

La notizia è di qualche giorno fa. Un ingegnere di Google, che lavorava a LaMda (acronimo di Language Model for Dialogue Applications), una tecnologia che consente di avviare conversazioni complesse ed usata per far funzionare gli assistenti vocali come “Alexa”, ha detto che ormai l'intelligenza artificiale è diventata senziente e, quindi, umana. Non sappiamo se le cose stiano realmente così. Da tempo, però, in tutto il mondo si discute di questa possibilità. E molte organizzazioni internazionali, a cominciare dall'Unione europea, stanno redigendo i principi etici che i ricercatori devono rispettare nel progettare macchine sempre più “intelligenti”. Un punto su cui si insiste molto, in queste pioneristiche Carte di principi sull'intelligenza artificiale, è che i robot dovranno adeguarsi ai valori umani. Ma quali sono i valori di noi umani (intesi come specie planetaria)? Nei giorni scorsi, la Corte Suprema degli Stati Uniti, cancellando la sentenza che dava copertura costituzionale federale al diritto all'interruzione volontaria della gravidanza, ha riaperto il dibattito tra chi crede che la vita del nascituro sia un valore da tutelare sin dal momento del concepimento e chi, viceversa, crede che il valore prevalente, (almeno nelle prime settimane di gestazione) sia quello delle donne alla propria autodeterminazione, o magari, più semplicemente, la considerazione di buon senso che nessuno può essere obbligato a fare qualcosa contro la propria volontà.

Intelligenza artificiale

Macchine senzienti: con quali valori?

GIOVANNI PASCUZZI

Ripetuti fatti di cronaca ripropongono la querelle tra chi sostiene che nessuno ha il diritto di disporre della propria vita (perché la vita è un valore in sé) e chi ritiene che lo Stato debba offrire gli strumenti per porre fine ai propri giorni quando l'esistenza diventa insopportabile a causa di gravi patologie. Nei mesi scorsi aveva assunto toni anche violenti la discussione tra chi sosteneva che per salvaguardare il diritto alla salute della collettività dovesse introdursi per tutti l'obbligo vaccinale contro il Covid e chi rivendicava il diritto a poter scegliere se vaccinarsi o no. La siccità di questi giorni ripropone lo scontro tra chi denuncia i rischi del riscaldamento globale (in nome della necessità di tutelare l'ambiente) e chi vuole continuare ad emettere nell'atmosfera agenti inquinanti in nome del diritto dei paesi emergenti a raggiungere un maggior benessere economico. Una guerra che si trascina ormai da mesi nel cuore dell'Europa ripropone addirittura un terreno di scontro (nel vero senso della parola) che sembrava definitivamente

consegnato al passato: quello tra chi ritiene di avere il diritto di imporre con le armi la propria visione (che si tratti di una contesa territoriale o di cose più importanti poco rileva) e chi ritiene che l'agredito debba essere comunque difeso. Gli esempi potrebbero continuare. Ma sono già più che sufficienti a dimostrare che viviamo in un tempo di polarizzazioni molto spinte delle posizioni. Si fronteggiano valori diversi se non contrapposti e sembra una chimera immaginare una possibile sintesi anche a causa della facilità con la quale si ricorre alla violenza: basti pensare ai primi giorni del 2021 quando un assalto armato al Parlamento degli Stati Uniti fu l'epilogo del rifiuto di riconoscere l'esito di una consultazione elettorale. L'idea che le macchine possano diventare sempre più intelligenti se non addirittura senzienti ci mette un po' i brividi lungo la schiena. Ecco che nasce impellente la necessità di imporre che esse non sfuggano mai al controllo umano. Di qui almeno l'auspicio che i robot siano progettati affinché facciano propri i valori di noi umani.

Si arriva così al nocciolo dei problemi: quello dei fini. Prima che la situazione ci sfugga di mano è indispensabile fare in modo che l'intelligenza artificiale: a) comprenda i nostri fini; b) adotti i nostri fini; c) conservi i nostri fini. Questa operazione viene definita: allineamento dei fini. Ed ecco il punto. Quali sono davvero i nostri fini, i valori in cui crediamo? Abbiamo come italiani, come europei, come terrestri, come umani dei valori davvero comuni da poter insegnare ai cosiddetti “postumani” affinché li custodiscano e diano loro attuazione quotidiana? E se la stessa diversità valoriale fosse un valore, esiste ancora un margine per il rispetto reciproco o l'astratta possibilità di immaginare visione e sintesi? Personalmente più che le macchine senzienti mi crea inquietudine la babele valoriale che potrebbe spingere qualcuno a pensare che i robot debbano essere usati non già al servizio di tutti ma di una parte (i droni intelligenti impiegati come arma già rispondono a questa logica). Qualche scienziato ha ipotizzato che il rischio vero non sia quello di ritrovarsi una intelligenza artificiale “cattiva”, bensì una intelligenza artificiale estremamente brava nel realizzare propri fini non coincidenti con i nostri quali essi siano: una intelligenza artificiale che pensi solo a se stessa, che forse è il vero insegnamento che oggi stiamo dando (anche) ai robot.

Giovanni Pascuzzi
Consigliere di Stato