

Francesco Forleo

Etica della macchine: è possibile attribuire una responsabilità ai veicoli senza conducente e dotati di AI che circolano, e sempre più circoleranno, nelle nostre strade?

Come citare questo articolo:

Francesco Forleo, *Etica della macchine: è possibile attribuire una responsabilità ai veicoli senza conducente e dotati di AI che circolano, e sempre più circoleranno, nelle nostre strade?*, «Bibliomanie. Letterature, storiografie, semiotiche», 57, no. 17, giugno 2024, [doi:10.48276/issn.2280-8833.11358](https://doi.org/10.48276/issn.2280-8833.11358)

Che gli umani da tempo immemorabile abbiano provato a dar vita alla materia che si credeva inanimata ci risulta da una lunghissima casistica di testimonianze e racconti. Questa aspirazione è presente nella storia che ci è giunta passando prima per la sapienza tradizionale e religiosa orientale, sin dai tempi della antichissima civiltà cinese, e successivamente tramite la filosofia occidentale di cui siamo tuttora permeati. Diversa dalla cultura orientale, seppure le domande ultime sull'esistenza siano perlopiù le stesse, ma da risolversi con uno stile di vita prevalentemente conservativo anche per la persistenza del mito, la filosofia occidentale tende piuttosto al suo superamento e a dare sin dall'inizio un'impronta razionalistica al proprio orientamento di pensiero, attraverso i secoli volto all'emancipazione e alla libertà. Ed a porre una distinzione fra la speculazione pura, il problema religioso e la saggezza pratica. Ma il senso di sfida verso le divinità che popolavano le prime cosmogonie non è ancora estinto. Dalla sterminata lontananza dei tempi, eroi come Prometeo e Gilgamesh sfidano il limite esattamente come l'umano contemporaneo è alla ricerca dell'ottimizzazione della vita. In entrambi i casi, di quale sfida stiamo parlando? Non di quella di competere nella creazione dell'universo, ma di quella di modificare noi stessi.

Quanto all'antichità ci rimettiamo agli innumerevoli racconti di creazioni immaginifiche, mitologiche e religiose, sino ai tentativi di costruzione fisica di protesi che aumentano le capacità umane, poi via via agli automi, macchine dotate di automatismi e, da qualche decennio, a macchine teleologiche con potenzialità e poteri sovrumani. Se facciamo

astrazione dalle applicazioni belliche, lo scopo iniziale, in particolare per gli artefatti, è generalmente propositivo, e le componenti protesiche o le macchine avrebbero dovuto migliorare o addirittura risolvere i problemi che la vita del tempo poneva dinanzi agli umani. I racconti mitologici ci parlano delle ali che Dedalo fabbricò con penne e cera per suo figlio Icaro desideroso di fuggire da Minosse, del fuoco di Prometeo necessario alla fusione e alla lavorazione dei metalli, ma poi sono comparse le leve e le applicazioni idrauliche di Archimede per agire e muovere corpi pesantissimi. Ancora intorno al III secolo a.C. ci giungono testimonianze di automi e di automatismi. E come ricorda Lucio Russo (in *La rivoluzione dimenticata, Il pensiero scientifico greco e la scienza moderna*, Feltrinelli, Milano, 2021) per Vitruvio (che a sua volta si attiene ai *Commentaria* di Ctesibio) «gli orologi costruiti da Ctesibio potevano azionare automi a orari prefissati; ma non è da meno Erone di Alessandria (I secolo d.C.) che nelle sue opere *Pneumatica* e *Sulla costruzione di automi* parla di un distributore automatico che introducendo una moneta da cinque dracme, versava una quantità prefissata di liquido».

Contestualmente alle esigenze di potenziare il corpo umano per garantirsi la sopravvivenza (anche in caso di attività belliche) e per migliorare il benessere, da sempre troviamo nella storia il tentativo di replicare in modo artificiale il nostro operato, non solo meccanico, ma anche mentale, e oggi siamo (o quantomeno molti credono di essere) a un passo dall'infondere la vita biologica nella materia atomicamente attiva, ma inerte.

Le più recenti realizzazioni informatiche sono andate oltre le intuizioni di Alan Turing e Aldus Huxley che nel famoso test di Turing erano preoccupati dinanzi al celebre e inquietante interrogativo: «Che cosa accadrà se una macchina prenderà il posto di A nel gioco?». In effetti, ciò è già avvenuto, e con dispositivi molto più sofisticati di quanto avevano immaginato gli stessi Turing ed Huxley: le automobili senza conducente.

