuscire a fare.

Pare opportuno presentare alcuni casi analizzati da Gigerenzer soprattutto nella prima parte del libro.

Il primo degli esempi di questo utilizzo dell’IA riguarda i siti d’incontri e cerca di rispondere ad una domanda sottesa al primo capitolo su come possa un algoritmo garantire di trovare l’anima gemella. Una delle più importanti aziende di questo settore dichiara¹: ogni 11 minuti un single si innamora! È sicuramente uno slogan interessante e funziona visto che sono milioni gli iscritti al sito. Ma a ben guardare – come fa Gigerenzer – se gli iscritti fossero soltanto un milione significherebbe che, in un anno, troverebbe l’amore soltanto il 5% di loro. In più, se sul sito fossero iscritti, come di fatto è, più di un milione di utenti ci vorrebbero più di dieci anni per trovare l’amore. Nonostante tutto però, i milioni di utenti restano e sperano e continuano a compilare questionari nella speranza che qualcun altro compili il questionario meglio compatibile. Gli algoritmi che utilizzano questi siti si fondano su tre principi cui vengono assegnati dei valori numerici. I tre principi base dell’algoritmo dell’amore sono: la somiglianza, la complementarità e l’importanza che insieme concorrono alla creazione

1. Il sito d’incontri cui si riferisce Gigerenzer è Parship, attivo con milioni di utenti e disponibile in 13 Paesi.

di un profilo utente; eppure, è raro che trovino l'anima gemella. Il motivo di questo fallimento risiede, secondo l'autore, nel fatto che un profilo non è una persona, infatti, manca l'apporto degli elementi fisici, chimici ed emotivi che sono fondamentali nelle relazioni umane, ma non solo. L'ambiente in cui operano gli algoritmi di questo tipo è colmo di incertezza che è, inoltre, alimentata anche dal fatto che i dati inseriti, molto spesso, non corrispondano alla realtà: reddito più alto e peso minore sono tra le "bugie" più comuni.

Il consiglio dello psicologo, oltre a calcolare in maniera più precisa le percentuali di successo dichiarate dalle aziende, è quello di andare a prendere un caffè e vedere come va. Un esempio ancora più chiaro di come le aziende progettino gli algoritmi al fine di manipolare il comportamento umano è fornito nel secondo capitolo. In questa sede Gigerenzer mostra come l'adozione da parte di aziende sanitarie statunitensi di un sistema basato su algoritmi per la creazione di cartelle sanitarie digitali che permettesse consultazioni, analisi, registrazioni e archiviazioni in modo veloce e funzionale, si rivelò un completo disastro. Furono previsti, in seguito all'adozione di un tale software già positivamente sperimentato in Gran Bretagna, circa 81 miliardi di dollari risparmiati all'anno. La realtà si rivelò ben diversa: «le fatturazioni degli ospedali aumentarono con costi che passarono da 2000 a 2800 miliardi di dollari» (2023, 31). L'IA avrebbe dovuto fungere da una sorta di velocissimo database interattivo, avendo a che fare con un mondo stabile di dati che venivano registrati e integrati. Il software, invece, era progettato non per snellire il sistema sanitario, ma per aumentare le richieste di interventi medici sui pazienti. Gigerenzer afferma infatti, che erano le stesse aziende sanitarie che si occupavano degli esami o delle operazioni o della produzione di dispositivi medici a progettare i software che, nell'ottica del profitto e della lealtà aziendale, consigliavano al medico inutili ulteriori esami o di cambiare, aggravandola, una diagnosi.

Due degli aspetti più importanti per gli esseri umani, l'amore e la salute, affidati all'Intelligenza Artificiale hanno subito e subiscono danni incalcolabili in nome del profitto. È, quello dello psicologo, un forte ammonimento al capitalismo. Dall'analisi che prende forma nel libro, appare chiaro come, anche se i dati raccolti nelle indagini dello studioso sembrano essere disastrosi, gli esseri umani continuano a fidarsi delle promesse delle aziende che sfruttano gli algoritmi. Ciò

avviene perché spesso le aziende si servono dei bias cognitivi, degli errori insiti nel modo di pensare delle persone. Uno dei più comuni bias, che confonde la predizione con l'adattamento, è la "Fallacia del cecchino texano": un cecchino spara alla parete di un fienile da una distanza considerevole; il 90% dei colpi sono raggruppati all'interno di un cerchio; come ha fatto a centrare così tante volte il bersaglio? Il cecchino texano ha prima sparato e poi disegnato il cerchio. Quando un'azienda tenta di vendere risultati sbalorditivi, consiglia al lettore Gigerenzer, è sempre importante chiedersi se «la performance esaltante di un algoritmo sia stata ottenuta attraverso una predizione o un mero adattamento» (2023, 35).

Il quarto capitolo è dedicato all'analisi dell'affidabilità delle auto a guida autonoma, tema delle più attuali indagini sulle effettive capacità dell'Intelligenza Artificiale. In questo tipo di tecnologia vengono investiti milioni di dollari e tollerati incidenti spaventosi. Le auto di nuova generazione hanno una autonomia di livello 5, il massimo possibile, il che significa che possono guidare in qualsiasi situazione e senza l'intervento dell'essere umano. Un nuovo principio viene introdotto da Gigerenzer, quello dell'adattamento dell'IA: «per migliorare la prestazione dell'IA, bisogna rendere l'ambiente fisico più stabile e il comportamento umano più prevedibile» (2023, 57). Questo significa che non è più la tecnologia a doversi adattare al comportamento umano, ma è l'essere umano a doversi adattare alla tecnologia in modo da evitare il più possibile le situazioni di incertezza. Inoltre, nonostante il migliore contributo possibile di una IA psicologica, che potrebbe istruire l'algoritmo delle auto a guida autonoma con le simulazioni delle percezioni, dei giudizi e dei processi decisionali umani, sarebbe impossibile per una AI imparare a guidare come un essere umano. Una delle regole, non scritte, della guida è: dietro un pallone c'è sempre un bambino. Un pilota umano saprebbe quando è il caso di rallentare, di frenare o di fare qualche altra azione sulla base di questa regola. Un pilota digitale frena, anche se dietro il pallone non c'è alcun bambino, e provoca un incidente a catena. Il vero pilota digitale sono le reti neurali profonde, che però commettono errori e scambiano, a 69 km all'ora, una donna in bicicletta prima per un oggetto sconosciuto, poi per un veicolo e infine identificano una bicicletta, a 4 secondi dall'impatto mortale. Secondo lo psicologo tedesco il problema delle reti neurali non consiste nel non avere abbastanza immagini di biciclette, ma nel

