

Rassegna

Alla ricerca di una bussola etica per le intelligenze artificiali

Leonardo De Chirico

La saggistica sull'etica dell'intelligenza artificiale è diventata un genere letterario in continua espansione. Sempre più pubblicazioni cercano di mettere a fuoco le implicazioni morali dello sviluppo delle tecnologie che affidano alle macchine la responsabilità di prendere decisioni che impattano la vita umana e dell'ambiente. Questa rassegna prende in considerazione una piccola rappresentanza di questa riflessione per testare il livello del discorso e le questioni ricorrenti.

PAOLO BENANTI, *Le macchine sapienti. Intelligenze artificiali e decisioni umane*, Bologna, Marietti1820 2018, pp. 157.

Benanti è docente di teologia morale alla Gregoriana ed è diventato una figura anche televisivamente conosciuta sui temi dell'intelligenza artificiale sui quali sta producendo una notevole quantità di pubblicazioni. Questo volume è scritto in modo agile e ha un taglio divulgativo. In quattro capitoli e due interludi tratteggia una mappa dei temi sul tappeto e offre spunti di approfondimento.

L'A. parte dalla constatazione che siamo arrivati ad un punto in cui è "la macchina ad interpellarci". Eravamo abituati ad avere un ruolo attivo e asimmetrico sulla macchina, ma dobbiamo abituarci ad avere la macchina fare con noi quello che facevamo con essa. Alcuni esempi sono già di uso quotidiano. Nel campo del lavoro, ad esempio, le intelligenze artificiali sono diventate pervasive: tutti gli smartphone sono venduti con un assistente "intelligente": Cortana, Siri o Google Hello, che diventa un vero

e proprio partner personale. Per quanto riguarda la salute, l'applicazione Ginger.io registra in maniera automatica e continuativa i parametri vitali, segnalando i comportamenti "non conformi". In ambito sociale, la Cina ha introdotto un sistema di "social credit" che assegna un punteggio ai cittadini in base alla fiducia meritata (il "citizen score"). Sulla base di queste rilevazioni, lo Stato poi riscontra con ricompense e privilegi. Si tratta di un metodo di sorveglianza sociale che si presenta sotto forma di un quasi gioco. Sono algoritmi predittivi che determinano se siamo buoni o cattivi cittadini. Siamo arrivati alla situazione in cui "quello che facciamo on line e off line si fonde con la nostra vita fisica" (p. 41): riprendendo un'espressione di Floridi, stiamo acquisendo "una personalità onlife: un mix tra on line e life".

Fatta questo sorvolo sui cambiamenti in corso, la domanda che, a questo punto, si pone è: "cosa significa essere umani?" Gira e rigira, si passa e si torna sempre lì. Le intelligenze artificiali interrogano il senso dell'umanità. Giustamente, Benanti sostiene che prima che essere un problema tecnologico, esse "rappresentano una problema epistemologico e filosofico" (p. 68). In particolare, visto che le macchine imparano un linguaggio quindi trattano delle informazioni, cosa significa che ascoltino e parlino sulla base di credenze orientate da dati? Vista la centralità dello scambio e del trattamento di informazioni che le macchine elaborano e su cui agiscono, l'interrogativo metafisico "ruota intorno all'informazione" (p. 94). Sulla scia di Harari (che ne parla nel libro *Homo Deus*)¹, l'A. suggerisce un neologismo: "dataismo", per indicare il movimento di cui fanno parte anche le intelligenze artificiali in grado di connettere i dati incessantemente fluttuanti tra loro. Nel dataismo il principio di legittimità è dato dall'uso di algoritmi e dai Big Data. Esso è quindi "un sistema di elaborazione dati onnicomprensivo" (p. 98) che nel tempo si fonderà col mondo stesso. È evidente che il rischio per gli esseri umani è di essere ridotti alla stregua di chip inseriti in un sistema non più umano.