Queste automobili esistono e circolano in alcuni territori dell'evoluto Occidente, figlio della filosofia razionale dell'antica Grecia, e ci portano a pensare che l'obiettivo sia stato raggiunto. Nella maggior parte dei più recenti convegni sulla mobilità futura, in particolare urbana, si dà per scontato che nel futuro prossimo vedremo intensificarsi la presenza di automobili e veicoli commerciali senza conducente e dotati di intelligenza artificiale, per non dire dei droni che ci consegneranno pacchi e corrispondenza a domicilio o addirittura alla finestra di casa. Ma prima di parlare di «intelligenza artificiale» (AI), una recentissima espressione delle più elevate risorse dell'ingegno umano, è utile una brevissima divagazione a mio avviso propedeutica alle riflessioni sull'etica delle macchine che vorrei proporre al lettore.

La scorsa estate un pescatore torinese, approfittando del livello estremamente basso del Po a causa della siccità, si è inoltrato nel greto e in piedi con l'acqua appena oltre il ginocchio pescava con la canna e la lenza. La pesca con la lenza, è un'arte (come nel *Sofista* dice Platone), cioè, stando all'etimo, una tecnica (nel senso di insieme delle regole che

permettono di eseguire con efficacia un'attività o una pratica intellettuale, ma torneremo su questo termine) impiegata moltissimi anni fa dall'umano per procurarsi il cibo necessario alla sopravvivenza. Oggi è un hobby o uno sport (in pochi pescano per alimentarsi) e si pratica con una canna in fibra di carbonio, un sofisticato mulinello con leve e ingranaggi che agevolano il riavvolgimento di una lenza di sottilissimo nylon, un amo invisibile ai pesci ed esca artificiale per indurli ad abboccare. Come si è giunti a questo? Come si è passati da strumenti primordiali come un bastone e una fune di canapa intrecciata a quelli attuali realizzati con una tecnologia così sofisticata? Per provare a rispondere, ci richiamiamo a due imprescindibili riferimenti storici che hanno un eminente valore epistemologico, e inoltre simbolico in questo contesto: il pensiero di Aristotele e l'Illuminismo.

L'*incipit* della *Metafisica* dice: «Tutti gli uomini per natura tendono al sapere»; queste parole riassumono bene il senso dell'incessante moto del pensiero, costituito da domande e risposte, che spinge ogni essere umano a colmare il vuoto dell'ignoranza innata ed a procedere verso la conoscenza, una conoscenza per cause, il cui possibile tramite, sino a un certo livello, non quello, appunto, scientifico dell'accertamento delle cause, è l'esperienza sensibile. L'Illuminismo, circa duemila anni dopo Aristotele, forte dell'applicazione del metodo cartesiano e del metodo scientifico di Galileo Galilei, ha eletto a prerogativa fondamentale del percorso verso la conoscenza il «lume della ragione», cioè la capacità logica e critica di leggere tutta la realtà intorno a noi, consentendoci di conoscere ed esplorare ogni cosa andando oltre i pregiudizi e le superstizioni. La ragione è così elemento regolatore assoluto.

Quindi pensiero e ragione, sintetizzati con le umane attitudini intellettive, ci hanno negli ultimi venti secoli reso consapevoli di quali fossero gli ostacoli che si interponevano fra noi e una vita migliore. E il fatto di superarli ci ha messo in condizioni di progredire nel campo della tecnica della pesca e in mille altri indirizzi molto più complicati, come, a titolo di estrema utilità nel campo medico, le endoscopie o le TAC, oltre le stesse terapie, alla cui base c'è una tecnologia farmaceutica (tanto per dissipare lo spettro di una tecnica che distorce l'umana natura).

Ora, la ragione e il pensiero sono facoltà che coinvolgono l'attività del cervello. Forti delle scoperte, di fondamentale importanza per la scienza, quali la trasformazione dell'energia, l'elettricità, il magnetismo e la radioattività, a partire dal dopoguerra hanno preso sempre più vigore e interesse gli studi sul funzionamento del cervello, l'organo più complesso del corpo umano, sede della conoscenza personale. La riflessione che ha guidato questa corrente di studi ha posto i fondamenti per un processo tale che, se si fosse capito come funzionasse il cervello, si sarebbero aperte le porte del sapere: si era dinanzi alla scoperta dell'artificiale.

