

## ChatGPT: merce “intelligenza artificiale” nel recinto della stupidità e dell’orrore capitalista

Giovanni Pontano: *Quante volte oggi uno scrittore si domanda se lo scrivere non sia un istinto insopprimibile, ma antiquato. Questo lavoro così solitario, artigiano, questo mettere faticosamente una parola dietro l'altra. Questo lavoro che non si riesce a meccanizzare in nessun modo.*  
Sig. Gherardini: *Ma lei ne è veramente convinto?*  
Giovanni Pontano: *No.*  
Dal film “La notte” di Michelangelo Antonioni (1961)

L’idea che un oggetto da noi creato possa diventare vivo, veramente umano, affonda le sue radici in un immaginario atavico oltre che in un’ampia letteratura.

Già nella mitologia greca Pigmalione si innamorava perdutamente della statua di donna cui aveva dato forma egli stesso scolpendola da un blocco d’avorio. La scultura era infine diventata viva, e di Galatea assunse poi il nome, per intercessione della dea Afrodite su preghiera dello stesso Pigmalione.

Il Golem, letteralmente “materia grezza” o “embrione”, è nella cabala ebraica un gigante d’argilla in cui, a differenza di Adamo, non è stato ancora infusa l’anima.

Nel capolavoro di Mary Shelley il dottor Frankenstein (o, come recita precisamente il titolo, “moderno Prometeo”) aveva assemblato parti di cadaveri per realizzare la sua creatura e infonderle poi la scintilla della vita tramite potenti scosse elettriche. Le suggestioni degli esperimenti di Galvani permettevano già alla fantasia di attingere al progresso scientifico.

Pochi decenni prima circolava in Europa e negli Stati Uniti il cosiddetto Turco, un automa di legno in grado di giocare a scacchi, poi rivelatosi un imbroglio, avendo spazio per nascondere al proprio interno una piccola persona che ne muoveva i marchingegni.

*Le avventure di Pinocchio* di Carlo Collodi, pare il libro più tradotto al mondo dopo la Bibbia, cos’altro non è se non il desiderio che qualcosa da noi realizzato, proprio perché imperfetto ancorché animato, svolga un percorso per giungere finalmente ad essere pienamente umano: il burattino vuole disperatamente diventare un bambino vero e ci riesce grazie anche alla fata dai capelli turchini.

Non dissimile è la ricerca dell’androide Data, nella fortunata serie televisiva *Star Trek: The Next Generation*, di ottenere il chip emozionale ed essere finalmente equiparato alla dignità della specie del suo creatore.

«Ti chiedi io, Creatore, dall’argilla di crearmi uomo, ti chiedi io dall’oscurità di promuovermi...?», si chiede John Milton nel *Paradiso perduto*. Eppure, alcune delle migliori menti, nei diversi modi di produzione che si sono succeduti nella storia, si sono gingillate con l’idea di realizzare – dal marmo, dal legno, dal ferro – una creatura a nostra immagine e somiglianza, in una analogia calzante rispetto a come nella tradizione giudaico-cristiana han voluto raffigurare la Genesi.

Con l’informatica i processi meccanizzati, via via più complessi, che simulano quelli cerebrali, si sono schiusi alle possibilità dell’industria. Dal precursore della calcolatrice meccanica ideata da Blaise Pascal, la pascalina (datata 1642), passando per il primo calcolatore interamente elettronico (l’Eniac del 1946, composto da 18 mila valvole, dal peso di 30 tonnellate e poggiante su una superficie di 140 mq) e arrivando alla moderna Intelligenza Artificiale, i successi di questo ramo tecnologico sono stati stupefacenti.

Il lancio sul mercato delle chatbot, e di ChatGPT in particolare nelle sue ultime due versioni, hanno attirato l’attenzione mediatica del mondo intero, alimentando accesi dibattiti. In due

mesi la creatura di OpenAI, che nella maggior parte dei casi è in grado di superare egregiamente il famigerato test di Turing, ha raccolto cento milioni di utenti. Tik Tok ha impiegato 9 mesi per arrivare a quella soglia, Instagram 30 mesi, Uber 70 mesi.

Ma in Italia il servizio è stato sospeso d'ufficio per svariate settimane, mentre dalla Silicon Valley mille personalità di spicco – tra cui Elon Musk (Tesla), Steve Wozniak (Apple) e Jaan Tallin (Skype) – hanno firmato una lettera chiedendo di sospendere ChatGPT per sei mesi e riflettere bene sulle possibili implicazioni pratiche e sui rischi.

Fazioni (borghesi) interessate, e schierate in tifoserie avverse, prefigurano futuri radiosi grazie al salto tecnologico compiuto dall'IA e alle sue promesse di per sé redentrici, oppure paventano oscuri scenari distopici che si riallacciano a tanti spauracchi narrativi e cinematografici. Da *Io, Robot* di Asimov in cui vengono contraddette le leggi della robotica, ad *Hal 9000* di *2001: Odissea nello spazio* di Arthur C. Clarke e Stanley Kubrick (dove le iniziali del super-computer che ci si ritorce contro non sono che le lettere precedenti a quelle di IBM), dal film *War-Games – Giochi di guerra*, in cui un software rischia di lanciare veramente le testate termonucleari delle grandi potenze e avviare la Terza guerra mondiale, alla saga di *Terminator*, in cui Skynet prende coscienza e scatena il giorno del giudizio contro la specie umana.

