

L'ETÀ DIGITALE

# Il robot non è infallibile, siamo impreparati al trionfo delle macchine

Lo sviluppo impetuoso della cibernetica nella società rende sempre più urgente una nuova etica universale

PAOLO BENANTI\*

**L**e intelligenze artificiali stanno cambiando il mondo: ogni attività umana, dalla medicina alla sicurezza nazionale stanno subendo profonde trasformazioni. I sistemi dotati di AI non solo aiutano l'uomo ma in sempre maggiori situazioni danno luogo a sistemi, bot oppure robot, completamente autonomi.

Di fronte a questo diluvio di intelligenza artificiale è urgente la questione etica. Più le AI diventano universali più è necessario sviluppare un nuovo linguaggio universale che sappia gestire l'innovazione. Le intelligenze artificiali sono artefatti tecnologici. Ma differenti da tutti gli artefatti prodotti fino ad oggi. Tutti gli strumenti che abbiamo prodotto consentono all'uomo di svolgere alcuni compiti. Dalle clave primitive fino alle grandi macchine industriali tutti questi strumenti servivano a fare meglio più velocemente in maniera più efficace dei compiti precisi. Le AI, tanto nei bot quanto nei robot, superano il concetto di artefatto e di macchina che conoscevano fino adesso. Tutti i meccanismi automatici che abbiamo costruito durante la rivoluzione industriale sono stati costruiti pensando a quale sarebbe stato il loro scopo. Facevano quello per cui erano progettati e basta. Oggi le AI non sono progettate così. Non sono software programmati ma sistemi addestrati. Si supera

il modello classico *if this then that* in cui un ingegnere del software prevedeva prima tutte le possibili occorrenze.

Le AI rispondono in maniera autonoma a un problema che gli viene posto. Questi artefatti sono una nuova specie nelle macchine. Delle *machine sapiens*. Oggi il mondo non è più abitato solo dall'*homo sapiens* ma anche da *machine sapiens*. Ma come decide la *machina sapiens*?

I *data scientist* ci dicono che il problema è alla qualità e alla quantità dei dati. Quando avremo un database perfetto su cui far girare i nostri servizi di AI la macchina farà scelte perfette. Ma è così? Già in passato abbiamo avuto questa impressione. Laplace sosteneva che qualora noi avessimo conosciuto la posizione in un istante di tutte le particelle che contengono l'universo saremmo stati in grado di predire tutto il futuro e di conoscere tutto il passato dell'universo. Lasciamo stare la domanda filosofica su questa possibilità e affrontiamola da un punto di vista operativo. Se fossimo in grado di creare una mappa che è l'esatta copia della realtà, includendo al suo interno tutto, compresi i passanti, le foglie degli alberi, ecc, dovremmo riconoscere che la mappa che abbiamo creato è inutile. Questa infatti sarebbe complessa come la realtà, troppo complessa per prendere le

## Si rischiano danni alle persone e un aumento degli squilibri globali

decisioni e quindi inutili.

Ci troveremmo cioè di fronte al noto paradosso raccontato da Jorge Luis Borges in un frammento *Del rigore della scienza*, l'ultimo di *Storia universale dell'infamia* pubblicato per la prima volta nel 1935. Come sua abitudine, l'autore argentino attribuisce la citazione a un libro che in realtà non esiste:

«... In quell'Impero, l'Arte della Cartografia giunse a una tal Perfezione che la Mappa di una sola Provincia occupava tutta una Città, e la mappa dell'impero tutta una Provincia. Col tempo, queste Mappe smisurate non bastarono più. I Collegi dei Cartografi fecero una Mappa dell'Impero che aveva l'Immensità dell'Impero e coincideva perfettamente con esso. (Suárez Miranda, *Viajes de Varones Prudentes*, libro IV, cap. XIV, Lérida, 1658)». (Dall'edizione italiana de Il Saggiatore, 1961).

I dati sono una mappa della realtà, rappresentano una riduzione della realtà e per questo sono utili per prendere decisioni. Inoltre, le AI lavorano sui database e sui sensori. Ma anche i sensori non leggono tutta la realtà: ne prendono solo una parte trasformandola in dati. Eccoci al punto chiave della questione. Siccome le intelli-

genze artificiali fondano loro decisioni sui dati e poiché questi non sono una coppia perfetta della realtà. La macchina *sapiens* sarà sempre e costitutivamente. Il robot non è infallibile. Le AI hanno costitutivamente bisogno di un'etica. Poiché le intelligenze artificiali possono sbagliare bisogna capire come gestire allora questo sbaglio. La questione etica è fondamentale importantissima è urgente. Bisogna trovare un sistema etico condiviso perché l'utilizzo di questi sistemi non produca ingiustizie, non danneggi le persone e non crei dei forti disequilibri globali.

L'esistenza di macchine *sapiens* chiede di mettere in piedi un nuovo linguaggio universale che sappia tradurre queste direttrici etiche in delle direttive eseguibili dalla macchina. Ma come fare questo? Il mondo nell'epoca del *Digital Age* è regolato degli algoritmi. Più di qualcuno parla di una algo-crazia. Per evitare che ci sia questo dominio dell'algoritmo anche grazie alle AI dobbiamo iniziare a sviluppare questo linguaggio comune dell'algor-etica.

Il problema è innanzitutto filosofico ed epistemologico. Le AI «funzionano» secondo schemi che connettono dati. Che tipo di conoscenza è questa? Che valore ha? Come va trattata e considerata? Insomma, la domanda prima che tecnologica è etica e filosofica: nella misura in cui vogliamo affidare competenze umane, di comprensione, di giudizio e di autonomia di azione a dei sistemi software di AI dobbiamo capire il valore,

in termini di conoscenza e capacità di azione, di questi sistemi che pretendono di essere intelligenti e cognitivi.

\* Paolo Benanti, francescano del Terzo Ordine Regolare, teologo, docente, si occupa di bioetica e etica delle tecnologie —



Paolo Benanti  
«Digital Age»  
San Paolo  
pp. 216, €18