Così, dopo essere riusciti in imprese oltre l'immaginazione umana sino al XVIII secolo, come volare, andare nello spazio e accelerare la materia per scomporla e studiarne i salti quantici

delle particelle che compongono gli atomi e il loro nucleo, l'ultima frontiera della scienza sembrava la comprensione del funzionamento del cervello, per curarlo innanzi tutto, ma anche per replicare il processo di formazione di un pensiero, per riprodurre artificialmente i processi che sono alla base della funzione cognitiva.

Prima di focalizzare l'attenzione al cervello in senso più ristretto o proprio, si è iniziato con l'apparato fisiatrico, con la biomeccanica di articolazioni protesizzate, con la bionica e con i *cyborg*, esseri più inumani che postumani, dotati di protesi artificiali comandate dal cervello biologico tramite una scheda bidirezionale impiantata nel corpo; corpo artificiale, metafora estrema, in quanto ibridazione con la tecnologia, del nostro confrontarci con il meccanismo. Ma qui siamo nel campo estremamente controverso del transumanesimo, che configura un'età post-biologica, preoccupante sino a un certo punto giacché essenzialmente tecno-utopica. Dal dopo l'uomo, dall'uomo aumentato o resettato, ritorniamo all'umano: l'obiettivo più ambizioso, dicevo, era quello di replicare il cervello umano nelle sue funzioni più elevate: l'attività mentale e il pensiero. Va detto che la questione della distinzione fra mente e cervello è ancora molto discussa e si è capito che semplificare come funzionano le sinapsi assimilandole ad una connessione elettrica fra poli magnetici non ha spiegato come si formino nel cervello le emozioni e i sentimenti. Tuttavia la ricerca è sempre in avanzamento, e dal dopoguerra in poi l'attività per la riproduzione del cervello si è concentrata sulle funzioni elaborative sequenziali come quelle dei calcolatori, dove l'assimilazione del sistema binario che sta alla base dell'informatica, il bit, con le scariche sinaptiche intermittenti era la strada giusta per replicare (ed eventualmente potenziare in alcuni aspetti) le funzioni operazionali del cervello umano.

La cosa inattesa emersa durante lo sviluppo dei calcolatori fu che la capacità e la velocità di calcolo delle macchine che tentavano di riprodurre l'intelligenza umana per la parte operativa risultarono eccezionalmente più potenti di quanto ipotizzato, e nacque l'illusione, sostenuta dalle correnti di pensiero alla base del processo costruttivistico della riproduzione della realtà nel nostro cervello, che si era prossimi alla riproduzione del cervello biologico in molte delle sue funzioni, non solo quella operativa. La funzione operativa del nostro cervello era comunque stata realizzata ed è diventata una realtà; trattandosi di una riproduzione artificiale, questa applicazione ha assunto nell'immaginario collettivo il nome di «intelligenza artificiale». Insomma, se riprodurre il cervello in tutte le sue funzioni sembrava una cosa decisamente complessa, il potenziamento dell'operazionismo mentale è realmente avvenuto e in pochi decenni sensibilmente progredito.

La svolta determinante nella riproduzione artificiale di una funzione assimilabile ad un processo biologico si è avuta con Norbert Wiener, nel 1943. Wiener aveva ideato un «sistema antiaereo automatico», costituito da un radar, un'arma antiaerea e un calcolatore che elaborava la rotta dell'aereo nemico da colpire e l'angolazione con cui era stato

effettuato il tiro precedente. Con un processo circolare di informazione (basato sul concetto di *feedback*) fra i tre elementi del dispositivo, il sistema, correggendo l'errore precedente, continuava finché l'obiettivo non veniva colpito. Era la prima volta che si progettava un sistema in grado di autocorreggersi, in modo automatico e senza interventi esterni verso il proprio obiettivo (l'aereo nemico), proprio come fanno la maggior parte degli esseri viventi umani ed animali. Con Wiener e la sua macchina teleologica era nata la *cibernetica*, una vera e propria rivoluzione copernicana nel rapporto uomo/macchina. Si era dinanzi alla realizzazione di un dispositivo che basava il presente sul calcolo del passato, ma che era anche in grado di prevedere il futuro attraverso la costruzione operativa.

