

Nello Cristianini, *La scorciatoia. Come le macchine sono diventate intelligenti senza pensare in modo umano*, il Mulino, Bologna 2023, 216 pp.

Cristina Rebuffo*

«*La scorciatoia* di Nello Cristianini è un libro affascinante che esplora l'intelligenza artificiale e il suo impatto sulla società. Cristianini, un esperto nel campo, affronta temi complessi con chiarezza, spiegando come l'I.A. non solo automatizzi processi, ma prenda anche decisioni cruciali che influenzano la vita quotidiana. Il titolo del libro si riferisce alla "scorciatoia" presa dall'I.A. per risolvere problemi, bypassando la comprensione umana tradizionale. L'autore riflette criticamente su questa dinamica, interrogandosi sulle implicazioni etiche e sociali. Un saggio stimolante che invita alla riflessione sul futuro dell'umanità in un mondo sempre più dominato dalla tecnologia».

Il precedente capoverso è stato elaborato da ChatGPT, acronimo di *Chat Generative Pre-trained Transformer*, il più celebre, attualmente (Agosto 2024), *chatbot* basato su intelligenza artificiale e apprendimento automatico, specializzato nell'interazione testuale tra macchina e utente umano. Il testo è stato elaborato da questo *software* in risposta al semplice *input* «Scrivi una breve recensione su *La scorciatoia* di Nello Cristianini»: si è deciso di anteporlo alla recensione vera e propria in maniera forse un poco provocatoria, ritenendo che possa essere una curiosa guida per l'analisi del volume in oggetto e come esempio pratico dei suoi contenuti.

L'attività di ricerca di Nello Cristianini, professore di Intelligenza Artificiale presso l'Università di Bath, in Gran Bretagna, è volta, in generale, all'analisi su larga scala di contenuti multimediali (notizie e *social media*) e all'impatto sociale dei cosiddetti *Big Data* e dell'Intelligenza Artificiale stessa; nello specifico, l'obiettivo del suo saggio qui presentato è quello di porre importanti questioni in merito alla necessità di trovare delle vie, delle strategie, per imparare a convivere con le nuove tecnologie basate sull'I.A., limitando il più possibile rischi ed effetti collaterali derivanti dal loro utilizzo deregolamentato, improvvisato e non ragionato. L'autore ritiene, in particolare, fondamentale che, per rispondere a tali importanti questioni e per impostare una relazione proficua ma sicura tra individui, comunità umane e politiche e "macchine intelligenti", sia assolutamente fondamentale attivare un dialogo tra le scienze naturali, la scienza informatica e le scienze umane e sociali.

Non a caso, il volume si apre con il capitolo intitolato *Alla ricerca dell'intelligenza*, in cui viene esposta una interessante disamina del concetto stesso di "intelligenza", che noi attribuiamo tanto alle facoltà umane quanto a quelle di alcuni tipi di macchine e tecnologie

* Lessico di Etica Pubblica, e-mail: rebuffocristina@gmail.com.