1. YUVAL NOAH HARARI, *Homo Deus. Breve storia del futuro*, Milano, Bompiani 2017.

32

Possono le macchine essere considerate delle entità emotivo-razionali in analogia agli esseri umani se non proprio in contiguità con essi? Si tratta di avversari evolutivi o di cooperatori della persona? Si vanno a sostituire alla persona o la loro ragion d'essere è l'aumento della capacità cognitiva? Se si comincia a riflettere sulla moralità delle macchine artificiali, le domande si moltiplicano: hanno una volontà libera? Sono in grado di anticipare ed assecondare le intenzioni? Quanto trasparente ed efficace è la loro intelligibilità? Quanto sono adattabili? Il punto decisivo per l'A. è il mantenimento della proprietà operativa non nell'algoritmo, ma nella persona "che è luogo e sede di dignità" (p. 123). L'assunto ontologico è che "è il robot che coopera con l'uomo e non l'uomo che assiste la macchina" (ivi). Da ciò ne può derivare un codice etico per le intelligenze artificiali che dia luogo ad una loro governance che tenga conto dei problemi di cyber-security, di rendicontazione, del continuo discernimento etico necessario ad ogni passaggio. In conclusione, si tratta per l'A. di informare e orientare una "cultura organizzativa" che tenga insieme tutti questi elementi: una cultura che plasmi e guidi l'innovazione tecnologica. Viene suggerita anche la crescente necessità di un "ethics manager" delle intelligenze artificiali (p. 149) nelle aziende che ne fanno uso, oltre al bisogno di far crescere la responsabilità diffusa.

GIANNI MANZONE, *Morale artificiale: nanotecnologie, intelligenza artificiale, robot. Sfide e promesse*, Bologna, [EDB](#) 2020, pp. 244.

Il volume di Manzone, docente di teologia morale alla Lateranense, copre un campo più vasto e include il ventaglio di questioni morali che sorgono dall'impiego di tecnologie avanzate in settori quali le nanotecnologie e la robotica. Anche Manzone ha alle spalle un'ampia produzione sulla Dottrina sociale della Chiesa cattolica e su temi di teologia morale, soprattutto sul crinale tra tecnica/tecnologia ed etica. Per l'A., la questione chiave è se la tecnologia "faciliti piuttosto che frustrare l'avventura umana" (p.

7). Siccome può fare entrambi, qui sorge l'interrogazione morale sulle sue finalità.

L'A. riporta i termini del dibattito attuale sulla possibilità stessa di una nanoetica e sul suo statuto: è una branca specifica o un'altra applicazione della etica, tra le tante? Quali sono le implicazioni sociopolitiche e le responsabilità di regolazione, soprattutto per quanto riguarda il governo dei rischi e la protezione della privacy? Un altro capitolo tratta il tema spinoso del potenziamento della condizione umana. L'A. utilmente distingue la terapia dell'enhancement e, anche in questo caso, è il bene umano ad essere toccato. Come misurarlo? Bilanciando istanze di giustizia ed equità, autonomia e libertà, oltre a tenere in considerazione le esigenze della salute e della sicurezza. Sembrerebbe una griglia procedurale sulla linea della classica etica dei principi; l'A. le intreccia al suo impianto tomista finalisticamente legato al "bene" dell'umano. Anche in questo caso, segnala la necessità di una governance anticipatoria più che ad una solamente reattiva agli sviluppi tecnologici. In ogni caso, l'A. fa propria una visione non ansiogena ed allarmata: non esiste nessun determinismo tecnologico che rende ineluttabile l'oltrepassamento transumanista, (trattato nel cap. 7) anche se occorre sempre vigilare sulle suggestioni del "potenziamento radicale" che è portatore di una ideologia forte di superamento dell'umanità, di super-umanità.

Il cap. 5 tratta dell'intelligenza artificiale (AI) in modo specifico. In molti dei suoi sostenitori, essa viene vista come estensione delle capacità della mente. Vero è che questa concezione di mente è funzionalista e non sembra riconoscere il proprium umano che può essere riassunto nella coscienza. La mente è considerata una macchina computazionale che può essere riprodotta, aumentata, esternalizzata. Mentre la tecnologia può riprodurre funzioni della mente, può sostituire la mente? Eseguire un programma in modo veloce, anche in presenza di algoritmi sofisticati, non vuol dire sostituirsi alla mente umana. In gioco ci sono i fondamenti identificativi della persona. Sul piano morale, "molta ricerca sull'AI è completamente pragmatica, senza alcun riferimento alle strutture

34 morali nell'architettura cognitiva" (p. 113). Ci sono molte "zone di non conoscenza e non verificate, con eventuali conseguenze non prevedibili" (p. 115). Manzone richiama le tre leggi della robotica dello scrittore di fantascienza Isaac Asimov: 1. Un robot non può creare danno agli esseri umani; 2. Un robot deve sempre obbedire agli esseri umani a meno che ciò non contrasti con la prima legge; 3. Un robot deve proteggere la propria esistenza a meno che ciò non contrasti con la prima o la seconda legge (p. 121). Risultano veramente efficaci? Chi può pretenderne l'applicazione? Come tradurle in protocolli universali? C'è poi tutto il campo dell'etica dei robot virtuali, quelli che operano nella infosfera. Temi quali la protezione della privacy ricorrono qui. L'A. richiama la necessità di un "modello cooperativo" tra diversi tipi di intelligenza (p. 124). L'integrazione tra essere umani e apparati tecnologici sarà sempre più massiccia e la responsabilità di pensare alla qualità della relazione è un dovere ineludibile in vista di una governance trasparente e globale.