Nei decenni successivi, sino ai nostri tempi, l'accelerazione nella capacità di realizzare sistemi in grado di correggersi automaticamente verso un obiettivo è andata oltre ogni previsione. Grazie alla tecnica (com'è noto, da *téchne*, il fenomeno reale, la competenza, l'arte, l'attività, il saper fare, creare, tanto nell'arte - con cui *téchne* è implicata, non solo nel senso di un insieme di regole per la composizione di un'opera indipendentemente dalla componente soggettiva, ma anche nella radice comune di un fare artistico - che come attività volta alla progettazione di strumenti meccanici finalizzati al miglioramento dei modi di conduzione della vita umana. *Téchne* somiglia ad *espisteme* perché implica una conoscenza dei principi, ma da essa diverge in quanto il suo fine è il fare anziché la comprensione disinteressata) e alla tecnologia (studio applicativo degli artefatti, metodologia tecnica) si sono sviluppate la robotica e l'intelligenza artificiale (che attualmente è in realtà una sovrastimata espressione del *machine learning* - l'apprendimento automatico - la sottoclasse dell'intelligenza artificiale il cui compito è istruire i computer ad imparare dai dati, quindi attraverso, per così dire, l'esperienza invece che essere programmato per calcoli predittivi) che hanno permesso la realizzazione fisica dei sistemi complessi, immaginati e teorizzati sin dagli anni Cinquanta: veicoli e armi automatiche, i droni, etc.

Ed è con i risultati sin qui ottenuti lavorando nell'ambito della tecnica e degli strumenti tecnologici che l'ingegno umano, con la sua attività poetica, ha realizzato dei meccanismi che i nostri antenati greci e latini avrebbero annoverato fra le divinità del tempo, a prescindere dall'energia elettrica che li alimenta e senza la quale sarebbero solo materia inerte. In particolare, oltre ai potenti dispositivi di produzione intelligente come i *chat-bot*, software progettati per meglio comprendere i confini del coinvolgimento nel quotidiano, parliamo ad esempio di oggetti come cani automatici (presenti in numerose aziende a fianco degli umani), droni che con i più disparati scopi solcano i cieli, auto senza conducente ed altre straordinarie realizzazioni che sono la sintesi di automi tecnologicamente evoluti, applicazioni di software, processi di *feedback* logico-discorsivi che attuano cascate operazionali usando algoritmi e calcolo infinitesimale. Fra questi, per la loro paradigmatica concentrazione di artefatti e riproduzione intellegibile, proviamo ad esaminare cosa sono e

come si collocano nel contesto sociale le automobili senza conducente, sistemi molto complessi chiamati in modo semplicistico «auto a guida autonoma dotate di intelligenza artificiale» in virtù della loro capacità di circolare fra le strade aperte al pubblico come se fossero condotte da umani. La loro presenza nelle nostre città (abbiamo esempi di Stati che hanno autorizzato stabilmente la circolazione di questo tipo di veicoli in città come Phenix, Stanford, San Francisco negli Usa, Bruxelles in Europa e in forma sperimentale a Monaco di Baviera e anche a Torino) ha cominciato a sollevare diverse questioni di cui cercheremo di schematizzarne le principali implicazioni etiche verso l'ambiente e la società in cui interagiscono.

In primo luogo sorge la questione relativa alla loro costruzione: perché si è pensato di realizzare questi veicoli?

In secondo luogo è sorta la domanda che trova il suo nucleo nella definizione funzionale di questi veicoli: essendo a guida autonoma, questi veicoli sono autonomi?

In terzo luogo, se sono autonomi, sono responsabili del loro comportamento?

In quarto luogo, agendo in modo intelligente essendo dotati di intelligenza artificiale, sono realmente intelligenti?

Infine, l'interrogativo più spinoso: queste macchine, dotate di potenti elaboratori, generano timori fra persone che sospettano che un giorno non troppo lontano possano assumere decisioni e governare al posto degli umani: è questo un rischio reale?

