

IA vs mondo del lavoro

memoria sintetica

Alessio Mazzolotti, PhD – CEO Ice Cubes e Fishbone Creek, partner Bad Idea Group

Aaron Brancotti, MS – CTO Ice Cubes e Fishbone Creek, partner Bad Idea Group

Gli scriventi si occupano della relazione tra uomo e tecnologia da decenni¹, operando in contesti imprenditoriali dedicati all'ideazione, progettazione, prototipazione e produzione di contenuti digitali evoluti (simulatori interattivi, applicativi in Extended Reality, sistemi per l'accessibilità fisica e cognitiva, tool immersivi per la didattica...), collaborando con atenei e centri di ricerca, curando attività di alfabetizzazione tecnologica, attraverso pubblicazioni e eventi.

La prospettiva di lavoro dei team che gli scriventi si onorano di coordinare abbraccia istanze di natura socio-antropologica e filosofica in un combinato disposto con questioni più squisitamente tecniche e tecnologiche, in un approccio contech nel quale la progettazione raramente si discosta dal cosiddetto Human Centered Design².

In questa cornice lo studio dell'Intelligenza Artificiale è da tempo uno dei temi di maggiore interesse e ricerca, come dimostrano – solo per rimanere nella storia recente – i primi esperimenti condotti con i sistemi OpenAI sin dal “lontano” 2021 quando una delle società degli scriventi è stata, fra le poche in Italia, ammessa alla fase di test (beta chiusa) di quello che in soli due anni avrebbe poi originato il ben noto ChatGPT.

Già in quella fase embrionale il sistema creato da OpenAI appariva molto promettente.

Il suo utilizzo richiedeva competenze nello sviluppo di codice non alla portata di tutti ma i test esperiti sulla “capacità di analogia” (una delle caratteristiche della mente umana più difficili da replicare digitalmente, ovvero la capacità di produrre e seguire processi di significazione su livelli differenti non limitandosi alla mera identità) destavano stupore pur non inducendo a immaginare la portata dello sviluppo esponenziale dei mesi a seguire.

Al contempo, allora come oggi, anche dinnanzi ai più recenti sviluppi, di Open AI così come di Google, gli scriventi continuano ad avere un'idea dello stato dell'arte dell'Intelligenza Artificiale che per molti versi si discosta dall'opinione comunemente condivisa.

¹ <https://www.sciencedirect.com/science/article/abs/pii/S001048259400040W>

² cfr. Brancotti A., Mazzolotti A. (2023), #METAVERSE ARCHITECT, FrancoAngeli, Milano.

Un esempio: il concetto comune di "IA generativa" intesa come "IA che crea cose nuove" è, sempre a parere degli scriventi, decisamente opinabile poiché di fatto le IA basate su tecnologie analoghe a GPT (Reti Neurali/Deep Learning) sono sistemi che costruiscono sempre e soltanto risposte plausibili. Quando le IA non hanno una risposta, la inventano di sana pianta, come uno studente colto impreparato all'interrogazione. Se, e soltanto se, creare risposte plausibili, ma non necessariamente "vere", significa generare, allora la definizione di IA generativa può essere considerata accettabile a dispetto però dell'accezione comunemente intesa che impone un'identità fra "generazione" e "creatività".

Questa chiave di lettura spesso sfugge ai più: comunemente le IA vengono descritte come enormi database o motori di ricerca o, peggio, come imponenti matrici di fattori di probabilità che concatenano le parole in sequenze ben definite (sostanzialmente il T9 di Nokia o l'odiosa definizione di "pappagallo stocastico"³, che dimostra quanto il funzionamento di questi sistemi non sia stato compreso), ma tutto ciò è ben lungi dall'essere vero.

