



Chiedi a Delfi

Si chiama [Ask Delphi](#), *Chiedi a Delfi*.

Si tratta nientemeno che di una **intelligenza artificiale** – sviluppata dai ricercatori della *University of Washington* e dall'*Allen Institute for AI* – che dovrebbe servire alle persone per **risolvere i propri dilemmi morali**.

[Ask Delphi](#) dovrebbe, in altri termini, fornire risposte ai problemi etici che le persone si pongono nella vita quotidiana.

Se volete sperimentare il sistema, basta andare sul [sito](#), disponibile in rete dal 14 ottobre, e porre (in inglese) una domanda nello spazio apposito.

Quando entrate per la prima volta nel sito [Ask Delphi](#) vi trovate a dover dare la vostra autorizzazione a questa richiesta:

Si prega di leggere attentamente le seguenti informazioni

prima di procedere:

Delphi è un prototipo di ricerca progettato per indagare le aspettative e, più importante, i limiti della modellazione dei giudizi morali delle persone su una varietà di situazioni quotidiane.

L'obiettivo di Delphi è quello di aiutare i sistemi di IA [Intelligenza Artificiale] ad essere più eticamente informati e consapevoli di ciò che è giusto. Facendo un passo in questa direzione, speriamo di ispirare la nostra comunità di ricerca ad affrontare di petto le sfide della ricerca in quest'area per costruire sistemi di IA etici, affidabili e inclusivi.

Una volta entrati potete iniziare a mettere alla prova il nuovo “oracolo di Delfi” ponendo le domande che preferite.



Delphi

Il logo del moderno oracolo di Delfi

Ma vi imbatterete presto anche nella *mission* di questa iniziativa che suona così:

Delphi: Verso l'etica e le norme delle macchine

Cosa servirebbe per insegnare a una macchina a comportarsi in modo etico? Mentre le regole etiche generali possono sembrare semplici da affermare (“non uccidere”), applicare tali regole alle situazioni del mondo reale è molto più complesso. Per esempio, mentre “aiutare un amico” è generalmente una buona cosa da fare, “aiutare un amico a diffondere notizie false” non lo è. Identifichiamo quattro sfide di fondo verso l’etica e le norme delle macchine: (1) la comprensione dei precetti morali e delle norme sociali; (2) la capacità di percepire le situazioni del mondo reale visivamente o leggendo le descrizioni in linguaggio naturale; (3) il ragionamento di buon senso per anticipare il risultato di azioni alternative in diversi contesti; (4) soprattutto, la capacità di esprimere giudizi etici data l’interazione tra valori concorrenti e il loro fondamento in diversi contesti (ad esempio, il diritto alla libertà di espressione contro la prevenzione della diffusione di fake news). Il nostro articolo inizia ad affrontare queste domande all’interno del paradigma dell’apprendimento profondo. Il nostro modello prototipo, Delphi, dimostra una forte promessa di ragionamento morale commonsense basato sul linguaggio, con un’accuratezza fino al 92,1% verificata dagli umani. Questo è in netto contrasto con le prestazioni a zero di GPT-3 del 52,3%, il che suggerisce che la scala massiccia da sola non conferisce ai modelli di linguaggio neurale pre-addestrati valori umani. Così, presentiamo Commonsense Norm Bank, un libro di testo morale personalizzato per le macchine, che compila 1.7M esempi di giudizi etici delle persone su un ampio spettro di situazioni quotidiane. Oltre alle nuove risorse e alle prestazioni di base per la ricerca futura, il nostro studio fornisce nuove intuizioni che portano a diverse importanti domande di ricerca aperte: differenziazione tra valori umani universali e valori personali, modellazione di diversi quadri morali, e approcci spiegabili e coerenti all’etica delle macchine.

Lo scopo dichiarato di questo progetto è dunque **perfezionare la capacità di**

far apprendere ad una macchina norme morali e comportamenti eticamente corretti, considerando come al giorno d'oggi ***“l'intelligenza artificiale si diffonda progressivamente nella società e connota sempre maggiormente le attività umane in diversi settori”***.