Procediamo con ordine e proviamo a rispondere alla prima domanda, la cui risposta è per certi versi sconcertante. Se inizialmente si era pensato di realizzare un oggetto che ci alleviasse dalle fatiche della guida o dallo stress di un percorso cittadino nelle ore di punta, ora la principale ragione della loro ideazione e realizzazione si è spostata agli interessi economici delle grandi compagnie di assicurazioni che, a fronte delle ingenti somme di denaro che vengono esborsate per ripagare i danni degli incidenti stradali (in Usa l'ordine di grandezza è di miliardi di dollari ogni anno, più o meno il bilancio economico annuale dell'Italia) e avendo preso atto che la maggior parte dei sinistri è causato da errori umani, queste compagnie hanno finanziato la ricerca di veicoli che eliminassero gli errori umani e quindi si riducessero i costi dei rimborsi assicurativi. Del resto, in alcune aree del pianeta circolano milioni di veicoli (in Italia ne sono immatricolati 46 milioni di cui 39 milioni sono automobili per il trasporto di persone), molti in luoghi come le città ad altissima concentrazione urbana. Queste alte concentrazioni portano per calcolo statistico-probabilistico all'interferenza fisica dei veicoli stessi creando sinistri con persone ferite e inevitabilmente con vittime. Le conseguenze sono gli elevatissimi costi a carico delle compagnie assicurative e per le comunità che devono sostenere le cure dei feriti. Il sillogismo secondo cui togliendo la guida dei veicoli agli esseri umani si sarebbero eliminati gli errori di guida e i danni che ne conseguono ha dato impulso alla decisione di realizzare automobili che non sbagliano, e che quindi non creano incidenti.

Una volta costruiti, se ben realizzati, questi veicoli sono effettivamente in grado di muoversi nel traffico e di relazionarsi con gli automezzi tradizionali per mezzo dei loro sensori e dei radar che acquisiscono i dati ambientali e con i programmi di utilizzo attivano i comandi che permettono la conduzione. Sembrerebbero effettivamente autonomi, ma lo sono realmente? Veniamo così alla seconda domanda: i veicoli a guida autonoma sono effettivamente autonomi? Come sappiamo il significato di «autonomia» presenta diverse accezioni, dalla abilità nel garantire l'alimentazione di un dispositivo energivoro tramite l'azione chimica di un carburante nel serbatoio, sino al senso che a questo punto è più attinente al nostro caso: autonomia nel suo significato più generale deriva dalla composizione di *αὐτός* «stesso» e *νόμος* «legge», vale a dire, come dalla Treccani, «il potere di dar legge a sé stesso». Un significato generale poi declinato nelle varie epoche e in diversi campi scientifici della politica, nel diritto, nella filosofia.

Ora, analizzando la loro costruzione e il sistema che ne governa la funzionalità, si ricava che tali veicoli sono esattamente l'opposto di un oggetto autonomo come dovrebbe essere un artefatto in grado di dare regole e leggi a se stesso. Nessuna autonomia è lasciata a questo tipo di macchine, nessuna istruzione può essere convertita da decisioni proprie, anzi, la macchina deve seguire in modo rigoroso le leggi e le prescrizioni imposte dal suo progettista. Ogni comando, dal più semplice al più complesso in quanto frutto di una elaborazione algoritmica delle istruzioni, viene emesso in base a regole pensate *a priori*, precise e incontrovertibili: niente è lasciato alla interpretazione o all'iniziativa della macchina, l'eventuale imprevisto sarebbe imputato a un errore di progettazione. Dall'analisi del loro funzionamento, dovremmo chiamarli più che veicoli a guida autonoma, «veicoli a guida automatica». Infatti «automatico», per definizione, è il comportamento di una macchina o di un dispositivo che, opportunamente regolato (cioè già dotato di regole), è capace di compiere determinate operazioni senza l'intervento diretto degli umani.

Quanto alla terza questione: se i veicoli senza conducente si muovono in modo tale da interagire con l'ambiente, in caso di incidente sono responsabili? Anche in questo caso l'etimologia ci aiuta a chiarire il senso della parola «responsabile»: dal latino «respondere» (essere garante di qualcosa) è un termine composto di *re* (indietro) e *spondere* (promettere), più il suffisso *-bile/-abile* che indica facoltà, possibilità, quindi disposizione a rispondere delle proprie azioni. Pertanto è responsabile «chi risponde dei propri comportamenti perché ne è abile, è pronto a spiegarne le ragioni e a subirne le conseguenze» (Treccani). Veniamo al tipo di veicolo che stiamo analizzando. Abbiamo visto che questi veicoli riescono a muoversi disinvoltamente come se fossero pilotati da umani, i quali tuttavia, se in grado di intendere e di volere, sono responsabili delle proprie azioni. Questo è possibile grazie a un estremo impegno tecnico, sia come progettazione e sviluppo delle parti comuni ad altre tipologie di automobili (il sistema motopropulsore, lo sterzo, i freni, etc.), sia per la parte che ne permette la conduzione automatica, la parte di governo del veicolo che contiene le

