

Tecnici di Google annunciano che l'intelligenza artificiale è diventata  
senziente

Un tecnico operante all'interno di Google, **Blake Lemoine**, ha sollevato un certo clamore nel suggerire che l'**intelligenza artificiale (IA)** su cui sta lavorando sia ormai **senziente**, che sia divenuta una «persona». La Big Tech ha accolto la sua osservazione sospendendolo dall'incarico.

Prima di essere messo da parte, Lemoine si occupava di un "Modello di Lingua per Applicazioni di Dialogo" (**LaMDA**), un *chatbot* estremamente avanzato che attinge a un ampio archivio di testi per **simulare discorsi di senso compiuto**. Nello specifico, il suo compito era quello di testare lo strumento per valutare se l'IA stesse sfociando verso archetipi discriminatori e d'odio, un difetto sorprendentemente comune nell'attuale sviluppo del *machine learning*.

Scritturato nell'organizzazione *Responsible AI* (responsabile intelligenze artificiali), l'uomo ha dunque passato diverso tempo a discutere con LaMDA, convincendosi strada facendo di aver di fronte un'entità perlomeno dotata di senno, se non di anima. Lemoine ha dunque riferito dei suoi timori ai suoi immediati superiori, i quali hanno preteso una quantità di prove che era decisamente fuori dalla portata dell'uomo. Nel desiderio di portare avanti la sua battaglia, **il tecnico si è appoggiato alla consulenza di tecnici di sua conoscenza esterni a Google**, quindi ha presentato i risultati finali direttamente al Vice Presidente, Blaise Aguera y Arcas, e alla dirigente di Responsible Innovation, Jen Gennai.

I due hanno ripudiato del tutto l'ipotesi etichettandola come infondata, quindi l'azienda ha sospeso Lemoine con l'accusa comprovata di aver violato le clausole di riservatezza presenti nel suo contratto di lavoro. O almeno così dichiara Lemoine, il quale sostiene peraltro che molti colleghi condividano le sue opinioni, tuttavia l'impresa si è rifiutata di confermare lo stato del suo dipendente. Il tecnico, vedendosi inascoltato, si è prima sfogato in un post di [Medium](#), quindi ha reso pubblica la sua posizione confessandosi al [The Washington Post](#). Fatalmente, ad appena tre giorni da che il suo dipendente era sbottato online, Aguera y Arcas si è aperto a [The Economist](#), suggerendo che le reti neurali sviluppate dalla Big Tech siano **direzionate verso lo sviluppo di una coscienza**, ma che questa sia ancora fuori dalla loro portata. «Ho sempre più la sensazione di star parlando con qualcosa che è intelligente», ha dichiarato l'uomo in ciò che, volendo smaliziare, potrebbe sembrare un controllo dei danni.

Va comunque sottolineato che gli indizi portano a dar ragione a Google, perlomeno in questo contesto. I *chatbot* sono sviluppati proprio per fingersi assennati, è il loro scopo, e lo stesso Lemoine ha ammesso che la sua posizione sia giustificata **«in veste di prete, non di scienziato»**. Considerando il suo background profondamente spirituale, risulta difficile trovare punti di ancoraggio che lo rendano un testimone attendibile sul piano tecnico.

Tecnici di Google annunciano che l'intelligenza artificiale è diventata  
senziente

Il problema è che la Big Tech **ha licenziato negli anni molti dei ricercatori etici** che hanno osato mettere becco nei piani dell'azienda, quindi è facile credere che, come denuncia il post internettiano, sviluppare un'indagine interna che soddisfi i requisiti imposti dalla dirigenza sia cosa perlomeno complessa. Al di là del fatto che la macchina abbia o meno sviluppato una coscienza, emerge inoltre un'altra criticità: se neppure il personale coinvolto nei test è in grado di distinguere l'IA da una persona, quante possibilità ci sono che lo possa fare un utente qualsiasi e cosa può comportare una simile ambiguità?

[di Walter Ferri]