

Chat GPT: l'intelligenza artificiale che ci costringe a pensare

Ci siamo progressivamente familiarizzati, e anche divertiti, con un'ampia gamma di applicazioni che realizzano prestazioni di vario genere, alcune delle quali oramai di uso tanto comune nei luoghi specializzati e nei nostri dispositivi, da non percepirle più come Intelligenza Artificiale: il riconoscimento e l'interpretazione di una forma (es: un carattere tipografico, una scena, un'impronta digitale, un volto, una radiografia, una configurazione di cellule), la traslazione di messaggi da una forma fisica a un'altra (es: suono-scrittura e viceversa). Perché ChatGPT ci ha sorpreso, non l'abbiamo vista arrivare, come si dice, e ci coinvolge forse nel più grande gioco filosofico di massa mai visto? Forse perché tutte le altre prestazioni dell'Intelligenza Artificiale ci simulano in quanto esseri capaci di percepire, riconoscere, trasformare, mentre ChatGPT pretende di simularci in quanto intellettuali.

Qualche premessa sull'Intelligenza Artificiale

L'idea di automatizzare i processi mentali ha dato vita, molto prima di essere un programma di ricerca, a domande scientifiche, filosofiche, etiche. La domanda base è quella formulata da Alan Turing in un saggio del 1950 (*Calcolatori e intelligenza*): *"Le macchine possono pensare?"*

La risposta di Turing è che, così posta, la domanda è priva di senso perché rimanda alla impossibile spiegazione di cosa è il pensiero. Per questo Turing propone un criterio empirico, un *test*, il famoso *"gioco dell'imitazione"*, che in sintesi invita a sostituire la domanda iniziale con una nuova domanda: *"Le risposte che una data macchina dà alle domande sono o no distinguibili da quelle di una persona?"*

Gli scienziati che lanciarono l'Intelligenza Artificiale come programma di ricerca la definirono così: *"...il problema dell'intelligenza artificiale è quello di far sì che una macchina agisca con modalità che sarebbero definite intelligenti se un essere umano si comportasse allo stesso modo"*

Quel progetto si è sviluppato nei quasi settant'anni successivi con un intricato labirinto di pratiche e paradigmi, senza giungere a una sistemazione teorica. L'approccio iniziale fu quello di chiedere aiuto alle scienze che si occupano della mente.

Da una parte alle *scienze cognitive e della linguistica formale*. L'idea era quella realizzare nella macchina strutture algoritmiche capaci di sviluppare, ad esempio, abilità linguistiche basate sulle grammatiche generative e memorie basate su reti semantiche. Oppure algoritmi capaci di imitare il pensiero euristico nel problem-solving, come nel *General Problem solving* di Newell e Simon.

Dall'altra parte ci si rivolge alla *neurologia* con reti neuroniche realizzate

con circuiti digitali o simulate nel computer. Il *Perceptrone*, studiato nell'ambito della Cibernetica, fu il primo prototipo.

In realtà né gli approcci teorico-formali, né quelli neuronali hanno dato risultati pratici apprezzabili, se non prototipi: il più delle volte semplici giocattoli. Il successo dell'IA è dovuto assai di più alle soluzioni ad hoc dell'ingegneria informatica. Così come la realizzazione degli automi industriali non nasce dai trattati di anatomia umana, ma dalla meccanica e dalla teoria dei controlli.

Due sono state le scelte decisive. Da una parte il *Machine Learning* che è la risposta a un'osservazione di Turing: gli uomini non nascono con le strutture cognitive già nella testa, ma con una struttura iniziale che li rende capaci apprenderle grazie alle interazioni e all'esperienza. Dall'altra parte la disponibilità di enormi masse di dati a disposizione creati ad hoc o disponibili in rete e l'applicazione di algoritmi di tipo probabilistico.

Produzione di testi per mezzo di testi come automazione di processi mentali. Il caso ChatGPT

Il sistema ChatGPT è capace di produrre, su richiesta o in un contesto dialogico, risposte di ogni tipo: spiegazioni, interpretazioni, illustrazioni, soluzioni di un problema ecc. E' dunque una *IA Generativa*. Ma può anche essere vista come un caso di *automazione di processi mentali*. Un'idea antica, coltivata per esempio da Leibniz, Babbage e Ada Lovelace.

