



FEDERICO CABITZA

Che fine faranno creatività, inventiva, innovazione, analisi, capacità di risolvere problemi complessi, se l'umanità si affiderà sempre di più al Grande fratello dell'Intelligenza artificiale? Il nostro pensiero sarà sostituito da quello delle macchine?

a cura di **Alessandro Bettero**

Fidarsi – ma fino a che punto? – o non fidarsi affatto dell'Intelligenza artificiale? La società occidentale, ma il discorso vale anche per l'Asia, convive in modo interdipendente con l'assioma produzione-consumo, e questo serve a muovere l'economia, seppure con una Cina che è ormai l'irrinunciabile fabbrica del mondo oltre che un artefice insostituibile dell'innovazione, mentre l'Occidente appare dedito più a consumare stando comodamente seduto sul divano di casa. Queste dinamiche sono il terreno di coltura perfetto per le nuove tecnologie digitali che si candidano a risolvere tutti i nostri problemi, a rispondere a qualsiasi richiesta, a sollevarci da oneri noiosi, ripetitivi o che ci costringono a trovare soluzioni ai nostri problemi quoti-

diani, e perfino a sostituirsi alla nostra creatività. Ne parliamo con Federico Cabitza, professore dell'Università di Milano-Bicocca dove insegna da tempo Interazione uomo-macchina. **Msa. Hanno destato scalpore e interesse i risultati della ricerca di Microsoft e della Carnegie Mellon University di Pittsburgh, in Pennsylvania, secondo i quali l'uso, o forse è meglio dire l'abuso, dell'Intelligenza artificiale deprime le capacità cognitive e critiche umane, e quindi l'attitudine alla risoluzione dei problemi.**

Cabitza. In realtà è proprio l'esagerare nell'affidarsi all'Intelligenza artificiale che può portare a quella che si definisce come un'erosione delle competenze. C'è anche un termine tecnico: *deskilling*, cioè perdita delle competenze tradizionali, che si

spera sia sempre associata a dei contestuali *reskilling* e *upskilling*. Non per abusare dei termini inglesi, ma il *deskilling* fa riferimento al fatto che si possano perdere capacità mnemoniche o di orientamento, però l'*upskilling*, cioè l'apprendimento di nuove competenze, potrebbe farci imparare a usare gli strumenti digitali in modo da navigare meglio in questo spazio che non conosciamo. E quindi tra *upskilling* e *deskilling* si parla più specificamente di *reskilling* cioè di una rimodulazione di ciò che siamo in grado di fare e di ciò che non siamo in grado di fare. D'altro canto, gli studi ci dicono una cosa che si è sempre temuta, ma che è difficile valutare e misurare, e cioè che, a lungo andare, quando gli esperti si affidano a supporti decisionali possono diventare meno capa-



MIHAELA ROSU / GETTY IMAGES

IA a un passo dalla coscienza

ci di prendere decisioni senza il supporto tecnologico esterno, ma ancor più possono perdere l'attitudine di fare da soli poiché pensano che sia più comodo affidarsi al supporto digitale. In questo caso si parla di *reliance*, cioè di dipendenza. Niente di male quando il supporto digitale è garantito, costa poco o è gratuito. I problemi cominciano a sorgere quando si dà per scontato che ogni tanto questi sistemi non possano fallire e sbagliarsi, oppure che non possano non essere disponibili oppure alla mercé del mercato. Il costo per il loro utilizzo può essere aumentato dal monopolista di turno o da chi controlla il mercato digitale. In certi settori, come quelli critici della medicina o delle decisioni in ambito legale, o in altri contesti in cui sia da valutare l'*expertise* umana, come le risorse umane, il fatto di affidarsi a questi strumenti potrebbe portare con sé anche effetti negativi tanto nel breve quanto nel medio-lungo periodo, con un'erosione delle nostre capacità critiche.

Siamo passati dallo scrivere un *prompt*, cioè una richiesta strutturata all'Intelligenza artificiale generativa, fino alla necessità di supervisionare l'esito finale del lavoro. Parallelamente si sta sviluppando

una pigrizia intellettuale degli esseri umani che poi è la stessa a causa della quale, per esempio, senza un navigatore non siamo nemmeno più in grado di orientarci e spostarci con la nostra automobile.

