

Il dubbio che serve alla Ai

intervista a Paolo Benanti, a cura di Domenico Agasso

in "La Stampa" del 3 gennaio 2024

Il teologo: "L'intelligenza artificiale sarebbe più umana se sapesse dubitare dei suoi stessi responsi. L'etica è tutto ciò che non è esigibile per legge, quindi si può insegnare solo a chi possiede una soggettività".

«La misericordia e l'uguaglianza, ma prima di tutto nei robot sarà decisivo il dubbio. Se riusciremo a inserire in ChatGPT la capacità di dubitare delle proprie risposte avremo realizzato qualcosa che rispetterà molto di più la natura di donne e uomini». Sono i principali valori morali da trasmettere all'Intelligenza artificiale (Ai) affinché «resti umana», secondo l'appello del presidente della Repubblica Sergio Mattarella. Parola del teologo francescano Paolo Benanti, membro – unico italiano – dell'Organo consultivo delle Nazioni Unite per l'Ai. 50 anni, il Religioso del Terz'Ordine Regolare è docente di Etica e Bioetica, Etica della Tecnologia e Artificial Intelligence.

Siamo all'alba di un mondo segnato dall'Ai. Padre Benanti, è possibile infondere un'educazione agli algoritmi?

«Bisogna premettere che l'Ai non può essere educata moralmente, perché non è una soggettività personale. I valori su cui si basa sono numerici. Però possono essere in qualche misura adeguati, controllati da algoritmi utilizzabili come una sorta di Guard rail etica. Diventa perciò fondamentale una nuova e rinnovata educazione morale di quegli utenti, cioè dei cittadini, che non solo si interfaceranno con l'Ai, ma che daranno all'Ai i limiti o i campi applicativi. Quindi è urgente alzare il livello di insegnamento etico. Dicendo questo però potrebbe sorgere un'ambiguità».

Quale?

«Abbiamo conosciuto in passato forme di Stato etico che tutto erano tranne che etiche. Allora vorrei chiarire che cosa si intende per morale in questo senso».

Ci spieghi.

«Un giudice inglese molto autorevole, Lord Moulton, in un discorso a fine carriera ha affermato che l'etica è tutto ciò che non è esigibile per legge. Sul Titanic che stava affondando qualcuno ha detto "prima le donne e i bambini", e hanno mandato prima le signore e i piccoli: quello era lo spazio dell'etica, lo spazio di civiltà. Ecco, noi dobbiamo edificare un nuovo strato di civiltà che abbia la funzione di cuscinetto etico: così si potrà gestire l'Ai nel migliore dei modi possibile».

Che cosa intende?

«Dare priorità all'essere umano. Tra l'obiettivo della macchina e lo scopo dell'umano, il progetto dell'umano deve venire prima. L'Ai deve dare la possibilità all'uomo di poter definire le sue priorità e quindi la sua autodeterminazione sociale».

Se l'Ai fosse un bambino, come lo si potrebbe istruire?

«Questo bimbo non ha tutte le abilità dell'adulto, allo stesso tempo ha una velocità eccezionale e non si stanca mai. Di fronte a una macchina inesauribile, e che lavora con una rapidità straordinaria nell'esecuzione del suo compito, ci deve essere vigilanza. Ray Kurzweil sosteneva che una macchina programmata per costruire graffette per i fogli - missione apparentemente innocua - senza mai fermarsi potrebbe portare alla fine del pianeta».

Immaginando l'Ai come un cervello nuovo, quali valori vi introdurrebbe per il bene dell'umanità, delle società odierne e del futuro?

«Dopo la nascita il cervello di noi umani si sviluppa nei primi mesi e anni, e questo fa sì che le relazioni che viviamo con le persone che ci accudiscono diano l'imprinting dei valori. Li assorbiamo così. Il bene che ci hanno voluto la nostra mamma e il nostro papà e altri familiari ci fanno apprendere la reciprocità dei comportamenti generosi, solidali, responsabili, educativi. Invece il "nuovo cervello" delle Ai non è aperto a quello che vivrà, ma è strutturato su quello che qualcun altro ha vissuto nel passato. Perciò secondo me i valori più importanti da collocarvi sono tre su tutti».

Il primo?

«Innanzitutto il dubbio. Detto così non sembra un valore morale. Ma se noi chiediamo qualcosa a ChatGPT una risposta ce la dà sempre. E a volte si tratta di allucinazioni. In questo momento la "virtù" più grande a cui possiamo aspirare noi uomini e donne non è la capacità di rispondere, ma di interrogarci se la risposta è valida o no. Se riuscissimo a dare alla macchina la predisposizione a dubitare dei suoi responsi avremmo fatto qualcosa che rispetterebbe molto di più la natura umana».

Ci dica il secondo.

«La misericordia, cioè la capacità di non essere determinati solo dagli avvenimenti negativi che una persona ha subito e registrato nella sua storia. Non si può prescindere dalla benevolenza».

Il terzo?

«Nell'ambito più sociale è cruciale l'aspetto dell'equità e dell'uguaglianza. Perché altrimenti le Ai alimenteranno la pericolosa divisione in classi sociali attraverso i pregiudizi».

E Dio che ruolo avrà?

«Il tema non sarà il rapporto tra Dio e l'Ai, ma sempre l'atteggiamento di noi uomini con l'Ai. Perché noi abbiamo la tendenza a crearci degli idoli, e potremo avere persone che si rivolgeranno alle Ai in forma oracolare, cioè come se fossero degli oracoli a cui chiedere delle divinazioni, di risolvere i nostri dilemmi esistenziali in una maniera che diventerebbe falso divino. E allora la questione sarà se l'uomo cercherà nell'Ai un dio che può controllare, una specie di bacchetta magica alla Harry Potter che sistema ciò che non va. In pratica il rischio è che si guardi al Cloud, dove risiede l'Ai, invece che al Cielo, e che si considerino le Ai delle nuove divinità».