

Né ridere né piangere, ma capire

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

Parole chiave

IA Big data, controllo sociale, disorientamento informato, indisciplinarietà

Mariella Berra, docente del Dipartimento CPS dell'Università di Torino, si è occupata di innovazione tecnologica e digitale, nuovi modelli produttivi e di scambio socio economici e di donne e nuove tecnologie nella società 4.0. Sui temi delle trasformazioni del lavoro, della innovazione digitale ha al suo attivo fra libri e articoli in riviste italiane e straniere 150 pubblicazioni (mariella.berra@unito.it)

Dopo 80 anni di alterne vicende, le tecnologie della Intelligenza Artificiale (IA) sono uscite dai laboratori, pervadono ogni ambito della nostra vita quotidiana e sono al centro delle scelte tecnologiche di imprese e governi. Si parla con algoritmi nello svolgimento dei compiti più banali, ad esempio prenotare un biglietto del treno o una visita medica. Alexa e Siri di Amazon o Google bot assistant ci forniscono servizi, altri intervengono autonomamente nelle nostre comunicazioni come ad esempio Smart Reply di Gmail. Si utilizzano traduttori automatici, algoritmi di IA sono usati per vagliare i *curricula* di lavoro e riconoscere la veridicità di chi fa domanda di asilo e, come il recente e discusso ChatGpt di AI, l'ultima frontiera dell'IA, sono capaci di

generare testi, articoli, musiche e filmati in diverse lingue. L'IA ha avuto uno straordinario successo pratico, ma ha riaperto, come spesso è capitato con le innovazioni tecnologiche radicali, un acceso dibattito tra grandi e utopiche aspettative e allarmate preoccupazioni. È del primo marzo di quest'anno la lettera manifesto pubblicata da un gruppo di imprenditori, accademici e personalità pubbliche, tra cui Elon Musk e Yuval Noah Harari, sul sito [Futureofflife.org](https://futureofflife.org) per richiedere a governi e aziende una moratoria di sei mesi sullo sviluppo delle nuove intelligenze artificiali generative, come Chat GPT, di cui non si conoscono le conseguenze sociali e politiche.

Capirle per poter convivere è spiegato nei dieci capitoli del libro di Nello Cristianini, pubblicato anche in versione inglese sempre nel 2023 da Taylor e Ford e scritto, dunque, dopo la apparizione di Chat GTP. Nell'ambito del ricco e multidisciplinare dibattito sulla IA, il lavoro di Cristianini ha il merito di esprimere il punto di vista interno di uno studioso universitario che non solo ha una esperienza teorica e pratica trentennale nella costruzione degli algoritmi alla base della IA, ma è informato sulle sue potenzialità e anche molto attento alle crescenti e irrisolte criticità sul piano sociale, giuridico e etico. L'autore è infatti professore di *machine learning* da 25 anni all'Università di Bath, ha contribuito a generare alcuni degli algoritmi di quella stessa IA di ultima generazione di cui parla nel libro ed è anche attivo collaboratore con le istituzioni dell'Unione Europea per predisporre la normativa per la prossima e auspicabile futura *governance* della IA. *Scorciatoia* in realtà è la sintesi di tre scorciatoie che rappresentano altrettante svolte fondamentali nella evoluzione storica delle tecnologie della IA, dalle origini degli anni '60 del secolo scorso, alla fase dei sistemi esperti degli anni '80, alla diffusione e applicazione delle reti neurali e del *machine learning* di questi ultimi decenni. Si è perfezionato così quel cambiamento che rappresenta un nuovo paradigma rispetto ai progetti originali di IA, una vera e propria rivoluzione in senso kuhniiano.

La prima scorciatoia si può sintetizzare nella relazione tra teoria e dati. La IA delle origini aveva raggiunto risultati particolarmente interessanti nei campi della logica e della dimostrazione automatica

di teoremi, però non era riuscita a individuare modalità di intervento efficace nei campi del cosiddetto senso comune, cioè quell'insieme di conoscenze sterminate che intervengono nel facilitare la vita quotidiana. Si è abbandonato quel modello di carattere deduttivo, che dopo cinquant'anni di lavoro non aveva prodotto grandi risultati e si è sviluppato, quindi, un approccio metodologico statistico che, abbandonata la pretesa di una conoscenza teorica a priori, puntasse, invece, ad addestrare le macchine ad apprendere dalla esperienza, rappresentata dai dati. Di conseguenza, se contano solo le previsioni e non la comprensione, non è più necessario capire il fenomeno che si sta riproducendo nella macchina. Tanti dati favoriscono l'apprendimento automatico da parte di macchine intelligenti in qualunque lingua e con qualsiasi collegamento. Le macchine sono state, infatti, addestrate a fare previsioni statistiche e sono in grado di completare su grandi matrici le variabili mancanti. La logica, spiega l'autore, del *machine learning* è paragonata a quegli esercizi, i cosiddetti *cloze text*, utilizzati per l'apprendimento delle lingue. L'IA del *machine learning* impara quindi a scoprire processando miliardi di dati, una serie di correlazioni che nessuna mente umana, né singola né collettiva, riuscirebbe a cogliere.

