

# Se all'Intelligenza artificiale diamo il potere di vita o di morte



di **Andrea Granelli**

In un celebre incidente stradale durante un test, una macchina con guida autonoma Uber ha rilevato la presenza di un pedone – Elaine Herzberg – sei secondi prima di colpirla e di ucciderla, ma l'ha identificata come una bicicletta (la Herzberg stava infatti spingendo la sua bici).

Come spiega il rapporto preliminare pubblicato dalla National Transportation Safety Board, Uber non aveva impostato il suo sistema per "agire in autonomia" su quel tipo di decisione. Gli ingegneri hanno impedito alla loro auto di frenare improvvisamente da sola in situazioni analoghe, afferma il rapporto, "per ridurre il potenziale di comportamento erratico del veicolo".

L'azienda ha deciso di affidarsi all'operatore umano dell'auto in tali situazioni per evitare incidenti. E questa scelta apre a un sistema complesso di interazioni possibili uomo-machina.

La fallibilità degli algoritmi di guida non sono però l'unico problema legato al corretto funzionamento delle macchine a guida autonoma. Ve ne sono altri due potenzialmente problematici e dalle molteplici implicazioni.

Il primo è che questi sistemi di guida usano un algoritmo "di morte" – peraltro protetto dal segreto industriale – che decide chi uccidere in una situazione critica: il/i passanti o il/i passeggeri. Un'indagine fatta dalla rivista del MIT – Technology Review – sul criterio che tali algoritmi debbano adottare dava una risposta unanime: minimizzare il costo delle vite umane.

Ma la complessità del tema non finisce qui: e se le persone che stanno attraversando hanno la fedina penale sporca? Se sono antagonisti di chi è al potere? Se hanno semplicemente un colore diverso della pelle? Nell'analisi "numerica" che vuole minimizzare il costo in vite umane valgono sempre uno? Oggi la capacità di riconoscimento facciale, unita alla possibilità di costruire algoritmi parametrici – che in questo caso non si limiterebbero a "contare" tre persone, ma ne valuterebbero le caratteristiche per scontare quel numero – sono un fatto



## **I SISTEMI DI GUIDA AUTONOMA CHE DECIDONO CHI SALVARE SONO ALGORITMI PROTETTI DAL SEGRETO INDUSTRIALE**

consolidato, non una ipotesi futuribile.

Il secondo aspetto problematico è che è relativamente semplice impossessarsi del sistema di guida e controllarlo direttamente da remoto per guidare il veicolo dove si vuole. I recenti casi del terrorismo ci ricordano che questa opportunità non è un semplice caso di scuola: guidare auto, camion sulla folla inerme è una delle strategie adottate in tempi recenti dal terrorismo islamico.

Per questi motivi il dibattito sugli algoritmi che guidano la "macchina senza guidatore" ad evitare gli incidenti "scegliendo il male minore" incominciano a destare preoccupazione. E vengono recuperate le riflessioni dei filosofi etici. Una per tutte il cosiddetto "Trolley Problem" formulato per la prima volta dalla filosofa Philippa Foot nel 1967: «un tram ferroviario ha perduto il controllo. Il guidatore non può frenare, ma può solo azionare lo scambio tra i binari. A un certo punto si trova di fronte a un bivio: seguendo il percorso previsto, ci sono cinque persone sul binario; mentre sull'altro binario – che può percorrere solo decidendo di azionare lo scambio – ce n'è solo una. In entrambi i casi, le persone moriranno nell'impatto. Cosa deve fare il guidatore? Subire passivamente quanto programmato e assistere alla morte di cinque persone o azionare deliberatamente lo scambio e ucciderne "solo" una?». Questo esercizio del pensiero vuole porre l'attenzione sul fatto che non basta la comparazione numerica del male minore; vi è anche la differenza tra assistere e determinare una morte.