Nelle IA di oggi, gli scriventi, se da un lato non ravvisano una capacità generativa incondizionata, dall'altro non possono però non riconoscere la presenza di un "qualcosa in più" che rimanda immediatamente a una riflessione tutt'altro che banale sui comportamenti emergenti e sulla teoria della complessità, invisa a chi vorrebbe spiegazioni analitiche per tutto lo scibile scientifico e che non accetta che esistano fenomeni che non possano essere spiegati in maniera riduzionistica. Proprio in tale senso e in vista di una corretta lettura del fenomeno IA sarebbe importante accettare la complessità e trasformarla in una opportunità di crescita, soprattutto per un establishment scientifico/ingegneristico fermo a Laplace e che persevera nel credere che il Problema dei Tre Corpi sia "un problema di precisione" e non una dimostrazione matematica di un limite ben preciso a quanto possiamo calcolare⁴.

Il processo di apprendimento delle IA massive come GPT porta alla creazione di quello che tecnicamente viene chiamato uno "spazio latente", una struttura di dati multidimensionale e stratificata

³ se le IA fossero semplici reti markoviane, non potrebbero scrivere una storia su una zuccina spaziale che parte alla ricerca del vello d'oro (test che abbiamo fatto addestrando una IA con gli *Esercizi di Fantasia* di Gianni Rodari), perché nessuno l'ha mai scritta prima. Quindi da dove verrebbe l'"idea"? La somma delle probabilità che legano la sequenza di tutti quei simboli è zero. È quindi evidente che nei processi sottostanti a una IA vi sia "qualcosa di più" di una semplice macchina stocastica.

⁴ Si pensi a Gödel e al Teorema di Incompletezza.

ove concetti simili sono tra loro "vicini" in senso topologico⁵. Se è vero che questo spazio viene inizialmente creato a partire da collegamenti puramente statistici –quelli della semplice macchina stocastica per intenderci–, non bisogna trascurare il fatto che alla fine del processo si ottengano invece delle vere e proprie menti "aliene" che, operando collegamenti "spontanei" fra i dati, diventano in grado di manifestare comportamenti e ragionamenti molto vicini a quelli umani.

Ovviamente il comportamento di questi sistemi dipende in prima istanza dai dati con i quali essi vengono addestrati.

Sorvolando in questa sede sulle evidenti istanze a livello normativo volte a disciplinare privacy, accesso e gestione alla basi dati, normazione della catena del valore in termini di diritti (copyright vs diritto d'autore), pare importante segnalare il ruolo cruciale dei data set con i quali il training di una IA è esperito: IA addestrate con dati "ideologici" esibiranno posizioni tendenziose, IA addestrate con dati parziali esprimeranno miopia e distorsioni⁶.

In tale senso il tema etico assume proporzioni enormi. Non esiste un modo "giusto" di addestrare una AI: avremo sempre dei bias, esattamente come con gli esseri umani. Queste IA rappresentano esattamente l'estrema variabilità delle menti umane; ci saranno quelle con le quali andremo d'accordo e quelle che riterremo totalmente idiote, esattamente come con i nostri simili.

Il meccanismo con il quale queste IA imparano è molto simile a quello con il quale impara un bambino, senza però avere la possibilità di sperimentare il mondo in prima persona e quindi di crearsi un senso critico, per formare il quale è (anche) necessario commettere errori subendone le conseguenze dirette. Da ciò derivano le forze e le debolezze di queste IA: eccezionali nello svolgere molti task, ma assolutamente inaffidabili dal punto di vista della autonomia operativa, intesa come "seniority".

Tutto ciò ci conduce ad alcune considerazioni sugli effetti delle IA sul mondo e sul mercato del lavoro.