La seconda scorciatoia riguarda le possibilità e il metodo della raccolta ed elaborazione dei dati. Produrre dati *ex novo* con lavori di programmazione sarebbe stato un processo molto costoso, in termini di mezzi e di tempi. La diffusione di Internet su scala planetaria e del WWW ha permesso di mettere a frutto un nuovo metodo economico di estrapolazione dei dati, quello di riciclare le informazioni esistenti prodotte da altri e diffuse su ampia scala su quella miniera a cielo aperto della rete. Con un lavoro di *data-mining* questi possono essere raccolti, elaborati, archiviati e analizzati in tempi estremamente rapidi.

Tante piccole scelte producono grandi scelte: osservare l'utente e individuarne le preferenze costituisce la terza scorciatoia. I dati già esistenti diffondono implicitamente informazioni sugli interessi degli utenti: ad esempio, il tempo con cui ci si sofferma ad analizzare una notizia attinta da Google, oppure quante volte si clicca un sito. Queste informazioni generano dei meta-dati che offrono indicazioni sul tasso

di interesse. Diventa così possibile costruire un gemello digitale che permette di fare predizioni e comunicazioni personalizzate.

Le tre scorciatoie hanno consentito la crescita enorme di prodotti collegati ad uno sviluppo *data driven* e segnato il successo di grandi multinazionali dell'IT, come Amazon e Google. Di fronte alla esplosione dei big data e ai successi della nuova IA, già in un articolo del 2008, Chris Anderson, direttore della rivista Wired, la cosiddetta Bibbia di Internet, ipotizzava, forse in modo troppo entusiastico, l'ipotesi di costruire teorie esclusivamente *data driven* (cfr. Andersen 2008). Secondo questo approccio, peraltro, all'epoca molto discusso, i dati sarebbero di per sé sufficienti a generare la loro teoria senza bisogno di uno schema teorico che li sappia ordinare e interpretare. La posizione di Cristianini è meno radicale: infatti, l'ipotesi di una 'fine della teoria' rischia di limitare le possibilità di esaminare e sottoporre a critica e revisione le applicazioni e l'uso degli algoritmi. Piuttosto, la logica dell'IA non è di cercare cause, ma riconoscere correlazioni anche tra variabili lontane. L'aver sostituito modelli teorici obsoleti e non funzionali con regolarità statistiche e riconoscimento di *pattern* appresi dai dati ha reso possibile le principali storie di successo sullo sviluppo e uso della IA. La macchina impara cosa è vero osservando dati generati da processi che non comprendiamo pienamente, come ad esempio le interazioni fra molte persone e gruppi di persone, apprende gli obiettivi dell'utente osservandone il comportamento o durante una interazione diretta. In questo modo di operare della IA, più che al sapere che guarda agli antecedenti causali passati, si guarda alle conseguenze future e al fatto che esse siano le più statisticamente attendibili, in modo da favorire il processo decisionale in ogni ambito di applicazione. Con la digitalizzazione e con la crescita del *machine learning* e del *deep learning*, si assiste a uno spostamento del paradigma epistemologico, da uno basato sulla spiegazione a uno basato sulla sola previsione. Le macchine sono diventate molto affidabili nel predire statisticamente le occorrenze di determinate parole e della loro sequenza sintattica nei vari contesti e i margini di errori, pur sempre presenti, sono sempre più ridotti ai minimi termini. Il *Deep Learning*

e il *Machine Learning*, sfruttando un sistema complesso di reti neurali che simulano il comportamento cellulare del nostro cervello, sono in grado di apprendere autonomamente e di imparare dai propri errori. Con ciò non significa che pensano in modo umano, ma, per l'autore, il valore della IA consiste nel comportarsi in modo non solo efficiente, ma efficace nel proprio contesto e negli ambienti nuovi.