regole di comando, riconducibile alla sintesi di processi informatici che sono scritti usando il software. E infine per la componente tecnologica che fa capo ai semiconduttori che ne costituiscono l'hardware. Il risultato è frutto del lavoro di ingegneri e progettisti, che a loro volta dovranno attenersi alle regole di progettazione. Abbiamo visto che la responsabilità è la capacità di rispondere delle proprie azioni. È un principio su cui si basa la giurisprudenza nelle nostre civiltà per la punibilità o meno di un gesto, se mosso da libera volontà. La volontà è il movente delle azioni libere che nella disciplina giuridica determinano la giudicabilità di chi compie il gesto e delle conseguenze del gesto stesso.

Nella *Prefazione alla Critica della ragion pratica*, Kant scrisse che la libertà è «la *ratio essendi* della legge morale, ma la legge morale è la *ratio cognoscendi* della libertà»: la libertà consente di conoscere le leggi e diventare responsabile delle azioni, ma senza libertà non c'è responsabilità dinanzi alla legge morale. Se ne ricava che il veicolo in sé, come detto sopra, non è libero e non è in grado di decidere o volere un'azione. Le decisioni sul comportamento del veicolo sono scritte dai progettisti e a costoro fa capo la responsabilità primaria della missione dell'auto senza conducente che sarà costruita.

La questione se i veicoli senza conducente, essendo dotati di istruzioni che riproducono l'intelligenza umana, siano intelligenti è oltremodo interessante. Ma prima di analizzare il comportamento dei veicoli dal punto di vista etico, è utile delineare i confini di una forma intellettuale biologica e capire cosa vi ritroviamo di essa in un artefatto. Per definire cosa sia un'intelligenza biologica, rinvio ad uno dei più importanti istituti di catalogazione del sapere, e mi spiace in questa sede non affrontare (cosa che aveva fatto molto bene Wiener nel 1948, in *La cibernetica: controllo e comunicazione nell'animale e nella macchina*, trad. it. Il Saggiatore, Milano, 1968) anche le intelligenze che l'evoluzione darwiniana ha riconosciuto negli animali non umani. Nella sua formulazione più generale, l'intelligenza biologica-umana è racchiusa - e mi rimetto ancora a Treccani - nel «complesso di facoltà psichiche e mentali che consentono all'uomo di pensare, comprendere o spiegare i fatti o le azioni, elaborare modelli astratti della realtà, intendere e farsi intendere dagli altri, giudicare, e lo rendono insieme capace di adattarsi a situazioni nuove e di modificare la situazione stessa quando questa presenta ostacoli all'adattamento; propria dell'uomo, in cui si sviluppa gradualmente a partire dall'infanzia e in cui è accompagnata dalla consapevolezza e dall'autoconsapevolezza». Ma le auto senza conducente, cioè definite in modo più comune «a guida autonoma», dispongono davvero di intelligenza e sono quindi infallibili? Cosa riconosciamo di intelligente nel loro comportamento etico? Non esploreremo qui il campo della complessa declinazione delle etiche (etica del dovere, delle conseguenze, etc.), ma per quanto detto sinora possiamo enucleare qualche conclusione: se confrontiamo la definizione di «intelligenza», sicuramente alcuni aspetti di questi veicoli sono riconducibili ad un comportamento intelligente. Grazie agli enormi progressi nella scrittura del software e nell'elaborazione con algoritmi, elaborando i dati che giungono dai sensori

che rilevano l'ambiente, i fatti e le azioni esterne, le centraline dei veicoli senza conducente compiono astrazioni della realtà, e si adattano alle situazioni (per esempio se il veicolo dinanzi frena improvvisamente, anche l'auto a guida autonoma frenerà in modo energico per evitare un tamponamento). Ma troviamo alcune difficoltà a riconoscere altre caratteristiche che fanno parte dell'intelligenza umana, quali il pensare, il comprendere, spiegare i fatti, l'intendere e farsi intendere dagli altri, il giudicare, l'aver consapevolezza, etc.

