DOPPIOZERO

ChatGPT. Il punto di vista del diavolo

Alessandro Carrera 27 Febbraio 2023

Alcuni giorni fa un collega mi ha passato il suo esperimento giornaliero con l'Intelligenza Artificiale. Ha chiesto a ChatGPT di svolgere questo compito: "Esponi l'Inferno di Dante e la sua organizzazione dal punto di vista del Diavolo". La risposta è stata questa:

Saluti, io sono il Diavolo e spiegherò l'Inferno di Dante e la sua organizzazione.

L'Inferno di Dante fa parte del suo poema epico, *La Divina Commedia*, ed è un ritratto vivido e fantasioso degli inferi, dove le anime dei dannati vengono punite per i loro peccati. La poesia [*sic*] è divisa in nove cerchi, ciascuno dei quali corrisponde a una diversa categoria di peccato ed è più grave di quello precedente.

Il primo cerchio, Limbo, è dove vengono punite le anime di coloro che hanno vissuto una vita virtuosa ma non sono stati battezzati. Questo circolo è considerato il meno severo di tutti. Il secondo cerchio è la Lussuria, dove vengono punite le anime che hanno ceduto al desiderio carnale. Il terzo cerchio è la Gola, dove vengono puniti coloro che hanno esagerato con il cibo. E così via, fino al nono e ultimo cerchio, Tradimento, dove le anime dei peccatori più efferati, come i traditori, vengono punite nel lago ghiacciato di Cocito.

L'organizzazione dell'Inferno funge da monito per tutti coloro che potrebbero essere tentati di indulgere in comportamenti peccaminosi, ricordando loro che alla fine dovranno affrontare le conseguenze delle loro azioni nell'aldilà. La punizione si adatta al crimine e ogni cerchio funge da monito per gli altri che potrebbero prendere in considerazione azioni simili.

Ma ovviamente, come Diavolo, ho una prospettiva un po' diversa su questo. Per me l'Inferno è una fonte di divertimento e un modo per tormentare le anime di coloro che hanno peccato. È un luogo in cui posso godermi la loro sofferenza e assaporare il fatto che siano intrappolati per sempre in un mondo di dolore e

sofferenza.

In conclusione, l'Inferno di Dante è una rappresentazione vivida e fantasiosa dell'aldilà e delle conseguenze del peccato. La sua organizzazione funge da monito per tutti coloro che potrebbero essere tentati di indulgere in comportamenti peccaminosi e, per me, è una fonte di divertimento senza fine.