La statistica è il linguaggio della IA e del suo percorso continuo di ottimizzazione, la quantità e qualità dei dati il suo alimento. L'utilizzo esteso dei dati forniti dagli utenti inizia nel 2001 quando, dopo la crisi delle dot.com Google, brillante e tecnologicamente ben attrezzata *start up*, per fronteggiare la pesante crisi economica ha fatto dell'offerta pubblicitaria, basata sulle preferenze degli utenti rispetto alla fornitura di servizi e informazioni, il suo *core business*. Come racconta la sociologa Soshana Zuboff in un interessante e documentato libro, che spiega in modo approfondito la trasformazione del sistema capitalistico nell'era digitale, il fondatore di Google Larry Page aveva intuito che esperienze, pensieri, emozioni e sentimenti, la vita cosiddetta *off line*, avrebbero potuto essere la nuova materia prima, estraibile, a differenza del petrolio, a costi bassi o tendenti a zero, un territorio vergine da sfruttare (cfr. Zuboff 2019). Proprio il *web* e i *social network* sono tra i principali luoghi dove si incontrano quegli agenti intelligenti, fonti di forti guadagni per il settore della IA e essenziali per il modello di *business* di molte imprese del Web che ci osservano, influenzano e imparano da noi. I dati di cui si avvale la IA – in modo differente a seconda delle loro diverse tipologie utilizzate: grezzi, ibridi, sintetici, supervisionati – sono viziati spesso dai pregiudizi, i cosiddetti *bias*. Non sono solo i *bias*, ad esempio di genere e razza determinati dalle caratteristiche dei programmatori, prevalentemente maschi e bianchi, a creare pregiudizi, ma sono i dati su cui si addestrano le macchine, attinti da un processo di continua, permanente digitalizzazione e registrazione delle nostre esperienze di vita quando lasciamo tracce di noi nel web, delle nostre opinioni, e dei contenuti diffusi veri e falsi nelle varie piattaforme digitali. Sono questi i dati su cui prima si addestrano e poi fanno predizioni le IA di ultima generazione, dati che sono poi modellizzati per

le diverse esigenze. Questo pone un problema di qualità dei dati che alimentano la IA e anche dell'uso della IA. Di conseguenza, le previsioni possono essere ingiuste o sbagliate e condurre a decisioni ingiuste o sbagliate. Una preoccupazione, quella di capire in che modo vengono raccolte e utilizzate le informazioni e cosa viene scelto dall'algoritmo di intelligenza artificiale per produrre i propri contenuti, che ha spinto l'Unione Europea nell'aprile 2023 a creare a Siviglia l'*European Centre for Algorithmic*. Il Centro, cui partecipano scienziati, esperti membri dell'industria e della società civile, intende contribuire a definire la strategia comunitaria sul digitale per capire meglio il funzionamento degli algoritmi, analizzare l'etica e valutare i rischi, proponendo nuovi approcci trasparenti e *best practice*. In linea con i progetti e le proposte dell'Unione Europea, l'autore mette in evidenza come l'obiettivo del suo lavoro consista nel non fare male alla gente: "noi tecnici abbiamo le responsabilità di dire le cose come stanno" nei limiti, comunque, delle possibilità e conoscenze acquisite.

E sottolinea come quando si è completata la prima bozza della IA, essa sia già stata integrata nelle nostre infrastrutture pubbliche e private, si sia insinuata tra di noi al punto che non risulta facile staccare la spina o rinunciare a essa. Di qui le domande sul perché devolvere ad algoritmi che non controlliamo appieno tante decisioni e che fare e anche l'invito a capire come ci influenzeranno e come goderne i benefici e limitarne i rischi.

Se la informatica e la statistica sono stati i principali artefici del salto di paradigma costituito dalle tre scorciatoie, per imparare a conoscere gli effetti della IA sulla nostra vita quotidiana, per comprendere i principi base del suo comportamento e regolarne lo sviluppo futuro, diventa impellente una collaborazione estesa tra la cultura scientifico-tecnica e quella umanistica. Cristianini, con un occhio attento al presente, affronta in modo brillante, da un punto di vista storico, l'evoluzione della IA, un percorso necessario per fare crescere la consapevolezza delle nuove sfide che sono di fronte a noi e alle generazioni future. Conoscere il percorso che ha portato a questa ultima fase e le modalità di funzionamento degli algoritmi è un obiettivo importante

per tutti gli studiosi di diverse discipline, per collaborare insieme e fornire risultati socialmente e politicamente efficaci. Con un'attenzione ai concetti tecnico-scientifici alla base della tecnologia e alle implicazioni più ampie per la società e con riferimento a indagini empiriche e contributi teorici, soprattutto nell'ambito del *marketing*, della psicologia cognitiva e dell'economia comportamentale, ripercorre i passaggi che hanno portato allo sviluppo di questa tecnologia e propone una descrizione attenta dei possibili effetti indesiderati. Come possiamo assicurarci che le macchine non violino norme sociali eseguendo alla lettera le nostre richieste, o prendano scorciatoie sbagliate, un timore, peraltro, già paventato dal fondatore della cibernetica Norbert Wiener?

