

Salvare il ciclista

con il casco o quello senza?

di Franco Marra

Guglielmo Tamburrini

ETICA DELLE MACCHINE
DILEMMI MORALI PER ROBOTICA
E INTELLIGENZA ARTIFICIALEpp. 151, € 13,
Carocci, Roma 2020

L'autonomia crescente delle macchine pone temi etici importanti e oggetto di continuo dibattito. Significativo in merito questo contributo di Guglielmo Tamburrini. Secondo l'autore, il problema dell'etica nelle macchine si scontra con uno scoglio allo stato dell'arte tuttora insuperabile: l'assenza della coscienza di sé. Malgrado siano disponibili meccanismi che consentono a livelli crescenti decisioni autonome, la responsabilità di ciò che una macchina fa non può esserle in alcun modo attribuita, perché l'incapacità di agire in base a convinzioni morali sentite come proprie non le consente di assumere il ruolo di agente morale. L'uomo quindi non può abdicare al compito etico, ma deve mantenere nel *loop* (come direbbe un ingegnere dei controlli automatici) capacità adeguate e calibrate di controllo e responsabilità: livelli opportuni di Controllo Umano Significativo (CUS). Questa conclusione discende da una analisi dei problemi etici indotti dall'autonomia delle macchine condotta empiricamente da Tamburrini su una serie di casi di studio significativi, acclarata l'impossibilità di dedurre da principi generali un insieme formale di norme coerenti codificabili in macchina.

Come sa bene Valeria, la giovane capo-progetto dell'informatica di bordo di un veicolo a guida autonoma (VA), fortemente motivata nel suo lavoro dalla prospettiva di contribuire alla diminuzione delle vittime di incidenti stradali ma frustrata dalle continue contraddizioni che emergono nelle discussioni del gruppo di lavoro. In caso di un incidente inevitabile, nella drammatica scelta di chi sacrificare tra il ciclista con il caschetto e quello senza e di fronte alla ragionevole prospettiva di penalizzare quello protetto perché si suppone corra meno rischi, come rispondere all'obiezione che alla lunga questa decisione codificata *ex ante* nel software del VA avrà la conseguenza di disincentivare i ciclisti all'uso del casco, aumentando così il numero finale di vittime della strada? Oppure, come convincere una mamma come lei ad acquistare un VA quando lei stessa o il bimbo seduto accanto sono potenziali vittime al posto dell'ignaro pedone che passeggia sul bordo della strada, come ipotizzato dalle linee guida della commissione critica del ministero dei Trasporti e dell'Infrastruttura digitale della Repubblica federale tedesca nel 2016, per il solo fatto che quest'ultimo "non è parte coinvolta nella generazione di rischi da mobilità"? Contraddizioni implicate nel confronto tra l'etica delle conseguenze, che spinge a scelte che privilegiano futuri benefici accettando compromessi nel presente, e l'etica dei valori, che proclama l'assoluta superiorità dei valori non negoziabili indipendentemente dagli esiti futuri.

Ma ecco altri casi di studio esemplari utilizzati dall'autore a sostegno

della sua tesi: gli algoritmi che governano le transazioni finanziarie a una velocità al di là di ogni possibile controllo umano causano crisi finanziarie inattese, mentre quelli che stimano la reputazione e l'affidabilità degli individui nei processi giudiziari e nel credito finanziario commettono grossolane ingiustizie. Soggetti a *bias* cognitivi dovuti all'addestramento o a pregiudizi inseriti magari inconsapevolmente nel codice dai loro progettisti, possono compromettere talvolta in modo irreparabile la vita di chi è sottoposto all'analisi. Gli algoritmi sociali di classificazione di massa utilizzati in occidente soprattutto a fini marketing arrivano in Cina a condizionare la vita di intere popolazioni, penalizzando chi non si comporta secondo i dettami di un pensiero medio unico, standard di riferimento, e discriminando minoranze etniche e religiose. Su tutti gravano poi asimmetrie informative e di potere che giocano sempre a favore del sistema politico o economico consolidato e che aggravano le differenze sociali. In scenari nei quali il problema del lavoro umano e della redistribuzione della ricchezza non compare certo tra le priorità, anche se ancora non si sono avverati i foschi scenari profetizzati da Alan Turing e Norbert Wiener.

In guerra, le armi autonome sono il sogno di ogni comandante sul campo: ma in azione dovrebbero aderire a principi di discriminazione del bersaglio, di proporzionalità del danno inflitto, di mantenimento della catena delle responsabilità e di umanità nei confronti del nemico reso innocuo, principi dettati dagli accordi internazionali oltre che da comuni sentimenti di umanità. Ma come possono decidere in modo corretto questi sistemi autonomi quando i loro apparati di percezione e classificazione possono essere disastrosamente ingannati da disturbi non percepibili dall'uomo? Al punto da far loro confondere scuolabus con struzzi e tartarughe con fucili? E cosa resta della dignità umana di chi sta per essere ucciso da una macchina freddamente estranea ad ogni forma di empatia e di comunanza di valori? Insomma, l'uomo deve stare nel *loop* a prendersi le sue responsabilità. Altrimenti, nell'opinione del recensore, si rischia di finire come narrato da Robert Sheckley in *L'ultima battaglia* (Il secondo libro della fantascienza, Einaudi, 1961). Ai tempi del Secondo Avvento, gli uomini schierano le loro truppe per la battaglia di Armageddon, contro le forze del male. Grazie ai progressi tecnologici, l'umanità è in grado di schierare un esercito di robot, micidiali armi autonome che si spera possano costituire la mossa vincente nella battaglia contro gli angeli ribelli di Satana. Lo scontro avviene, ed è terribile, ma l'umanità vince, grazie all'estremo eroico sacrificio dei bot. E arriva il premio del Signore: dai rotami fumanti del campo di battaglia risorgono sotto lo sguardo sbalordito dei comandanti umani e ascendono al cielo nella folgorante luce della gloria di Dio i corpi di metallo dei robot caduti.

marrafranco@gmail.com

F. Marra è informatico