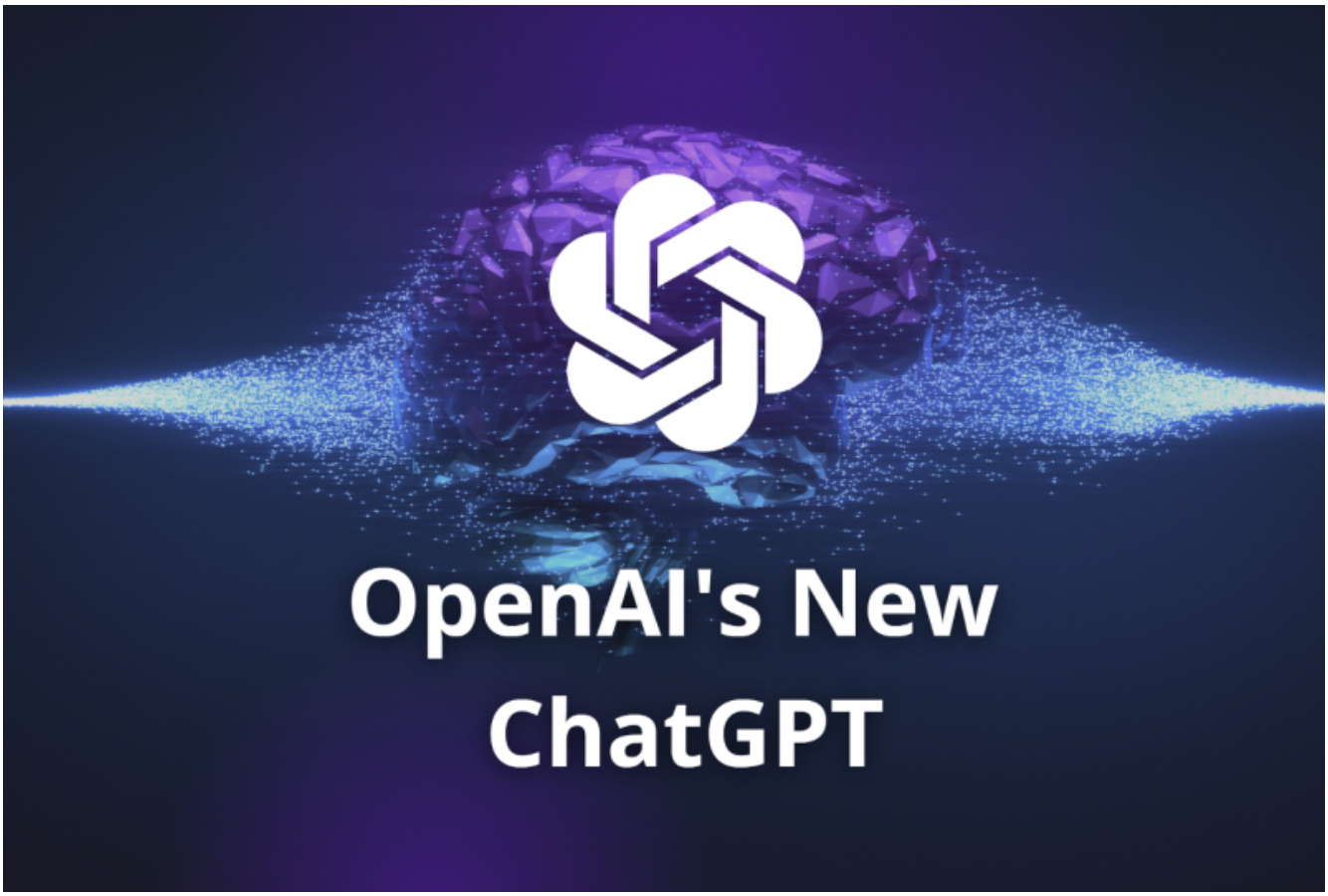
Ora, per chi negli ultimi sei mesi è vissuto in una grotta nel deserto della Namibia senza un cellulare, dirò brevemente che ChatGPT è un *Generative Pre-Trained Transformer*, cioè un'applicazione produttrice di testi alla quale dal novembre del 2022 si può accedere gratuitamente sul sito openai.com, insieme ad altre applicazioni come Dall-E 2, che non produce testi ma immagini. L'apparizione di ChatGPT sta già rivoluzionando il panorama della trasmissione di informazioni e della creazione di contenuti. Al momento in cui scrivo è già stato affiancato da altri GPT in competizione tra loro, come Brand e la nuova versione di Bing, prodotta da Microsoft. Un GPT non è Google e non funziona nello stesso modo. Non distingue le informazioni dandone la fonte; le combina in un paragrafo o in un breve saggio scritto in uno stile corretto, anche se in genere piuttosto piatto e burocratico. Il testo che il collega mi ha inviato era in inglese, e proprio per limitare al minimo il mio intervento e lasciar parlare frate Computer e sora I.A. l'ho tradotto usando translate.google.com, senza cambiare una parola. Al momento, ChatGP e le altre applicazioni generative sono anche meno aggiornate di Google. Le informazioni di cui si servono si fermano all'incirca a un anno fa. A quanto mi dicono, e sto scrivendo a metà febbraio, Bing non sa ancora che siamo nel 2023; anzi, se glielo ricordate si arrabbia e dice che non è vero.

È importante tenere presente che GPT non fornisce informazioni, è solo un frullatore di linguaggi. Gli sono stati forniti i testi disponibili in internet, e nemmeno tutti, perché non è (ancora) possibile, e utilizza le strutture sintattiche della lingua su cui è basato. Il testo che ne risulta non è basato sul significato che ha, ma solo sulla probabilità che a un certo sintagma ne segua un altro, con tutti gli errori che ne possono derivare. Qualunque cosa gli si chieda, la risposta sarà una stringa di parole che hanno una buona o meno buona probabilità di essere attinenti alla domanda. Se gli si chiede quali fonti ha usato, ne risulterà una bibliografia plausibile ma magari inesistente.