Tutti i campi applicativi nei quali un giudizio e una validazione umani siano pratica intrinseca e sottintesa possono vedere grandi spazi per l'impiego di IA: se una immagine generata con MidJourney è brutta o non risponde alle esigenze di chi dovrà farne uso, l'operatore si limiterà a chiedere la generazione di una nuova immagine giungendo via via all'esito atteso. L'operatore è obbligato a esprimere un giudizio e, quindi, a validare l'operato della IA. Lo stesso può valere per la sintesi di una

⁵ Lo Spazio Latente è tecnicamente una varietà differenziabile (un "manifold") a centinaia/migliaia di dimensioni, intuitivamente non dissimile da uno spazio topologico euclideo n-dimensionale. Viene generato spontaneamente da algoritmi di apprendimento non supervisionato e ogni dimensione rappresenta "qualcosa". Una forte semplificazione è immaginare che esista una dimensione per il "colore", un'altra per la "forma" e via così, e che quindi il sistema raggruppi concetti in maniera topologica: tutti i concetti "rossi", come una Ferrari o la cappa del torero, saranno tra loro "vicini" lungo la dimensione del colore, così come lo saranno le parole "rosso", "red", "rouge" e "rot". In realtà è impossibile "aprire il cofano" di una IA e determinare quali siano le partizioni semantiche dello spazio latente, perché questo viene creato in maniera totalmente automatica ed imperscrutabile. Ogni dimensione "significa" molte cose diverse, tra loro mescolate. Questo spiega perché sia così difficile e tutt'altro che immediato correggere una IA che "si comporta male", ad esempio fornendo informazioni eticamente inaccettabili: sarebbe come agire a livello chirurgico sul cervello di un essere umano e pretendere di ottenerne gli stessi risultati.

⁶ Nel 2021 le IA di OpenAI non sapevano alcunché della pandemia poiché il loro ultimo training risaliva al 2019.

serie di testi che conosciamo o il meeting report di una riunione alla quale abbiamo partecipato. La validazione diviene pressoché immediata o, comunque, contestuale all'erogazione dell'output.

Molto più critico è invece l'utilizzo di una IA in contesti nei quali vi è il rischio che essa sia preda di allucinazioni: in mancanza di dati sufficienti per fornire una risposta plausibile "corretta" (e anche sulla correttezza si potrebbero sollevare temi a livello epistemologico e semiotico), la IA inventerà una risposta plausibile ma "sbagliata".

Si pensi quanto questo diventi potenzialmente critico nell'elaborazione di enormi volumi di dati, umanamente non affrontabili, o nello sviluppo di migliaia di righe di codice software nelle quali l'errore è riscontrabile solo a valle del processo ma dove la sua individuazione può diventare quasi impossibile. Sono ormai storia gli innumerevoli casi di ChatGPT chiamata a fare da avvocato con risultati sconcertanti⁷.

Questi sistemi sono progettati per fornire sempre una risposta e ciò porta a una conclusione: ad oggi non possono essere usati in maniera totalmente autonoma poiché una supervisione umana è imprescindibile.

Tale dinamica interessa peraltro anche il rapporto lavorativo tra esseri umani. La fiducia nei collaboratori è ciò che distingue un comportamento "senior", al quale possiamo demandare completamente un compito ed aspettarci dei risultati validi, da uno "junior" che dobbiamo invece in qualche modo supervisionare.

L'introduzione di ogni innovazione tecnologica impatta dapprima il *modus operandi* degli esseri umani per poi modificarne le strutture mentali. È ciò che accade sin dall'invenzione della scrittura, anch'essa oggetto di grandi paure e perplessità al momento della sua introduzione "massiva", come Platone descrive nel *Fedro*.

Parafrasando Stewart Brand, il grande filosofo ponte fra controcultura hippie e rivoluzione digitale, per cambiare il modo di pensare delle persone non serve educarle, basta semplicemente mettere loro a disposizione nuovi strumenti, perché possano fare in modo diverso le cose che già facevano o iniziare a fare cose completamente nuove.

La storia dell'umanità è infatti compenetrata a quella delle *techne*. Ce lo ha insegnato Andy Clark nel suo fondativo *Natural Born Cyborgs*⁸: il confine fra mente contenuta nel wetware, il nostro cervello, e in forme variabili di hardware esterno è destinato a continuare a spostarsi. È successo con i libri, con gli orologi da polso, con i computer e oggi con il fenomeno emergente chiamato Intelligenza Artificiale.