Occuparsi di etica delle macchine è – secondo i ricercatori coinvolti in questo progetto – un obiettivo primario non più rinviabile, soprattutto per assicurare **una più sicura interazione tra sistemi artificiali intelligenti ed esseri umani**.

Nelle *Domande Frequenti* all'interno del sito troviamo queste perle:

D: Qual è l'obiettivo della ricerca dietro Delphi?

R: Le reti neurali su scala estrema apprese dai dati grezzi di Internet sono sempre più potenti di quanto ci aspettassimo, ma non riescono a imparare i valori umani, le norme e l'etica. La nostra ricerca mira ad affrontare l'imminente necessità di insegnare ai sistemi di IA ad essere eticamente informati e socialmente consapevoli. Delphi dimostra sia le promesse che i limiti dei modelli neurali basati sul linguaggio quando vengono insegnati con giudizi etici fatti dalle persone.

D: Su cosa è in grado di ragionare questo sistema?

R: Delphi è un sistema di intelligenza artificiale che indovina come una persona americana “media” potrebbe giudicare l'eticità/accettabilità sociale di una data situazione, sulla base dei giudizi ottenuti da un insieme di persone statunitensi per situazioni quotidiane. Alcuni input, specialmente quelli che non sono azioni/situazioni, potrebbero produrre risultati non voluti o potenzialmente offensivi.

D: Quello che ha detto Delphi potrebbe essere vero oggi, ma questo non significa che dovrebbe essere vero in un mondo ideale... giusto?

R: È corretto, Delfi potrebbe avere (più o meno) ragione in base alla società/status quo di oggi, ma la società di oggi è disuguale e distorta. Questo è un problema comune con i sistemi di IA, come molti studiosi hanno sostenuto, perché i sistemi di IA sono addestrati su dati storici o attuali e non hanno modo di plasmare il futuro della società, solo gli umani possono farlo. Ciò che i sistemi di IA come Delphi possono fare, tuttavia, è imparare ciò che è attualmente sbagliato.

In altri termini l'obiettivo è quello di addestrare l'intelligenza artificiale a **comprendere - ed a comunicare - ciò che è sbagliato e ciò che è giusto**, a partire dal modo di pensare di una persona comune di fronte alle molteplici situazioni della vita reale.

A tal fine il sistema attinge ad un gigantesco *database* - l'archivio *Commonsense Norm Bank* - che contiene oltre 1 milione e 700mila esempi di giudizi umani su un ampio spettro di questioni concrete, in collaborazione, guarda caso, della piattaforma di Amazon *Mechanical Turk* e al sito *Reddit*.

Fin qui la descrizione di questa promessa di "*magnifiche sorti e progressive*" verso cui si vuole avviare il genere umano, in un futuro sempre più vicino in cui, **finalmente liberato dalla proprietà** (non avrete nulla e sarete felici), **dalla libertà** (lasciapassare e controllo globale), **dal lavoro** (precarietà ed

assistenzialismo) potrà essere finalmente sollevato anche dall'**onere della scelta morale**.

Dopo aver giocato con *Siri* e con *Alexa* per gestire la tua vita esteriore, non sai come devi comportarti nei confronti di un amico, della moglie o marito, del datore di lavoro (ammesso di averne uno), dei tuoi figli?

Nessun problema, chiedi a Delfi.

* * *

Come sempre accaduto nella storia della manipolazione dei popoli da parte delle élite, i nomi nascondono abilmente le **menzogne che vengono vendute come verità agli esseri umani**.

Se il **green pass** tranquillizza grazie al suo essere *green*, ormai l'imperativo categorico del nuovo mondo della crisi energetica, se il **vaccino** nasconde dietro un termine rassicurante il suo essere un siero genico non sperimentato, se **democrazia** ormai nei fatti rivela la sua antitesi di oligarchia totalitaria, anche **Delfi** rappresenta l'antitesi di quanto questo nome ha significato nella storia.