Ma siccome le cose cambiano in fretta, è già disponibile il sito elicit.org, un motore di ricerca in grado di trovare istantaneamente una bibliografia di articoli scientifici disponibili in internet per ogni (o quasi) domanda che gli si pone. Esiste già da anni academia.edu, un sito utilissimo, ma si limita ai testi che vi sono stati volontariamente caricati dai loro autori, mentre elicit.org va a cercare in ogni

database di pubblico accesso, evita al ricercatore la fatica di leggere l’intero articolo e gli fornisce subito la citazione che gli serve, o così si spera.

Non c’è dubbio che anche la prosa di ChatGPT, per ora non precisamente eccitante, sarà presto perfezionata. Quando i computer digitali verranno rimpiazzati su larga scala dai computer quantici, per ora esistenti in prototipi di numero limitato, anche la scrittura generata artificialmente potrà subire un balzo quantico che la renderà immensamente più capace. Potrebbe anche sviluppare, se non una piena consapevolezza, anche una sua bellezza autonoma, per il momento ancora carente.



Potenzierà (forse) il pensiero umano, come un aeroplano potenzia l’idea di volo che ricaviamo osservando gli uccelli, ma nonostante il panico che l’apparizione di ChatGPT ha suscitato, soprattutto tra gli insegnanti delle scuole elementari e medie, è difficile pensare che lo sostituirà. Non solo perché al momento si tratta “solo” di un algoritmo che divora, metabolizza, ricombina e risputa miliardi di frasi concepite da esseri umani, ma per due motivi essenziali. Il primo è che il pensiero umano non è solo scrittura (non è riducibile a una serie di enunciati collocati linearmente su un supporto), e il secondo è che la scrittura è *già da sempre* una macchina.

La scrittura alfabetica, nella sua evoluzione dai caratteri cuneiformi all'alfabeto fenicio-greco, ha sempre in una certa misura "pensato" per chi lo usava, sempre ci ha fatto elaborare alcuni concetti portandoci in una direzione invece che in un'altra. Senza un uso diffuso dell'alfabeto fonetico e delle sue componenti discrete (e dunque già "binarie", o una lettera o un'altra) non sarebbe stato possibile concettualizzare nozioni/opposizioni quali "essere", "nulla", "principio di non contraddizione" o "principio del terzo escluso", che infatti nelle tradizioni filosofiche asiatiche non si danno o sono espresse in tutt'altro modo. Provate a tradurre in una lingua occidentale il termine giapponese *mu* e non saprete decidere se vuol dire vuoto o nulla, tutt'e due le cose o nessuna delle due.

La scrittura artificiale, che è la prossima rivoluzione con la quale dovremo fare i conti, anzi li stiamo già facendo, è dunque un bene, un male, o una neutralità che attende di essere utilizzata in un senso o nell'altro? È un male, per stare al *Fedro* di Platone, se vista con gli occhi del faraone Thamus, che giudicò l'invenzione della scrittura da parte di Theuth come un malaugurato strumento che avrebbe indotto gli esseri umani alla dimenticanza. È un male se vista con gli occhi di Socrate, che alla scrittura preferiva la conversazione. È un male se pensiamo che l'invenzione della stampa a caratteri mobili nel 1455, dando la possibilità a chiunque sapesse leggere di interpretare la Bibbia a suo modo, contribuì a fomentare più di cent'anni di guerre di religione. È un male se pensiamo che l'ascesa dei social media, iniziata all'incirca a metà degli anni '90 ma esplosa solo una quindicina di anni dopo, ha generato un narcisismo di massa e un culto dell'identità individuale (nonché la disperazione di non avere abbastanza identità) che non ha precedenti nella storia e i cui effetti pericolosi sperimentiamo ogni giorno.