di cui ChatGPT fornisce un esempio particolarmente noto e controverso. Qui, l'autore assume un atteggiamento critico nei confronti delle conseguenze derivanti dall'uso, appunto, del concetto di “intelligenza” da una prospettiva esageratamente antropocentrica, che ci porta ad assumere l'aggettivo “intelligente” come sinonimo di “somigliante a un essere umano”; al contrario, sottolinea l'autore, «dimostrare intelligenza non significa assomigliare agli esseri umani, ma essere capaci di comportarsi in modo efficace in situazioni nuove. Questa capacità non richiede un cervello [...] non c'è un singolo modo di essere intelligenti [...] È fuorviante attribuire qualità umane a tutti gli esseri intelligenti»¹, siano essi animali non umani o macchine. Un comportamento intelligente è, dunque, secondo questa definizione, un atteggiamento teleologico, in quanto consiste nella comprensione e nello sfruttamento dell'ordine dell'ambiente in cui l'individuo opera col fine di trovare una soluzione a un certo problema, e questo può essere fatto indipendentemente dalla presenza di un cervello umano, di una coscienza umana, di un corpo umano; semmai, a essere necessario per poter agire con intelligenza, è che l'ambiente in cui si agisce sia un ambiente “regolare”, perché solo in presenza di regolarità è possibile, tanto per un individuo umano, quanto per un calamaro, quanto per un software, fare previsioni più o meno precise e certe sul futuro e dunque individuare la risposta più corretta a ogni situazione, ed è solo in presenza di regolarità che è inoltre possibile sviluppare esperienza e apprendimento. Ciò che propone Cristianini è, per sua stessa ammissione, una nuova “rivoluzione copernicana” che ci conduca finalmente a riconoscere intelligenze radicalmente alternative a quella umana, con l'obiettivo di «calibrare meglio quello che possiamo aspettarci da queste nostre creature»², che non saranno forse mai in grado di armonizzare una sinfonia per orchestra con lo stile di Beethoven o di comprendere se sia eticamente corretto utilizzare dati biometrici per scopi di lucro ma sono, ad esempio, in grado di interpretare un QRCode. Proprio a partire da questa “rivoluzione copernicana” circa la concezione dell'intelligenza si innesta l'idea della “scorciatoia” che dà il titolo al volume e che indica il *modus operandi* delle cosiddette intelligenze artificiali: ripercorrendo gli studi svolti negli ultimi decenni in merito al loro funzionamento, Cristianini perviene a indicare con tale termine il modo, alternativo o alieno all'intelligenza umana, con cui alcuni tipi di *software* sono in grado, a partire da regolarità statistiche presenti in un certo numero di dati o informazioni, di elaborare previsioni o soluzioni: questa scorciatoia è alla base, ad esempio, del funzionamento dei motori di ricerca sul Web o dei correttori automatici dei programmi di scrittura e messaggistica istantanea o dei suggerimenti personalizzati sui più comuni siti di *shopping online* o, ancora, della procedura con la quale ChatGPT ha elaborato il testo introduttivo a questa recensione. In tutti questi casi le intelligenze artificiali hanno però dovuto operare anche una seconda scorciatoia, relativa al reperimento di dati e informazioni in numero sufficiente per poter operare la prima, e una terza, consistente nell'apprendimento delle tendenze degli utenti a partire dall'osservazione dei loro comportamenti in risposta alle soluzioni via via proposte: ed ecco dunque che ritorna l'importanza, per l'agire intelligente, di un ambiente “regolare” a partire dal quale apprendere quale via seguire per proporre soluzioni sempre più precise ed efficienti. È questo il principio sul quale si fonda la tecnologia chiamata *machine learning*, che consente alle macchine di perfezionare le proprie abilità se non addirittura di acquisirne di nuove

¹ N. Cristianini, *La scorciatoia. Come le macchine sono diventate intelligenti senza pensare in modo umano*, il Mulino, Bologna 2023, p. 9.

² Ivi, p. 25.

senza aver ricevuto una specifica programmazione esterna, umana: «Il risultato sarà un comportamento originale, che potenzialmente il programmatore non avrebbe immaginato»³, un comportamento che quindi potremmo definire, con l'Autore, “sovrumano”, reso possibile dal fatto che la velocità e la mole delle esperienze maturate da una macchina sono del tutto estranee alle abilità umane.

Ora, alla luce di tutto ciò, è urgente chiederci «come possiamo assicurarci che le macchine non violino le fondamentali norme sociali, eseguendo “alla lettera” quello che viene chiesto loro»⁴ o nel loro agire intelligente “sovrumano”. Come possiamo, cioè, interagire con tecnologie che non sono solo teoreticamente ma anche eticamente aliene rispetto a noi? A partire dal capitolo 5 del volume, *Comportamenti imprevisti*, l'Autore inizia a evidenziare i potenziali (o reali) rischi ed effetti indesiderati provocati dall'uso indiscriminato delle intelligenze artificiali in molteplici ambiti nei quali se da un lato esse risultano estremamente efficienti nel pervenire a risultati attesi dagli utenti, dall'altro ignorano il rispetto di norme sociali ed etiche comunemente condivise. In merito a ciò, Cristianini fa presente che

al momento, gli scienziati e i filosofi stanno esplorando diverse dimensioni della «fiducia» (*trust*): trasparenza, equità, responsabilità, accuratezza e verificabilità (o ispezionabilità). Con verificabilità (o ispezionabilità) intendiamo che ogni strumento software sia concepito sia dal principio in modo da poter essere facilmente ispezionato da terzi, per esempio un'agenzia istituzionale.