Il fortunato termine “robot” deriva dal ceco “robota”, lavoro forzato, e lo si trova per la prima volta nel romanzo *R.U.R.* di Karel Čapek del 1920 (anche se i robot in questo caso non erano esseri meccanici quanto piuttosto dei replicanti alla *Blade runner*). Come già si troverà esplicitato in questo documento è chiaro che, stante il sistema capitalista, anche le migliori e sbalorditive meraviglie tecnologiche, come ChatGPT, non sono che merci sofisticate, al cui servizio viene piegata l'intelligenza viva, umana, di esseri che per casualità si ritrovano a vendere la propria forza-lavoro in cambio di un salario, alla stregua di robot fatti di carne e sangue. Quali moderni schiavi, gli schiavi salariati contribuiscono a realizzare ed affinare una merce per ora morta, l'Intelligenza Artificiale, che solletica però l'immaginario collettivo. Ma la tecnologia fa anche altro, e proveremo a renderne conto in prossimi elaborati.

Quali tipologie di lavori, in virtù dell'ascesa dell'IA e in particolare delle chatbot, verranno velocizzati suscitando quindi importanti ristrutturazioni? Che mansioni saranno spazzate via e relegate nel museo e nei libri di storia, generando perciò, nell'immediato, disoccupazione? Quali settori, e in che misura produttivi di plusvalore o parassitari, verranno ad essere concentrati? Che riflessi vedremo sulla composizione della forza-lavoro internazionale?

E ancora: quali gruppi industriali capitalistici sono coinvolti nella corsa alle nuove frontiere dell'intelligenza artificiale? Come gli Stati si doteranno di propri strumenti di utilizzo e controllo di questi ritrovati che non saranno ideologicamente neutri e né tanto meno automaticamente sovversivi rispetto alle leggi fondamentali di funzionamento del capitalismo? Che implicazioni militari sono in corso, e sono ipotizzabili, per l'IA applicata ai dispositivi bellici di cui dispongono le borghesie nazionali in lotta tra loro?

Cominciamo però dapprima col vedere concretamente alcuni aspetti di interesse relativi a ChatGPT.

### ***Di cosa si tratta, come funziona***

Partiamo dal funzionamento, prendendo in esame ChatGPT, la chatbot di OpenAI rilasciata a novembre 2022 nella versione ChatGPT-3 e, a marzo 2023, ChatGPT-4.

Questo programma, così come le alternative delle altre grandi aziende (spesso costrette a rilasciare versioni incomplete solo per tenere il passo in questa nuova corsa che ricorda quella allo spazio) permette di dialogare con un assistente virtuale, in grado di rispondere in maniera coerente alla maggior parte delle richieste e delle domande. L'aspetto che più colpisce è però la capacità di scrivere interi paragrafi su qualsiasi argomento, ma il funzionamento è ben diverso dagli ormai ubiquitari assistenti vocali (Siri, Alexa, Google...).

Si tratta infatti di un modello di linguaggio probabilistico, che va a completare la frase aggiungendo volta per volta una parola che possa avere senso nel determinato contesto, integrandosi con la porzione di testo che la precede.

Per riuscire a fare ciò, è necessaria una prima fase nella sua programmazione, detta di addestramento. Per formare delle frasi di senso compiuto e rispettanti particolari parametri, va infatti considerata la relazione di ogni parola con quelle precedenti, necessitando però così di potenza di calcolo altissima (per una frase di 13 parole esistono  $10^{70}$  combinazioni diverse possibili). Per approssimare e semplificare queste relazioni tra le parole ci sono più modi, tra cui la serie di Fourier, la serie di Taylor e, soprattutto, la rete neurale. Le reti neurali vengono utilizzate per costruire dei “large language models”, modelli linguistici addestrati con miliardi di parole (provenienti da internet e dai libri disponibili digitalmente), in grado di replicare con una buona accuratezza varie abilità, dal calcolo alla scrittura alla programmazione e alla creazione di immagini.

Il language model utilizza sistemi di calcolo probabilistico per predire quale parola sarà la più appropriata per proseguire una frase, tenendo conto della domanda e delle parole che la precedono. Questo viene fatto assegnando diversi valori di probabilità ad ogni nuova parola (basandosi sul data set originale estratto da libri, articoli, chat, forum e blog), e scegliendo quali utilizzare tra le varie opzioni plausibili (con una probabilità superiore a zero, e quindi “sensate”).