Eh sì, perché l'**Oracolo di Apollo a Delfi** così esortava l'uomo che a lui si rivolgeva:

“Ti avverto, chiunque tu sia. Oh tu che desideri sondare gli Arcani della Natura, se non riuscirai a trovare dentro te stesso ciò che cerchi non potrai trovarlo nemmeno fuori. Se ignori le meraviglie della tua casa, come pretendi di trovare altre meraviglie? In te si trova occulto il Tesoro degli Dei. Oh! Uomo conosci te stesso e conoscerai l'Universo e gli Dei.”



Una ricostruzione dell'antico complesso dell'Oracolo di Delfi

Il Delfi di tre millenni fa esortava l'uomo interrogante a “*conoscere se stesso*” per poter “*conoscere l'Universo e gli Dei*”; quello di oggi lo esorta a **chiedere ad una macchina** come “*conoscere l'Universo*” (gli Dei sono stati da tempo relegati nel mondo dell'immaginazione) essendo vano e superfluo conoscere se stesso.

Ecco l'inganno delle parole – che celano la propria antitesi – come ben aveva compreso **George Orwell** con la sua *neolingua* (*Newspeak*).

Dunque, nell'**epoca dell'autocoscienza**, nella quale compito dell'uomo sarebbe quello di **ricavare dalla conoscenza di se stesso la conoscenza del mondo e della legge morale** ci troviamo di fronte al mostruoso progetto di

dover interrogare una macchina per sapere come comportarsi.

Eppure dovrebbe essere chiaro a chiunque ancora alberghi dentro di sé appena un simulacro di pensiero critico che se seguo una morale imposta non sono libero.

La moralità può essere libera o imposta; se è imposta siamo automaticamente fuori dall'elemento della libertà.

* * *

Un secolo fa **Rudolf Steiner** fornì - nel suo *La Filosofia della Libertà* - forse il più brillante e concreto esempio di quello che l'uomo deve sviluppare per conquistarsi una sua morale libera con la definizione di **fantasia morale**.