Nell'ultimo capitolo, l'A. affronta il ruolo della teologia nella cultura tecnologica. Nell'ambito del dibattito interdisciplinare richiesto dalla riflessione etica, la teologia contribuisce in modo particolare nel sottolineare temi quali la dignità umana, la visione del limite, la consapevolezza apocalittica del progresso e la tensione alla trascendenza del dato biologico. Mentre è interessante notare la validità della *imago Dei* come "utile benchmark per pensare il futuro della natura umana nel mondo tecnologico" (p. 203), non altrettanto presente è il tema del peccato e delle sue conseguenze sulla sfera etica. Certamente, l'A. parla di rischi e problemi connessi all'uso della tecnologia, ma la teologia non può dire di più circa la deviazione morale che è figlia della ribellione a Dio e che ha spinto la direzione della storia verso il parossismo etico? È anche vero che Manzone parla del contributo teologico alla nozione di limite, ma anche qui la teologia cristiana del peccato non traspare dal libro. È presente il limite della creazione/creatura, ma è assente il tema della rottura del peccato e del suo impatto dirompente. L'etica cristiana invita alla responsabilità

perché, oltre a conoscere i limiti insiti nella costituzione delle persone, sa anche della nostra insaziabile hubris introdotta dal peccato che trova nella tecnologia un altro campo in cui manifestare la sua irrequietezza.

35

TIMO PLUTSCHINSKI (ed.), *The 4th Industrial Revolution from an Ethical Perspective*, Amersfoort, Sallux 2021, pp. 113.

Il volume è frutto del lavoro della Business Coalition dell'Alleanza Evangelica Mondiale (WEA) che ha messo insieme studiosi di diverse competenze e provenienze per analizzare le sfide etiche della "quarta rivoluzione". È un bell'esempio di un servizio reso da un organismo come l'Alleanza evangelica per promuovere un confronto tra esperti evangelici di varie provenienze regionali e fornire piste di riflessione allargate. Si compone di una breve introduzione e di sei capitoli che trattano le sfide etiche dell'AI, le relazioni tra AI e l'agenzia umana, le trasformazioni dell'AI nel mondo dei media, la chiesa e le nuove tecnologie e l'impatto della quarta rivoluzione sulla vita personale e delle famiglie.

Di particolare interesse sono le cinque tesi proposte da Klaus Henning (pp. 29-44), tanto da meritare una citazione estesa:

1. L'AI sta pervadendo tutte le macchine, i sistemi e i dispositivi, tutte le occupazioni e le vite private. I compagni digitali saranno onnipresenti senza dare nell'occhio. Sono estremamente utile e questa è la ragione perché li usiamo.
2. In un mondo pienamente connesso e digitalizzato, abbiamo bisogno di nuovi modi di bilanciare le tensioni tra le persone, le culture, le tecnologie, le realtà virtuali e l'AI con la nostra coscienza.
3. Le piattaforme digitali sono la chiave del mercato globale di beni e servizi.
4. Molte professioni spariranno e di completamente nuove ne sorgeranno, a tutti i livelli di competenza.
5. La più grande rivoluzione culturale da Gutenberg sta

36 prendendo piede grazie all'AI, un vero e proprio "game-changer". Abbiamo la possibilità di usare l'AI per governare la trasformazione digitale in un modo responsabile prima che altri lo facciano in modo irresponsabile.

Alla luce di queste considerazioni, l'età dell'ibridizzazione sta sorgendo e chiede un rinnovato sussulto di riflessione morale. È possibile una partnership proficua tra esseri umani e macchine intelligenti? La speranza sottesa ai vari interventi è possibilista, ma le sfide da affrontare sono dall'esito tutt'altro che scontato. Il tono del volume è descrittivo e si capisce che i vari autori hanno una certa competenza tecnica in diversi settori coinvolti. Quanto la loro riflessione sia informata da un'etica propriamente evangelica è un discorso a parte. In ogni caso, è un segnale di vitalità culturale non trascurabile.