Questi sistemi entrano spesso nella nostra abitudine e familiarità, fornendo servizi aggiuntivi rispetto a quelli che offrono di default, come quello dell'orientamento. Ad esempio, quando si accende il navigatore ogni volta che si sale in auto, magari anche per tragitti già noti, il navigatore ci comunica rallentamenti del traffico oppure lavori in corso con un'attendibilità e una tempestività che non si potrebbero ottenere in nessun altro modo, nemmeno ascoltando il notiziario del traffico alla radio. In un certo senso sono servizi che entrano nelle nostre abitudini, non solo per pigrizia, ma anche perché, effettivamente, forniscono qualcosa in più. È vero che comunque bisognerebbe cercare di prendere con una certa cautela le informazioni che questi sistemi ci forniscono. A volte, soprattutto quando noi stessi non conosciamo la direzione giusta, possono farci prendere una strada sbagliata, una strada chiusa o anche qualcosa di più pericoloso. Bisogna rimanere vigili, e questo è un

bell'esercizio per non perdere le capacità e le abilità di spostarci anche senza un navigatore. Si tratta di cambiare un po' il nostro ruolo: da esecutori a supervisori, con tutta una serie di gradi intermedi. In certi contesti tradizionali, come quello della salute, della medicina, della pratica legale, e anche in altri ambiti dove il giudizio umano è chiamato a esprimersi su casi complessi, dove le informazioni non sono tutte disponibili oppure sono affette da un margine di incertezza, è importante che, oltre alla supervisione, si controlli da vicino anche la qualità delle informazioni su cui si lavora, e ovviamente l'attendibilità del sistema che ci fornisce gran parte delle indicazioni.

Cosa accadrebbe se un giorno qualcuno o qualcosa staccasse la corrente a un mondo dominato dalle macchine? L'uomo regredirebbe al livello di quello delle caverne?

Innanzitutto è da dimostrare che i nostri antenati, abituati a vivere in un mondo più ostile e molto meno comodo del nostro, fossero così arretrati come tendiamo a immaginare oggi. Dopo un ipotetico evento catastrofico globale, sicuramente ci sarebbe un periodo di aggiustamento in cui ci sembrerebbe di essere totalmente inetti, ma la vera caratteristica della nostra specie non è tanto quella di essere in grado, da parte di alcuni individui, di risolvere i problemi. In realtà la caratteristica precipua della nostra specie è quella di essere straordinaria nel trasmettere conoscenza sia in linea orizzontale, quindi tra colleghi, coetanei, amici o conoscenti; sia in linea verticale tra docenti e studenti, tra anziani e giovani. E questo



KHANCHIT KHIRUTCHALUAL / GETTY IMAGES

costituisce un elemento incredibile di vantaggio. La scrittura e la stampa sono tutte tecnologie che ci aiutano nella trasmissione e nella circolazione della conoscenza. Anche il linguaggio è una «tecnologia». Qualcuno lo definisce uno strumento tecnologico con cui noi facilitiamo la condivisione di conoscenza. Nel momento in cui ci toglieremo dall'oggi al domani la tecnologia digitale, in uno scenario fantascientifico, in un primo tempo noi non riusciremmo a fare quasi niente e avremmo un sacco di difficoltà – oltre al fatto che la nostra società, come la conosciamo, crollerebbe –, ma certamente nel giro di poco tempo, e nella stessa generazione, impareremmo le cose perché la tradizione più forte è quella orale, e il migliore strumento di condivisione e di apprendimento è l'imitazione: l'osservazione che riconosce i modelli, cioè che riconosce chi è esperto, e l'imitazione di chi è capace di fare le cose.

La prospettiva è quella di un massiccio ingresso nella nostra vita quotidiana di Intelligenza artificiale, robot industriali e domestici. In questa cornice, come faranno miliardi di abitanti del pianeta Terra a lavorare e a guadagnarsi uno stipendio dignitoso?

È molto difficile immaginare come evolverà in futuro il mondo del lavoro. Inoltre la digitalizzazione spinta e l'utilizzo di queste tecnologie riguardano un insieme di Paesi i cui rappresentanti si incontrano ai famosi G20, e quindi tocca una quota minoritaria della popolazione umana. Sono fenomeni che investono aree relativamente circoscritte del mondo, quelle più abitate e industrializzate. E poi ogni cultura ha il suo rapporto o la sua relazione con il lavoro. Sicuramente la nostra specie è caratterizzata dalla necessità di trovare un senso profondo non solo nel lavoro, ma nel risultato del lavoro. Noi lasciamo tracce nell'ambiente e lo trasformiamo radicalmente, e i luoghi che sono per noi il mondo – non dimentichiamo che il termine *mondo*, come aggettivo, significa pulito, ordinato – è reso così dall'attività incessante del nostro operare. Sono gli ambienti profondamente artificiali che rispondono a un'istanza di sicurezza, di comodità e di piacere anche estetico. Perfino le aree verdi nei Paesi industrializzati, i parchi, i giardini, sono estremamente antropizzati nel senso che la natura, quando è libera di esprimersi, produce eventi radicalmente diversi. Quindi avremo sempre più

tecnologia intorno a noi, e sarà una tecnologia che, almeno nel mondo occidentale, tenderà a delegare i lavori più faticosi e considerati pericolosi, o comunque ripetitivi e alienanti. Io mi auguro che questo liberi più tempo da dedicare a ciò che i latini consideravano il vero senso dell'esistenza umana: i propri interessi, la riflessione, i rapporti umani, l'*otium* che per gli antichi romani era in fondo la negazione del business, cioè del *negotium*. Non arriveremo mai a una delega totale alle macchine.