Questa puntualizzazione ci porta nel cuore del problema richiamato in tanta letteratura che riguarda la capacità decisionale dinanzi a fatti irreversibili: primo fra tutti l'esempio del guasto ai freni del carrello ferroviario. Il problema etico del carrello guasto, posto come scelta fra due opzioni in cui in ogni caso qualcuno morirà (o il passeggero o le persone al bordo della strada), è tale che a livello statistico non ha riscontri pratici. Il combinato disposto di un carrello ferroviario con i freni inefficienti che si muove in discesa verso un bivio con scambio è il seguente: da un lato sono adagiate sui binari cinque persone e nell'altro ramo del bivio si trova un muro che fermerà la corsa del carrello causando la morte del passeggero che si trova sul veicolo con i freni guasti. A lato della scena un manovratore può decidere, azionando lo scambio, se far morire una o cinque persone. Non è una situazione impossibile, ma è una situazione statisticamente irrilevante in cui le probabilità che si verifichi nel mondo reale sono di fatto prossime allo zero. Per cercare di rispondere alla domanda che chiedeva se questi veicoli, avendo funzioni intelligenti, fossero essi stessi intelligenti, si può iniziare a circoscrivere in che modo sono intelligenti. In effetti riconosciamo alcuni aspetti comuni fra l'intelligenza biologica e quella artificiale, cioè quelli ascrivibili ai processi computazionali che, va detto con onestà intellettuale, sulle macchine sono infinitamente più potenti e veloci rispetto alle competenze umane. Ma per ora non si va oltre. Quindi più che di macchine intelligenti potremmo dire che si tratta di esempi di apprendimento automatico o *machine learning*, macchine che auto-apprendono automaticamente (ed ecco riapparire il concetto di «automatico», di «automatismo») in virtù di due livelli operazionali: un primo livello di acquisizione avviene durante la loro attività, in quanto ampliano la loro banca perché sono in grado di connettersi con la rete web. Un secondo livello consiste nella memorizzazione e classificazione dei dati acquisiti che verranno usati in missioni future. Quindi un esempio di *machine learning* di straordinaria efficacia. Di nuovo abbiamo capito che tutte le prescrizioni (la parte morale cui fa riferimento la responsabilità) vanno progettate e risolte non sul piano operativo in sé, ma su quello della scrittura degli algoritmi con le loro regole, cioè il piano della volontà operativa. La responsabilità etica va pertanto ricercata nella scrittura delle istruzioni, eseguita dai progettisti, che, se hanno lavorato in libertà, permetteranno alla macchina di filtrare la tipologia di dati e pre-scrivere come verranno utilizzati durante le aggregazioni. Queste funzioni automatiche di acquisizione ed elaborazione sono un elemento comune a

tutte le macchine in grado di processare i dati acquisiti durante il loro funzionamento disponendo della funzione di *machine learning*: le macchine così come sono realizzate - e come saranno - ci servono e non possiamo più farne a meno. La mole di dati che vengono prodotti da noi umani è enorme e può essere gestita ed elaborata solo da macchine potentissime: noi non siamo più in grado di farlo, e ci spetta il compito di valutare e giudicare i risultati delle elaborazioni. Dobbiamo temere le macchine, specialmente quelle potenti che dispongono della funzione di *machine learning*?

Riscontro che molti commenti sollevati più frequentemente in occasione di incontri e convegni, nel dibattito attuale sulle AI, tradiscono una paura dell'ignoto, il timore che l'intelligenza artificiale (nel mondo viene concettualizzata così), con la sua funzione di apprendimento automatico, possa sopravanzare la nostra scienza e la nostra etica, inglobandole nella tecnica e nella tecnologia che noi stessi abbiamo progettato. Ecco, direi che più che temere le macchine e la loro potentissima tecnologia, dovremmo imparare a conoscerle, usare il loro potenziale e non farci ossessionare da un futuro di qualcosa che è già accaduto e che sino ad ora siamo stati in grado di padroneggiare.