I processi mentali *guidati da uno scopo* (vedremo oltre questo punto) possono essere schematizzati come una catena:

- Formulazione di una Richiesta
- Ricerca di informazioni
- Selezione delle informazioni utili
- Elaborazione e produzione del risultato

Questa catena può essere *automatizzata*. Occorre però, prima di tutto, accettare l'idea che gli atti mentali debbono in qualche modo essere materializzati (registrati), cioè diventino *testi*, nel senso più generale di questo termine.

Gli strumenti di base per l'automazione sono un Computer e la Rete delle informazioni (Il magazzino dei testi a disposizione). I *livelli di automazione* aumentano con il crescere dei dispositivi di cui è dotato il computer:

1. *Un Browser* fa trovare nella rete un testo di cui si conosce già l'indirizzo Internet
2. *Un Motore di ricerca* (Google) trova un elenco di indirizzi Internet di testi relativi a un argomento richiesto. Fino a questo livello le successive operazioni di selezione dei testi utili e la generazione del testo finale sono opera dell'uomo.
3. Un sistema di *IA Generativa* che assorbe tutti i passaggi e passa direttamente dal testo-richiesta al testo-risultato

Si può quindi parlare di *produzione di testi per mezzo di testi*

Le domande sull'IA Generativa

Si possono rivolgere a ChatGPT richieste di varia complessità: dal dare una spiegazione allo scrivere un saggio su una questione filosofica, commentare o produrre una poesia analizzare un quadro, risolvere un problema di matematica, di fisica, di meccanica ecc. ChatGPT non ripete a "pappagallo", ma mostra comprensione contestualizzando le nozioni e applicando i procedimenti a situazioni specifiche. Ogni tanto commette errori clamorosi, ma questo di per sé non la rivela come non umana, visto che anche gli uomini sbagliano.

Ma ChatGPT, mentre risponde alle nostre domande, ci impegna a farne altre a noi stessi, molto profonde e complesse. Discutiamone sommariamente alcune.

Questioni filosofiche e scientifiche.

Il criterio di Turing, che sembrava "tagliare la testa al toro" diventa meno certo quando ci si avvia a sistemi complessi come ChatGPT per almeno due motivi:

1. a prima vista le risposte sembrano fornite da uno studente scrupoloso e onnisciente, ma proprio la sua *onniscienza* e la *forma standard delle risposte* fanno venire dei dubbi sulla sua 'umanità',
2. il giudizio sui testi prodotti dalla macchina può dipendere dal livello culturale e dalla capacità critica di chi li legge.

Nella discussione in corso, nei libri e nei giornali, circola la domanda di sempre: fino a che punto la macchina può imitare (riprodurre, sostituire) il pensiero umano? Dove sta la soglia invalicabile? In cosa sono diversi i processi mentali imitabili da quelli inimitabili? Può essere utile ricorrere, come esempio, a due schemi teorici proposti da due pensatori.

Jung, nel suo saggio *Le due forme del pensare*, propone uno schema:

- Il *pensare indirizzato* è un pensare la realtà o, come potremmo anche denominarlo, il pensare con parole,
- Il *pensare non indirizzato* è un pensare puramente associativo. È il sognare o fantasticare.

Heidegger propone un'altra dicotomia tra:

- Il *pensiero calcolante* è quello *rivolto a uno scopo*,
- Il *pensiero meditante*

Sembra ragionevole affermare che solo il pensiero indirizzato o calcolante è quello imitabile da una macchina. Cioè, semplificando, *le macchine non possono sognare o essere filosofi.*

Con certezza si sostiene da tempo che le macchine, a parte gli automi di

Blade Runner, non possono avere sentimenti. Ma il problema è: esistono, nella vita umana, atti di pensiero in cui intelligenza e passione non si mescolano?

Tutta un'altra categoria di domande possono sorgere all'interno dei vari campi del sapere, IA e scienza, arte, letteratura: l'IA è un oggetto estraneo, uno scomodo concorrente o un nuovo strumento?