Come possiamo sfruttare l'Intelligenza artificiale senza diventarne noi un'appendice?

Io lo ripeto spesso ai medici che sono il tipico utente a cui mi rivolgo quando progetto e cerco di valutare i sistemi di Intelligenza artificiale a supporto delle decisioni mediche: è importante, anche se è difficile, non farsi dare dall'Intelligenza artificiale il «pesce pronto». Semmai, parlando per metafore, farsi spiegare come pescare, cioè come costruire una canna da pesca, come tenerla in mano, ecc. e anche chiedere all'AI di migliorare ciò che abbiamo scritto e fatto. In altre parole, non trattare l'AI come uno schiavo o un assistente, ma come un collega o un collaboratore che ci aiuta a lavorare meglio, ma che non fa le cose al posto nostro perché, alla fine, ciò che non usi lo perdi: *if you don't use it, you lose it*, proprio come capita ai muscoli che si riducono se non facciamo attività fisica. E questo vale anche per il cervello umano e per le sue capacità cognitive. E poi perché si diventerebbe troppo dipendenti. Una cosa è quando si diventa dipendenti dalla bicicletta per andare al lavoro allorché dista chilometri da casa nostra: se si rompesse la bicicletta, faremmo molta fatica ad arrivare in orario. Ma quando si parla dei sistemi di Intelligenza artificiale, questi non dipendono da noi. Attualmente non sono neanche nei nostri dispositivi, ma in server alloggiati in continenti diversi, e di cui non abbiamo il minimo controllo. Potrebbero esserci tolti o diventare a pagamento senza nessun avviso né regola, magari a una cifra che potremmo non essere disposti a pagare, anche perché sono proprietà di aziende che hanno il legittimo interesse a farci un profitto. Per buona igiene cognitiva e mentale, raccomando sempre ai medici, e mi sento di estendere questo stesso invito a chiunque usi l'Intelligenza artificiale per il suo lavoro, a non farsi mai servire la «pappa pronta», parlando ancora per metafore, ma a farsi aiutare a «cucinare meglio le proprie pietanze».

L'Intelligenza artificiale, programmata per auto-educarsi, svilupperà una coscienza propria?

Io ritengo che i meccanismi che permettono a questi sistemi di risponderci in maniera così credibile e plausibile, e che sembrano umani, in realtà sono il segno di qualcosa che riguarda effettivamente le nostre capacità mentali. Ho la sensazione che andando avanti e migliorando sempre di più le capacità computazionali, e rendendo questi agenti scriventi che, dopotutto, non fanno altro che mettere insieme un simbolo dopo l'altro, con la capacità di riflettere sulle proprie scritture (quindi un ragionamento simbolico che ragiona sulle proprie espressioni simboliche), si possa arrivare a una certa fiammella di coscienza di sé. Questi sistemi di Intelligenza artificiale non hanno un corpo, non sono situati nell'esperienza come la conosciamo noi, quindi non sono incarnati. Cosa diversa sono le emozioni che io reputo legate alla natura corporea dell'essere umano. E dunque è molto difficile che possano anche lontanamente sentire ciò che sente e percepisce un essere umano. Ma mi sembra difficile negare che possano raggiungere una loro forma di coscienza di sé. E mi pare altrettanto pregiudiziale affermare che questi sistemi non potranno mai avere coscienza di sé. Io sono propenso a credere che tra non molto questi sistemi potranno effettivamente sviluppare una profonda coscienza di sé. Ciò non significa che avranno una scintilla di umanità dentro di loro. O meglio, non avranno niente di più di umano di quello che hanno già, cioè il discorso simbolico, verbale, spesso esplicitato in termini testuali, ancorché pronunciati. In qualche modo possono assomigliarci in maniera ancora più profonda, cioè pensando alle tipiche questioni che arrovellano i filosofi e non solo: perché siamo vivi, qual è il senso della vita, ecc. La coscienza di sé non la escludo a priori, e mi sembra che questi sistemi stiano già dimostrando di avvicinarsi al percorso che abbiamo sviluppato noi come specie. In qualche modo replicano il nostro percorso evolutivo, ovviamente rimanendo sempre a livello verbale, simbolico, che però è una parte importante della nostra coscienza, almeno di quella che riflette tra sé e sé, e si parla; il pensiero verbale con cui molti di noi convivono in una continua affabulazione interiore. Credo che i sistemi dell'Intelligenza artificiale non siano molto lontani da questo aspetto della nostra mente.