In particolare, i sistemi di raccomandazione, che indirizzano l'utente su informazioni sintonizzate sul suo profilo, sono indicati come i primi agenti intelligenti applicati ai diversi ambiti della vita sociale e come la più importante espressione della IA, della sua azione e dei suoi condizionamenti. Come d'altra parte dimostra molta letteratura sociologica italiana e internazionale, le diverse piattaforme digitali imparano a conoscere meglio ogni singolo utente, le sue preferenze e a proporre contenuti in linea con i suoi gusti. Con raffinate tecniche psicometriche, profilano gli utenti e ne predicono i comportamenti. Si tratta di una grande opportunità per i settori privati e pubblici, che valutano i rischi individuali come, ad esempio, la concessione di prestiti, le ammissioni scolastiche, il reclutamento o le possibilità di recidiva. Effetto importante di questi sistemi di raccomandazione è quello di chiudere il soggetto in una *echo chamber*, una camera virtuale dove idee, credenze e opinioni vengono amplificate all'interno di un contesto chiuso e omogeneo, nel quale il confronto con gli altri non è che l'eco della propria voce. Il pericolo è una polarizzazione dei gruppi con conseguenze di frammentazione sociale, riduzione della libertà di scelta e crescita della disinformazione nell'opinione pubblica, come ha ben messo in evidenza l'economista comportamentale Cass Sunstein (cfr. Sunstein 2017). Anche il *microtargeting*, quella forma di pubblicità *online* profilata che analizza i dati personali, contribuisce alla creazione di camere dell'eco nelle quali gli utenti vengono esposti a un unico tipo

di informazioni, che indirizzano e spesso distorcono le loro opinioni. Caso noto di pericoloso *marketing* politico, che peraltro si pensa di vietare in Europa, è l'uso fattone da Facebook, adoperando tecniche di apprendimento algoritmico induttivo, nelle presidenziali statunitensi del 2016. La manipolazione *online* rappresenta una seria minaccia per la società ed è anche un sintomo dell'opacità e della mancanza di responsabilità dell'ecosistema digitale. Indubbiamente, la diffusione di algoritmi sempre più potenti e sempre più complessi costituisce un pericolo per la democrazia. Ad esempio, con l'ultimo e discusso Chat GTP, grazie ai miliardi di informazioni elaborate, si aumentano di molto le persone che ricevono le stesse risposte per le stesse domande su questioni complesse e sensibili. Quale spazio verrà riservato al pensiero critico? Un altro rischio connesso alla diffusione di agenti intelligenti e artificiali sempre più complessi è quello di rendere disponibili quantità sempre maggiore di informazioni private e riservate. Quella libertà e autonomia, collegata al superamento di sistemi di intermediazione, all'abbattimento di barriere e cancelli offertaci dalla rete, rischia all'opposto di ridursi molto con algoritmi capaci di manipolare, di modificare, di predire, di gestirci. Il dilemma che suggerisce la lettura del testo è dunque se saremo noi i padroni dell'informazione o saremo i suoi schiavi, passando da un progressivo stato di "disorientamento informato", come lo definiva Castells, il grande studioso della società informazionale, a uno stato di servitù volontaria, come scriveva già nel 1600 il politologo e giurista francese Etienne de la Boétie.

Molti sono i problemi sociali ed etici che ne conseguono e toccano temi quali la giustizia, l'eguaglianza, la *privacy*, che richiedono una gestione collettiva e un intervento della politica e delle politiche degli Stati. Un problema etico molto discusso è la gestione della *privacy* che si dibatte nel vecchio conflitto fra libertà e sicurezza, un binomio che andrebbe specificato in libertà per chi e quale sicurezza tra gli opposti valori della tutela della *privacy* da un lato e del controllo e della sicurezza sociale dall'altro. Correttamente, sostiene Cristianini, considerata anche la competizione geopolitico-tecnica planetaria nel campo della IA, è irrealistico pensare che ci si possa liberare tornando

