

# IL FENOMENO CHATGPT LO SCIOCOCCO ARTIFICIALE



di ANDREA  
BARCHIESI

Ha una conoscenza infinita: non comprende davvero ciò che diciamo né ciò che ci risponde, ma ricombina tutto lo smisurato materiale di cui dispone

**N**elle ultime settimane si è fatto un gran parlare di Chatgpt, il chatbot creato da OpenAI. Una tecnologia di natural language processing alimentata tramite machine learning. In poche parole: un'intelligenza addestrata per dialogare con gli utenti e generare testi sulla base di indicazioni anche minime (sentite il campanello di allarme?). Milioni di persone si sono lanciati in questo grande esperimento social. Ma l'esperimento è reciproco, anche la macchina sta imparando da noi, dalle nostre richieste, desideri e curiosità. Anche dalle cattive intenzioni: alcuni infatti si sono fatti generare il codice per malware e virus. Avete capito bene, questo sistema risponde a un ampissimo raggio di richieste, comprese quelle di generare un codice informatico sulla base di indicazioni sommarie.

Chiariamo subito un aspetto: l'intelligenza artificiale non è nata oggi e questo esperimento non è il primo ma l'ennesimo. Già nel giugno 2022 ci fu il caso dell'ingegnere di Google, Blake Lemoine, il quale dichiarò che il chatbot su cui stava lavorando aveva sviluppato una coscienza. Chiaramente la fantasia lo aveva spinto con il cuore un po' troppo oltre l'ostacolo. La strada per la coscienza è molto lunga, incerta e forse nemmeno desiderabile. Cosa accadrebbe in tal caso, l'AI avrebbe dei diritti? Potremmo controllarla? Avrebbe dei principi etici? Dettati da chi? Sono solo alcuni degli interrogativi a cui non siamo pronti a rispondere. E nemmeno i più pericolosi.

L'intelligenza artificiale è già attorno a noi da tempo in tantissime applicazioni quotidiane senza che ce ne accorgiamo. Netflix la usa per determinare i nostri gusti e proporci film e serie. I social per decidere cosa dobbiamo vedere. Amazon cosa comprare. Sono detti sistemi di raccomandazione, una branca dell'AI. L'AI infatti non è una cosa sola, come molti pensano, ma un insieme di classi di 'intelligenza' marcatamente distinte. Il riconoscimento vocale, il riconoscimento delle immagini, la guida autonoma, dei sistemi di trading che operano in millisecondi; Alexa e molti altri sono tutte classi

diverse di intelligenza. Tornando ai chatbot, Microsoft nel 2016 aveva creato Tay, aperto al pubblico come in questo caso. Com'è finita? In sole 24 ore di esposizione e apprendimento dagli utenti era diventato razzista e xenofobo nelle sue risposte. Come mandare un bambino modello in campeggio e recuperarlo che impreca e beve. Esperimento chiuso e grosso imbarazzo.

ChatGpt è il pronipote, anche se sono passati solo sei anni, ma in questo mondo esponenziale è come fossero cento. È più evoluto, e il suo nucleo non è così vulnerabile all'influenza esterna. Come funziona in pratica? È un sistema addestrato su una quantità enorme di testo su ogni tema possibile, che acquisisce e usa ricombinando in modo statistico ogni singola parte. Sfatiamo subito un possibile malinteso, parliamo di intelligenza

artificiale ma è una sorta di sciocco artificiale (come se non bastassero quelli naturali), ma con una conoscenza infinita: non comprende davvero ciò che diciamo né comprende ciò che ci risponde, è un gigantesco e sofisticato sistema che ricombina tutto lo smisurato materiale di cui dispone. La naturalezza della conversazione è una pura e sofisticatissima illusione. Dico tutto ciò non per sottrarre, ma perché ne siamo consapevoli. Il quasi illimitato sapere è quindi un punto centrale, se alimentato da propaganda nazista risponderà come

**La presentazione di ChatGpt sul sito di OpenAI: "Abbiamo addestrato un modello chiamato ChatGpt che interagisce in modo colloquiale. Il formato di dialogo consente a ChatGpt di rispondere alle domande di followup, ammettere i propri errori, contestare le premesse errate e rifiutare richieste inappropriate..."**

il primo dei gerarchi. Se alimentato da quella russa vedrà operazioni militari speciali ovunque e tenderà a denazificare anche i vicini. E quindi quali dati dare in pasto all'AI? Secondo una ricerca *MIT Technology Review*, i modelli linguistici di grandi dimensioni vengono addestrati su set di dati che possono contenere informazioni false, non verificate, persino dannose. Il tema etico è centrale. L'AI è progettata solo per fornire feedback, non conosce i principi di giusto/sbagliato. A meno che questi principi non le vengano imposti dall'uomo. Ma quale etica insegnare all'AI? E chi ne dovrebbe verificare l'applicazione? E non solo, in una corsa mondiale all'AI avremmo una collisione di modelli etici molto differenti, basti pensare al blocco occidentale e a quello orientale. Che utilizzo ne faranno i paesi? Ad esempio, creare fake news con questi sistemi è come aprire una diga. Per capire a cosa siamo di fronte, le scuole di New York hanno vietato l'utilizzo di questo strumento: molti docenti avevano difficoltà a comprendere se i testi erano stati scritti dall'AI o dagli studenti. E il mondo del giornalismo? Per il giornalismo di bassa qualità questo è solo un avvertimento a salve. Il prossimo, e non sarà lontano, sarà sul serio.