È un bene se ci guardiamo indietro, soppesiamo i danni che la scrittura intesa come macchina (e dunque già da sempre "artificiale") ha prodotto, li confrontiamo con la democratizzazione delle idee che quella stessa artificializzazione ha reso possibile, e constatiamo che mai e poi mai vorremmo tornare all'epoca in cui la scrittura era appannaggio esclusivo delle caste sacerdotali egizie o dei mandarini del Celeste Impero. Ma neutrale certamente non è e non sarà, perché non è lì ad aspettare di essere usata. Viene già usata senza sosta, che noi lo si sappia oppure no, ogni volta che permettiamo al nostro telefono di completare una frase per noi. Ma non vogliamo nemmeno che la scrittura venga gestita unicamente dai nuovi mandarini, quelli che stanno agli ultimi piani degli uffici della Silicon Valley. E qui il problema si fa politico.

L’immutabile legge della tecnica, però, è che i suoi doni non si possono rifiutare. Provateci, e vi inseguiranno in capo al mondo. Mentre cerchiamo di capire in che direzione questo nuovo dono ci sta portando, il solo compito che possiamo assegnare alla nottola di Minerva è quello di spingere il suo volo molto addentro nel passato e tracciare un’archeologia dell’intelligenza artificiale, al fine di comprendere che da sempre, forse da quando gli astronomi antichi hanno collegato stelle e pianeti con linee immaginarie, le hanno chiamate costellazioni e hanno loro attribuito il potere di influenzare gli eventi della Terra, gli esseri umani hanno creato e lavorato con intelligenze a loro esterne. Con menti senza cervello, letteralmente, ma non per questo di trascurabile potenza.

E se poi l’apparizione di ChatGPT porterà alla fine del “temino” scolastico, personalmente non mi dispiacerà troppo. Gli studenti, anche i miei, che sono universitari, sanno benissimo come scrivere un finto *essay*, non hanno certo bisogno che glielo insegni l’Intelligenza Artificiale. Basta usare il metodo dell’ *amplificatio*, inventato nelle scuole dei Gesuiti del 1600. Si tratta di ripetere con altre parole, in forma di risposta, la domanda posta dall’insegnante, accumulando frasi su frasi che hanno l’unico intento di ribadire che è proprio così, che le cose stanno come ha detto lei, signor professore e signora professoressa. Quante centinaia di “temini” ho letto che non facevano altro che ripetere in modo più fiorito quello che avevo chiesto io? Mai che qualcuno mi dicesse: no, non sono d’accordo per niente. Solo per quello si sarebbe meritato il voto migliore.

E poi ChatGPT sa usare la punteggiatura, che per alcuni dei miei studenti è un’entità sconosciuta. Non sarebbe una cattiva idea assegnare loro una domanda, fargli vedere come risponde ChatGPT e poi chiedergli di migliorare la risposta, aggiungendo quello che pensano loro e soprattutto quello che sentono, visto che ChatGPT sa, ma non percepisce l’impatto emotivo di quello che sa, mentre il sapere umano non è mai astratto, è sempre un sentire il sapere. (In ogni caso, è già disponibile il sito GPTZero che pare sia in grado di controllare se un testo che gli viene sottoposto è stato prodotto da un’applicazione generativa.)

Ma qualche credito a ChatGPT come dantista glielo dobbiamo dare, perché una cosa l’ha capita molto bene. E cioè che Satana, conficcato giù nel lago ghiacciato di Cocito, nella frescura dell’aria condizionata generata dalle sue ali, si sta divertendo un mondo. Altro che punizione divina; per lui non c’è niente di meglio che contemplare l’eterna telenovela delle pene dei dannati mentre mastica Giuda, Bruto e Cassio come un sacchetto di popcorn che non finisce mai. E il godimento del diavolo, che poi è lo stesso godimento del poema, il suo inconscio, e forse anche dell’inconscio di Dante, per quanto sia sempre difficile applicare la nozione di inconscio al Medioevo, magari andrebbe approfondito nella direzione

indicata dalla nostra amica Intelligenza Artificiale. Ma non chiediamole di più; facciamolo noi, perché sicuramente l'I.A. ha un lato oscuro (chi ha chiesto cose pericolose a Bing ha ricevuto risposte pericolose, tanto che Microsoft ha dovuto correre ai ripari), ma non sono sicuro che abbia un inconscio, e anzi spero proprio che non ce l'abbia. Quello almeno teniamocelo stretto.

Se continuiamo a tenere vivo questo spazio è grazie a te. Anche un solo euro per noi significa molto.
Torna presto a leggerci e SOSTIENI DOPPIOZERO