non possederne il concetto. La “Fallacia del carrarmato russo” è il tipo di errore che commettono le reti neurali e riguarda il fatto che spesso queste si focalizzano su elementi secondari che sono presenti sia nell’addestramento sia nella serie di test che vengono fatti. L’Intelligenza Artificiale con le reti neurali cerca pattern basati su certe caratteristiche; invece, l’intelligenza umana è in grado di costruirsi una rappresentazione del mondo. Ancora più importante è il dilemma morale che le auto a guida autonoma impongono: tre anziani stanno attraversando la strada col semaforo rosso mentre sta per raggiungerli un’auto a guida autonoma con tre passeggeri a bordo; chi ucciderà l’IA, i tre anziani o i tre passeggeri, schiantandosi su un muro per evitare i pedoni? Secondo Gigerenzer questo dilemma non interessa nemmeno agli ingegneri, proprio perché risolverlo significherebbe aver progettato una IA capace di prevedere ogni possibile conseguenza e di essere perfettamente autonoma, cosa ancora molto al di là dal divenire realtà.

È particolarmente chiara l’analisi del divario tra l’intelligenza umana e quella artificiale che si va tracciando nelle pagine del libro. Ci sono profonde differenze tra l’intelligenza umana e quella artificiale. La prima e fondamentale differenza è che l’intelligenza umana si è sviluppata in condizioni di incertezza, mentre quella artificiale in condizioni di rischio. Questa differenza di condizioni ha comportato lo sviluppo di attività differenti. Gli esseri umani hanno sviluppato capacità che tutte insieme Gigerenzer chiama “buonsenso”: il pensiero causale, la psicologia intuitiva, la fisica intuitiva e la socialità intuitiva. Il “buonsenso” è la «conoscenza condivisa sulle persone e sul mondo, resa possibile dal cervello biologico, e richiede un’esperienza limitata» (2023, 82). Le reti neurali hanno sviluppato il calcolo veloce, la scoperta di associazioni nei big data, il ritrovamento di pattern nelle immagini o nelle informazioni sonore.

L’errore di fondo, sostiene l’autore, è stato quello di confondere la capacità di calcolo con l’intelligenza; un maggior numero di informazioni con la conoscenza; la memoria con l’intuizione. All’Intelligenza Artificiale manca il buonsenso e questo si riscontra nel linguaggio, che spesso viene stravolto dall’IA che non trova corrispondenze adeguate all’interpretazione, come anche nel riconoscimento facciale oppure di un certo scenario.

Gigerenzer fornisce al lettore altri strumenti per mantenere il controllo sulla propria vita e restare intelligenti, tra cui l’euristica della

recenza: «in situazioni di incertezza, fatela semplice e non scommettete sul passato» (2023, 105); oppure suggerisce di non fidarsi dei grandi numeri, proprio perché i numeri *non* parlano da soli è sempre necessario ricondurli ad un mondo stabile e interpretarli mediante una teoria.

Un altro interessante esempio della inadeguatezza dell'IA nelle situazioni di incertezza e dell'interesse delle aziende nella manipolazione del comportamento umano riguarda la privacy. Gigerenzer afferma, nell'ottavo capitolo, che la privacy non vale un fico secco! Tutti restano costernati alla notizia del furto di dati sensibili o alla loro perdita o della vendita dei dati personali a terze parti, sentendosi vittime di una società di matrice orwelliana. Ma qui sta il paradosso: le stesse persone indignate e spaventate dal fatto che un qualche tipo di organizzazione possa sottrarre informazioni private, non si fa nessuno scrupolo né a rivelarle sulle piattaforme social né a negare di essere disposti a pagare per mantenere quelle informazioni così importanti al sicuro.

Dunque, è enorme l'effetto che l'IA ha sulla vita degli esseri umani, ma questo non significa che l'umanità sia diventata succube delle macchine, non ancora almeno. Per far sì che questo non avvenga è necessario comprendere da quali bias siano affetti gli algoritmi e in quale mondo siano applicabili. Gigerenzer sprona il lettore a restare intelligente, ad indagare e ad analizzare gli scenari condivisi con l'IA e a sfruttarne le potenzialità quando queste si rivelano preziose ma, nello stesso tempo, a dubitare delle promesse stupefacenti che le aziende tentano di vendere.

L'intelligenza umana batte ancora gli algoritmi perché è basata sul buon senso e sulla corporeità, sull'intuito e la creatività, sull'improvvisazione e l'incertezza. Sintesi perfetta tra capacità divulgativa e rigore scientifico, *Perché l'intelligenza umana batte ancora gli algoritmi* è un libro interessante e piacevole, che coinvolge e stimola il lettore munendolo di importanti strategie per vivere la propria vita in maniera consapevole insieme all'IA.

Claudia Navarra
Università degli Studi di Palermo
claudia.navarra@unipa.it