indietro, rinunciando cioè a questa tecnologia che è ormai profondamente integrata nelle infrastrutture della comunicazione e della riproduzione sociale, o anche imponendo forti restrizioni da parte di organi istituzionali. Inoltre, queste tecnologie sono utili nel semplificare la nostra vita e nell'offrire servizi qualificati ed efficienti in ambiti quali ad esempio la sanità, l'istruzione, le telecomunicazioni, la logistica, la ricerca, l'industria. Di conseguenza, richiamando l'obiettivo del libro, per conviverci occorre conoscere i principi alla base del loro comportamento e i loro effetti, e questo richiede di costruire un rapporto di lavoro cooperativo fra scienziati sociali e tecnici del *software* in quell'ambito intermedio tra la creazione del prototipo e la sua applicazione, una collaborazione estesa su scala internazionale, come è avvenuto per recenti scoperte scientifiche. Le modalità della nostra relazione con le macchine e di come si vada delineando il rapporto fra tecnologie e società vanno gestite in modo collettivo e planetario. Diventa urgente, ricorda Cristianini, elaborare anche regole giuridiche sia a monte nella produzione, sia a valle sull'uso di tali macchine di IA. E bisognerà fare questo guardando non solo ai principi etici astratti, ma anche ai dettagli tecnici e alle conseguenze sociali.

Tre sono oggi a grandi linee i modelli che si confrontano su scala mondiale. Quello del grande fratello riguarda Stati autoritari che usano l'IA come sistema di controllo pubblico, un sistema di divieti normativi nell'uso e nella diffusione delle tecnologie. Quello liberale degli Stati Uniti, basato su una logica privatistica di libertà della innovazione e della competizione, un modello che si potrebbe definire delle suadenti sorelle, dove le grandi imprese, a cui si devono anche i principali progressi della IA, controllano un mercato parzialmente assuefatto o, per dirla ancora con Zuboff, inconsapevole e obbediente fornitore di un surplus di dati comportamentali. E, infine, il modello dell'Europa che, tra numerose difficoltà, dovute anche alla non omogeneità di indirizzi dei diversi Stati che ne fanno parte, a partire dal GDPR (Regolamento generale sulla protezione dei dati) del 2016, alla tutela della privacy e al piano coordinato di una regolazione etica per gli Stati membri nell'uso della IA, propone un esempio virtuoso. Con

l'ultima IA, si è infatti imbattuti nel problema della cosiddetta XAI (not-explainable AI). Le difficoltà, e spesso l'impossibilità, di chiedere spiegazioni sulla logica della macchina, penetrando la 'scatola nera' che regola il funzionamento della IA, pone nuove e impellenti riflessioni (cfr. Foryciarz, Leufer, Szymielewicz 2020).

Si tratta di un problema, a mio avviso, di fiducia tecnica e sociale sul modo in cui queste macchine sono state costruite e sul conseguente impatto sociale di una tecnologia sempre più complessa e anche sui gradi di autonomia che può raggiungere una IA, tale che non possiamo completamente spiegare e controllare. L'incontro fra le due culture, scientifica e umanistica, aiuterebbe non solo a valutare e a interpretare il significato della IA, ma a elaborare proposte su come governare le conseguenze difficili da prevedere di queste nuove macchine. Si tratta, infatti, di un compito che richiede la massima informazione e formazione anche dell'opinione pubblica, che deve diventare soggetto attivo per coltivare il pensiero critico e indipendente. Una democrazia ha bisogno di anticonformisti consapevolmente formati. Altrimenti, c'è il rischio di un *digital divide* nuovo e allargato non solo fra chi avrà accesso alle tecnologie, ma tra chi saprà e potrà interagire proficuamente con le varie forme di AI e chi invece no.

Riferimenti bibliografici

Anderson, C.

2008, *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, Wired June (<https://www.wired.com/2008/06/pb-theory/>)

Foryciarz, A., Leufer, D., Szymielewicz, K. 2020, *Black-Boxed Politics: Opacity is a Choice in AI Systems*, 2020, (<https://hiw.kuleuven.be/en/news-events/news/in-the-spotlight/agata-foryciarz-daniel-leufer-katarzyna-szymielewicz-black-boxed-politics-opacity-is-a-choice-in-ai-systems>)

Sunstein, C. R.

2017, *Republic: divided democracy in the age of social media*, Princeton University Press, Princeton, 2017.

Zuboff, S.

2019, *The Age of Surveillance Capitalism. The fight for a human future and the new frontier of power*, Public Affairs, New York.