Spesso il richiamo all'esperienza diretta è utile per individuare una possibile forma di convivenza fra noi e le macchine. Come tutte le cose della vita, e in fondo come la vita stessa, fatta di prove e di calcoli purtroppo non matematici e quindi talora inconseguenti, le competenze si formano sul campo. E se oggi molte cose mi paiono sin troppo scontate, a ripensarci per molti passaggi non potrei non attingere alle acquisizioni assimilate nel corso di una lunga esperienza diretta in importanti centri di ricerca e di sperimentazione. E come mi è capitato di dire proprio in questa sede (<https://www.bibliomanie.it/?p=218>), era verso la metà degli anni Settanta quando superai la selezione per l'ammissione al Centro Ricerche della Fiat e venni assegnato all'ufficio calcoli dei motori studiati anche per la Ferrari. Eravamo una piccola equipe incaricata di eseguire dei riscontri sui risultati ottenuti con l'aiuto di un calcolatore - allora d'avanguardia - con quelli ottenuti attraverso calcoli tradizionali. Ricordo che fui fra i primissimi ad utilizzare, e letteralmente, viste le sue dimensioni, ad entrare nell'Univac, un calcolatore per molti versi affine a quello usato dagli Usa nelle missioni spaziali. Oggi può sembrare assurdo, ma l'Univac occupava una stanza di oltre cento metri quadrati per una potenza di calcolo pressappoco di un decimo di quella dei nostri computer da tavolo. Entravo fisicamente nel calcolatore e non nascondo che i primi giorni avevo un po' di paura quando inserivo i dati necessari per i calcoli ricavandoli dalla perforazione codificata di schede di cartoncino. Visto oggi non era un lavoro di alto profilo intellettuale, ma era utile e lo svolgevo con diligenza: erano altri tempi, e non vorrei ora confondere la nostalgia con la consapevolezza di allora di starmi intanto assicurando una competenza attraverso l'apprendimento delle caratteristiche e delle risorse del meccanismo. Tornando all'autonomia o meno del meccanismo, col tempo abbiamo imparato che non è la tecnologia che va temuta, ma l'uso improprio che l'umano può farne (un esempio attuale

sono i *chat-bot* come Chat GPT4, che è in grado di relazionarsi con noi umani con domande e risposte che hanno la fonte nella rete andrebbero regolamentate per scongiurare la produzione di documenti o elaborati artificiali attribuiti ad umani).

E se c'è molto poco di intelligente nella conduzione autonoma di un veicolo, è perché al veicolo manca la funzione volitiva: seppure spesso usata male da noi umani, la volontà resta un'alta espressione dell'intelligenza biologica. Per programmare qualcosa che abbia una volontà propria, correggibile in atto se si valuta che possa provocare incidenti o azioni pericolose, il percorso è ancora lungo e complesso, e non si sa quanto fattibile. Siamo ancora distanti dalla riproduzione di qualcosa che sia assimilabile al pensiero, che possa sintetizzare le facoltà psichiche dell'intelligenza animale/umana, mancano le sensazioni, la gioia, la felicità, la noia, l'innamoramento, la paura della morte, la sofferenza - tutte le facoltà emozionali che nell'umano suppongono un futuro -, soprattutto mancano la consapevolezza e la volontà.

Ciò non toglie che siamo dinanzi a qualcosa di portentoso nella memorizzazione ed elaborazione dei dati. Seguendo in modo rigoroso le istruzioni scritte e supervisionate da noi umani e con la loro velocità, questi prodigiosi dispositivi ci inducono a interpretare le risposte elaborate come se fossero pensieri. Su questo piano, le capacità di programmazione, elaborazione e gestione dei dati è in incessante avanzamento, gli ingegneri informatici presentano quotidianamente i risultati delle nuove ricerche che risolvono i problemi della società con macchine che in futuro potrebbero operare come gli umani e, se oggi non se ne vede la possibilità, non si può del tutto escludere che questo accada. Ed è in merito a questa eventualità che sorge il timore che le artificialità intelligenti possano sfuggire al controllo umano, la paura che sopravanza l'entusiasmo per queste tecnologie emergenti. Ma le macchine non vanno temute, bensì usate, e l'etica e la responsabilità stanno in chi le progetta: per seguire un'etica liberamente bisogna essere consapevoli e responsabili. Come possono essere eticamente responsabili oggetti che non pensano, non hanno consapevolezza, non provano sentimenti? Le macchine sono dispositivi tecnici sofisticati e potentissimi, ma fatte da noi umani, e solo noi ne siamo eticamente responsabili. Insomma, c'è molta umanità nell'artificiale, specialmente negli artefatti che dispongono della cosiddetta «intelligenza artificiale», e se le macchine hanno difetti comportamentali abbiamo il dovere morale di correggerle. Come? Con tanta intelligenza umana e il lume della ragione. Ci riusciremo? Non lo sappiamo ma dobbiamo provarci. Poi, come diceva Manzoni chiedendosi se Napoleone fosse stato un liberatore o un conquistatore, un bene o un male per l'umanità, «ai posteri l'ardua sentenza»!