Scegliendo sempre la parola più probabile, però, il discorso risulta artificiale e piatto, motivo per cui questi programmi spesso modificano la scelta delle parole, dando maggior varietà e realismo. Si tratta infatti di un modello stocastico, e non deterministico (in cui a ogni input corrisponde un output sempre uguale), per cui la stessa richiesta potrà dare esiti molto diversi tra loro, sebbene ciò non sia dovuto ad una vera e propria “evoluzione” dell’IA.

Non si tratta di un semplice motore di ricerca che rielabora i dati ottenuti online, ma di un programma in grado di funzionare anche offline. Questo avviene poiché la parte più importante nello sviluppo è quella dell’addestramento, da cui dipenderà la quantità di informazioni a disposizione dell’IA e quindi la capacità di rispondere. Una volta conclusa la fase di addestramento, l’IA non sarà aggiornata, e non sarà quindi nemmeno in grado di parlare di fatti seguenti il rilascio (negli ultimi sviluppi, però, è stata implementata, in alcune chatbot, la possibilità di ricercare i risultati più recenti da internet per integrare le proprie risposte).

Questa fase di addestramento è quindi fondamentale per la realizzazione di un’intelligenza artificiale. Le tre fasi di addestramento di ChatGPT sono:

1. Pre-addestramento generale, in cui i dati testuali vengono “assorbiti” e catalogati;
2. Aggiustamento supervisionato, in cui si ha come scopo l’emulazione di una conversazione tra umani, in cui il contenuto delle risposte dipende anche dal contesto dei messaggi precedenti;
3. Apprendimento per rinforzo con feedback umano, in cui un feedback umano giudica le varie opzioni di risposta proposte dal programma, permettendogli di capire quali siano più adatte (nelle versioni più recenti, questo processo può continuare anche in seguito al rilascio ufficiale, sia sfruttando il parere dell’utente finale, che può lasciare un “like” o un “dislike” alla risposta prodotta, sia da lavoratori con questo specifico compito).

### ***Intelligenze umane appendici di macchine***

Quest’ultima fase si è rivelata essere fondamentale, poiché, limitandosi a rielaborare i testi trovati online, l’IA spesso rispondeva con contenuti inappropriati, perché inventati (è più difficile per un programma riuscire a distinguere tra libri di fantascienza, articoli di giornale e forum) o caratterizzati da forti bias comuni online, come razzismo, sessismo o violenza.

Si è così reso necessario un ulteriore step nell’addestramento, in cui si è cercato di “purificare” (impossibile farlo del tutto) le risposte. Nel fare ciò, è stata fondamentale la manodopera umana. Per poter riconoscere discorsi d’odio, di abusi sessuali e violenza, è stato necessario “nutrire” l’IA con decine di migliaia di testi provenienti dagli angoli più reconditi della rete, in cui si descrivevano con vividi dettagli “abusi sessuali su minori, bestialità, omicidi, suicidi, torture, autolesionismo e incesto”.

Un’inchiesta del Time (18 gennaio 2023, Billy Perrigo, “Exclusive: OpenAI Used Kenyan

Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic”) ha svelato come questa operazione sia stata condotta da lavoratori provenienti da Kenya, India ed Uganda, che hanno passato nove ore al giorno (per una cifra compresa tra 1,32 e 2 dollari l’ora) a visualizzare e catalogare i contenuti più tossici e violenti di internet.

I lavoratori intervistati (mantenendo l’anonimato per paura di ripercussioni sulla loro vita) hanno dichiarato di soffrire di visioni ricorrenti di bestialità in presenza di minori ed altre atrocità, definite come “torture” che li hanno lasciati segnati, e le promesse di incontri singoli con un terapeuta sono state disattese.

La ditta che si occupava di gestire questo addestramento secondario, Sama, collabora con Google, Meta e Microsoft, e si definisce un’azienda di “intelligenza artificiale etica” e che – afferma – ha “contribuito a far uscire dalla povertà oltre cinquantamila persone”.

Sama riceveva da OpenAI 12,50\$ orari, tra le sei e le nove volte più di quanto guadagnassero i lavoratori. Lo scaricabarile delle responsabilità tra OpenAI e Sama non si è fatto attendere, e, otto mesi prima del tempo previsto, la collaborazione si è interrotta. Prima, però, un nuovo accordo riguardante la categorizzazione di 1.400 immagini (per un totale per OpenAI di 787,50\$), spesso illegali per la legge Usa e sempre riguardanti pedofilia, bestialità, stupri e violenza era stato completato. La cessazione del rapporto fra le società ha portato al licenziamento di alcuni dei circa 36 membri del team di “filtraggio contenuti” e al declassamento di altri. In seguito allo scandalo di questa inchiesta, non sono stati rinnovati anche i contratti milionari per la moderazione di contenuti su Facebook, causando ulteriori centinaia di licenziamenti.

Anche nei maggiori progressi della tecnologia, quindi, in una società capitalista, lo sfruttamento dei lavoratori in compiti ripetitivi e traumatici è fondamentale per aumentare il margine di profitto. ChatGPT non fa eccezione.