⁷ <https://masterx.iulm.it/chatgpt/usa-avvocati-usano-chat-gpt-scoperti-e-multati/>

⁸ Clark A. (2004), *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*, Oxford University Press, Oxford

Compito della comunità scientifica e intellettuale –così come del decisore politico– è tentare di comprendere per poter interpretare e mitigare i disequilibri che fisiologicamente ogni innovazione produce prima che la società sia in grado di trovare spontaneamente un nuovo baricentro.

IA: quali rischi e quali opportunità?

I rischi in termini di ricaduta occupazionale, a parere degli scriventi, ad oggi riguardano tutte quelle attività a basso controllo, peraltro da anni già interessate da fenomeni di automazione assimilabili a quelli della Rivoluzione Industriale (si pensi a chatbot, assistenti virtuali, sistemi di content automation per i social media), e richiedono uno sforzo nel liberare il potenziale umano all'interno della catena del valore, perimetrando e valorizzando tutti quegli ambiti nei quali la mente umana rimane fattore strategico decisivo.

L'attenzione deve quindi andare in direzione dell'upskilling e del reskilling dei lavoratori più vicini alle competenze "junior": se il 9% dei lavori è destinato a sparire perché sostituito da IA, ciò significa che quei lavori erano subumani in partenza non che la IA li stia "rubando" agli esseri umani!

Nell'ancorché limitato scenario operativo degli scriventi ciò che emerge in modo prepotente in termini di opportunità è invece la capacità di "ascolto" che l'utilizzo di IA esprime all'interno dei contenuti digitali. La IA permette infatti una customizzazione capillare di contenuti e esperienze capace di superare in modo definitivo le logiche dei media generalisti⁹, ivi compresi i social, nel segno di un approccio dialogico fra contenuto e Utente. Ciò si traduce in territori di opportunità professionale per l'umano che sino a pochi anni fa non esistevano oltre a percorsi di inclusione davvero incoraggianti.

Qualche esempio:

- Creazione di simulazioni interattive per la formazione del personale (e.g. industria 4.0) in grado di assumere comportamenti adattivi nei confronti dell'Utente;
- Costruzione di strumenti didattici capaci di generare senso di "presenza" (la nuova frontiera della DaD);
- Creazione di contenuti capaci di adattarsi in tempo reale alle istanze dell'Utente (deficit sensoriali, cognitivi, bisogni educativi speciali, target socio-culturali eterogenei, localizzazione linguistica...) durante percorsi di visita a musei, monumenti, siti archeologici o di luoghi pubblici in genere in una logica di inclusione mai sperimentata prima;
- Ricostruzione di ambienti virtuali a partire da materiali di repertorio (foto e video);
- Analisi e elaborazione in tempo reale di feedback neurofisiologici e loro impiego per la creazione di strumenti interattivi sempre più vicini alle esigenze dell'Utente;

⁹ cfr. Mazzolotti A. (2007), *Il se' sospeso. Il pubblico televisivo nell'era dell'interattività*, Università degli Studi di Parma, Parma.



- Generazione di digital twin di esseri umani la cui vita fisiologica è terminata ma dei quali esistono tracce digitali (filmati, scritti, testimonianze..).

Ciò detto, la IA rimane uno strumento, come lo è un cacciavite o un telaio jacquard, e come tale non è buona né cattiva in sé. Di certo la velocità esponenziale con la quale questa tecnologia sta evolvendo non ha precedenti nella storia dell'uomo e impone una capacità di "rincorsa" dell'attività normativa e progettuale molto sfidante.

Formare e informare rimangono gli elementi strategici cruciali per vincere questa sfida, accanto al superamento del rumore mediatico prodotto intorno a questi temi foriero soltanto di risposte emotive basate su entusiasmi acritici o su derive neoluddiste, entrambi figli dell'ignoranza, unico vero grande nemico del nostro futuro.

Roma, 7 febbraio 2024