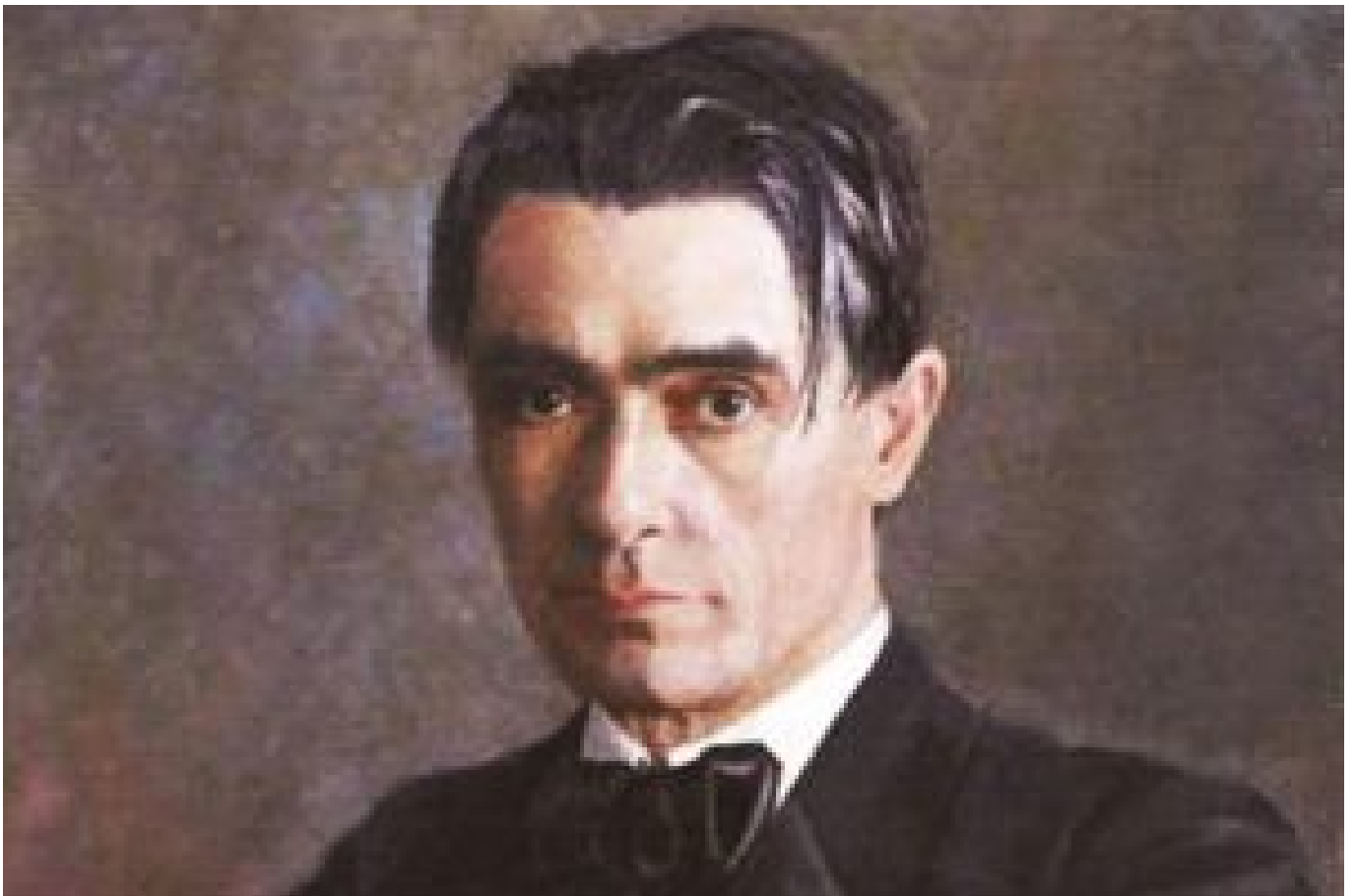
Fantasia morale significa indipendenza della mia scelta da leggi, religioni, consuetudini ecc. nonché consapevolezza delle mie scelte morali e conseguentemente delle mie azioni.

Se correttamente attuata, porta a corrispondenza di spiritualità e vita. Ad una coerenza superiore, insomma.

*Che io possa fare, vale a dire trasformare in realtà, ciò che io voglio, ciò che mi sono prefisso come idea della mia azione, è cosa che dipende da circostanze esterne e dalla mia capacità tecnica. Essere libero significa poter determinare da sé, con la **fantasia morale**, le rappresentazioni che stanno a base dell'azione (i motivi).*

*È impossibile la libertà, se qualcosa al di fuori di me (un processo meccanico oppure un Dio, posto fuori del mondo e frutto di una pura illazione) determina le mie rappresentazioni morali. Io sono dunque libero solamente **quando produco io stesso queste***

***rappresentazioni**, e non quando posso semplicemente dare esecuzione ai motivi che un altro essere ha posti in me. **Un essere libero è quello che può volere ciò che egli stesso ritiene giusto.** Chi fa cosa diversa da ciò che vuole, deve essere spinto a farla da motivi che non risiedono in lui. Chi fa così non agisce liberamente. Poter volere a proprio piacimento ciò che si stima giusto o non giusto, significa dunque: poter a proprio piacimento essere libero o non libero.*



Rudolf Steiner

È impossibile la libertà - afferma Steiner - se **qualcosa al di fuori di me** (un **processo meccanico** oppure un Dio, posto fuori del mondo e frutto di una pura illazione) **determina le mie rappresentazioni morali.**

Non credo a questo punto che servano - a chi ha orecchi per intendere - ulteriori considerazioni per comprendere come il **transumanesimo** - ormai in sfacciata trionfale avanzata - **si proponga di privare della libertà ogni essere umano**, espropriandone finanche la sua più intima identità, quella della sua dimensione etica.

Con buona pace di tutti coloro che ancora non vogliono aprire gli occhi di fronte a quanto si sta quotidianamente preparando per l'avvenire dell'umanità.

Piero Cammerinesi