Questioni socio-culturali

L'IA ripropone, in forma diversa, i problemi che già aveva posto l'automazione del lavoro manuale: accettare le conseguenze o combattere contro? La risposta prevalente è: governare questo nuovo salto tecnologico. Ma questo significa niente di meno che pensare a una società diversa.

Più concretamente e nel breve termine: come cambiano le professioni della mente più aggredibili (la traduzione e l'elaborazione di testi, la progettazione standard, l'espressione di intrattenimento ecc.)? Si può pensare a un modo di cooperazione fra IA e persone?

Questioni etiche

Forse nessuna attività umana è stata tanto discussa per i suoi rischi quanto quelle connesse all'uso delle nuove tecnologie e dell'IA in particolare. Nel suo testo *Etica dell'intelligenza artificiale* Luciano Floridi ricorda che un cospicuo di organismi culturali e politici hanno formulato cataloghi di principi etici relativi all'IA, che comunque convergono su cinque principi fondamentali:

1. *Beneficienza*: promuovere il benessere , preservare la dignità e sostenere il pianeta
2. *Non maleficienza*: privacy, sicurezza e "cautela della capacità" (automiglioramento ricorsivo dei sistemi, cioè il noto rischio della macchine che sfuggono al controllo)
3. *Autonomia*: il potere di decidere di decidere
4. *Giustizia*: promuovere la prosperità, preservare la solidarietà, evitare l'iniquità
5. *EsPLICABILITÀ*: rendere possibili gli altri principi tramite l'intellegibilità e la responsabilità

IA Generativa e educazione

Come tutte le rivoluzioni tecnologiche anche questa genererà scomparsa di professioni (i call center, per esempio potrebbero essere totalmente automatici) o la loro trasformazione (I traduttori di testi non letterari potrebbero diventare revisori delle traduzioni automatiche). Questo pone il problema di realizzare davvero una prassi di educazione degli adulti.

La scuola sarà inevitabilmente coinvolta nella comprensione dei problemi dell'IA e dovrà fare le sue scelte, come già succede da tempo per tutte le nuove tecnologie. E dovrà affrontare il problema a due livelli.

I fini educativi e le competenze

A che livello l'avanzamento delle tecnologie delle mente può o deve influire sulla ridefinizione funzione della scuola:

1. in relazione alle *finalità generali*: cambia qualcosa nel profilo culturale ed etico delle persone destinate a vivere in un mondo in cui l'IA diventa una presenza normale e diffusa?
2. in relazione alle *competenze*: che fare delle competenze che "non serviranno più" perché automatizzabili? (Es. Redigere un resoconto, produrre un testo di sintesi, risolvere un problema ecc.). Abbandonarle e alzare il livello? Praticarle comunque fino a un certo livello scolastico (in analogia a quanto è successo con il calcolo e la scrittura manuale)?

La pratica didattica

A livello della didattica si ripropongono, in modo più complesso, le domande che già da tempo ci rivolgiamo sul rapporto con le tecnologie digitali

– La scelta di ignorare e proibire sembra perdente in partenza. Il problema è sempre quello di imparare a convivere con i sistemi tecnologici, usarli con intelligenza. Sistemi che sembrano rubare l'intelligenza non possono invece richiedere e promuovere un'intelligenza più alta e complessa?

Per esempio spostare la mente dal rispondere alle domande all'imparare a farsi domande più complesse?

-L'esplosione della curiosità, l'esplorazione e le vere e proprie "recensioni" alle risposte di ChatGPT di commentatori non può essere un indizio di come affrontare il problema in modo creativo? I ragazzi sono già entrati in questa ottica in modo ludico, diamolo per acquisito lavoriamoci sopra.

Qualche riferimento.

L. Floridi – *Etica dell'Intelligenza Artificiale* – Raffaello Cortina Ed.- Milano 2022

N. Cristianini – *La scorciatoia. Come le macchine sono diventate intelligenti* – Il Mulino -Bologna -2023

M. Mitchell-*L'Intelligenza Artificiale- Einaudi- Torino 2022*