Potrebbe essere una buona idea decidere che solo tecnologie «verificabili» o ispezionabili possano essere usate in settori regolamentati.⁵

La verificabilità dell'agire intelligente dei *software* di cui ci serviamo giornalmente, talvolta anche inconsapevolmente, sarebbe particolarmente utile, come sottolineato nel capitolo 6, per quanto riguarda *Messaggi personalizzati e persuasione di massa*, cioè in tutti i casi in cui le intelligenze artificiali vengono utilizzate non solo per analizzare e comprendere ma anche per influenzare i nostri comportamenti reali, compromettendo quello che l'Autore definisce il nostro “diritto all'autonomia”. Ci si riferisce, qui, ad esempio, alle situazioni in cui i dati personali sia pubblici che privati degli utenti di un *social network* vengono utilizzati per inferire tratti del comportamento non solo degli utenti presi in esame ma del comportamento pubblico anche di altri utenti, con lo scopo di indirizzare loro in maniera mirata messaggi pubblicitari, consigli per gli acquisti, notizie più o meno vere, posizioni politiche e consigli elettorali, influenzando anche il comportamento reale di quegli stessi utenti. Sono molti i casi di cronaca, anche mediaticamente molto esposti, che hanno mostrato le conseguenze di questo tipo di utilizzo dell'I.A.; basti pensare alle vicende riguardanti Cambridge Analytica e al suo ruolo fondamentale per l'elezione di Donald Trump a Presidente degli Stati Uniti nel 2016, grazie a una pericolosa intersezione tra psicologia comportamentale, scienza dei dati e tecnologia *machine learning*, che ha mostrato che il cosiddetto “targeting psicologico” è in grado di influenzare il comportamento di masse di utenti tramite messaggi persuasivi mirati, costruiti sulla base di specifici bisogni psicologici della platea a cui si rivolge. In casi come questo è evidente che l'intelligenza artificiale non si limita a fornire un servizio ma manipola pericolosamente il comportamento virtuale e reale di masse di utenti, senza che questi ne siano consapevoli o

³ Ivi, p. 75.

⁴ Ivi, p. 83.

⁵ Ivi, p. 97.

senza che abbiano fornito il proprio consenso, compromettendo pesantemente la loro autonomia nel prendere decisioni. Inoltre, l'utilizzo deregolamentato delle tecnologie descritte nella prima parte del volume, potrebbe generare effetti indesiderati anche in termini di benessere sociale e individuale derivanti dall'esposizione prolungata, il che può generare problemi di autocontrollo, polarizzazione affettiva, distorsione della realtà. Molto si è detto e molto si è scritto a tal proposito, negli ultimi anni, anche in ragione dell'uso prolungato delle tecnologie di ultima generazione durante il lungo periodo della pandemia da Covid19; tuttavia pochi sono ancora gli studi scientifici rigorosi che sappiano esprimersi in merito con chiarezza: di qui l'urgenza di studiare più a fondo la questione, date le sue possibili implicazioni.

Il volume si conclude così, con forza, con un'esortazione con cui si apre il capitolo 10, *Regolare, non spegnere*, che suggerisce che «non possiamo realisticamente ritornare a un mondo senza Intelligenza Artificiale, così dobbiamo trovare un modo di convivere in sicurezza con questa tecnologia»⁶ e la ricetta per realizzare questo importante obiettivo si fonda, secondo Cristianini, su due principi-chiave: quello della responsabilità e quello della verificabilità:

Decidere chi è responsabile per gli effetti di un sistema di IA sarà un passo cruciale: è l'operatore, il produttore o l'utente? E questo si lega al secondo fattore: la verificabilità, ovvero l'*ispezionabilità*. Come faremo a fidarci di sistemi che non possiamo ispezionare, a volte addirittura perché sono intrinsecamente costruiti in tale modo? Ogni ulteriore regolamentazione del settore dovrà stabilire fin dall'inizio che gli agenti intelligenti siano costruiti in modo tale da essere ispezionabili, e che quell'onere deve cadere sul produttore o sull'operatore. Avendo stabilito questo, sarà poi possibile discutere la loro sicurezza, equità, e tutti gli altri aspetti, che possono essere chiariti solo ispezionando l'agente.⁷

⁶ Ivi, p. 187.

⁷ Ivi, p. 196.