

APhEx 25, 2022 (ed. Vera Tripodi)  
Ricevuto il: 14/05/2021  
Accettato il: 21/10/2021  
Redattori: Claudio Calosi & Pierluigi Graziani

**APhEx**  
PORTALE ITALIANO DI FILOSOFIA ANALITICA  
GIORNALE DI **FILOSOFIA**  
NETWORK  
N°25, 2022

L e t t u r e c r i t i c h e

Guglielmo Tamburrini, **Etica delle macchine. Dilemmi morali per robotica e intelligenza artificiale**. Carocci editore, Roma, 2020, pp. 151.

*Viola Schiaffonati*

### **1. Introduzione**

Intelligenza artificiale (IA) e robotica sono da qualche tempo al centro di un interesse filosofico crescente. Tale interesse è una conseguenza dello sviluppo che ha caratterizzato queste discipline negli ultimi anni. Si tratta di risultati strabilianti non solo agli occhi dei non esperti ma anche – e forse soprattutto – di quelli che conoscono la storia dell’IA. Per molto tempo queste discipline sono state relegate in laboratori di ricerca accademici dove, nel corso degli anni, hanno ottenuto risultati significativi, ma con uno scarso impatto sulla nostra vita quotidiana. Ultimamente le tecnologie prodotte in questi settori sono entrate a far parte della nostra quotidianità o si accingono

a farlo, suscitando grandi aspettative ma anche paure incontrollate. Non è un caso che la tradizionale riflessione filosofica, in gran parte dedicata a cercare di rispondere a domande su che cosa sia l'intelligenza e se e come essa possa essere realizzata in artefatti creati da esseri umani, si concentri oggi sull'etica di tali artefatti, dove la preposizione "di" può essere intesa col significato sia di "appartenente a" sia di "relativa a". Non sono più i tempi in cui il libro di testo più usato al mondo nei corsi di IA (Russel & Norvig, 2020) intitolava uno dei paragrafi del primo capitolo introduttivo "Cosa succede se ci riusciamo?"; oggi piuttosto pare giunto il momento di riformulare questo titolo in "Ci siamo ormai riusciti: e ora che cosa facciamo?".

Il libro *Etica delle macchine* di Guglielmo Tamburrini affronta questa domanda, dando forma agli interrogativi etici che nascono in questo contesto e discutendo delle politiche etiche che ne conseguono. Come il sottotitolo ben illustra ("Dilemmi morali per robotica e intelligenza artificiale"), il volume presenta i principali interrogativi etici che nascono dallo sviluppo di sistemi artificiali autonomi e intelligenti, insieme ai dilemmi morali e alle scelte individuali e collettive a cui siamo chiamati di fronte a queste tecnologie. Fin qui il libro potrebbe semplicemente sembrare uno fra i numerosi tentativi, anche editoriali, di presentare gli interrogativi etici di IA e robotica a un pubblico di lettori colti ma non esperti, proponendo casi di studio e la relativa discussione delle tematiche di carattere morale a essi connessi. In realtà si tratta di molto di più, e in parte, anche di diverso. In primo luogo, concentrandosi principalmente su veicoli e armi autonome, il volume non intende solo presentare dei casi paradigmatici delle tecnologie in discussione e dei temi etici a esse soggiacenti. Piuttosto questa specifica attenzione delimita in maniera molto netta e precisa l'ambito di indagine che riguarda i problemi etici dell'*autonomia operativa* nei sistemi cosiddetti intelligenti. Come Tamburrini spiega chiaramente nell'introduzione, le questioni morali discusse nel volume sono quelle che sorgono quando il passaggio fra la decisione algoritmica e l'azione robotica è così rapido da restringere in maniera considerevole, o addirittura eliminare del tutto, la possibilità di un intervento da parte degli esseri umani. Infatti nel caso dei veicoli autonomi o delle armi autonome l'intervallo temporale fra la decisione e l'azione non è abbastanza lungo perché gli esseri umani possano considerare la decisione frutto di un processo algoritmico, capire come è stata presa ed, eventualmente, contestarla o modificarla. L'autonomia operativa in contesti in cui il fattore temporale rende difficile l'intervento umano rappresenta dunque la

specificità di questo volume che riflette, sicuramente per primo nel panorama italiano, su questi problemi in maniera sistematica.

Nei prossimi paragrafi cercheremo di mettere in luce le peculiarità di questo lavoro. Ci preme sottolineare però, fin da subito, un elemento importante. Si tratta infatti di un'analisi etica basata su considerazioni epistemiche relative agli artefatti tecnici in discussione, ma anche funzionale alle scelte politiche che da queste analisi devono essere informate. Se lo specifico contesto dell'autonomia operativa rende il volume originale dal punto di vista dell'oggetto, l'etica epistemicamente fondata e in funzione della politica costituisce la sua unicità dal punto di vista del metodo.

## **2. Autonomia operativa e scelte morali**

Il libro prende avvio dalla discussione delle questioni etiche connesse ai veicoli a guida autonoma (VA d'ora in poi). Si tratta, oggi, di un tema al centro dell'attenzione: dotte discussioni su come dovrebbe comportarsi un VA in caso di incidenti inevitabili monopolizzano la ricerca accademica e il dibattito pubblico. L'etica dei VA pare diventata l'emblema di tutte le questioni alla confluenza fra etica e sviluppo tecnologico. Anche il volume di Tamburrini si apre con i VA; tuttavia la prospettiva originale che caratterizza questo volume si delinea fin da subito in maniera molto chiara: i VA sono usati come chiave di accesso alle questioni etiche che riguardano l'autonomia operativa dei sistemi alla confluenza fra IA e robotica. Si tratta di quella autonomia che non richiede intervento umano e si manifesta nel passaggio dal ragionamento algoritmico, reso possibile da tecniche di IA, alle azioni realizzate in un determinato ambiente per mezzo di un corpo robotico.

Per caratterizzare i gradi crescenti di autonomia degli autoveicoli sono stati proposti dalla *Society of Automotive Engineers* sei livelli che descrivono i loro compiti a partire da un livello base fino ad arrivare ai livelli di autonomia crescente. La guida di un veicolo è scomposta in compiti più semplici, attribuiti di volta in volta all'essere umano o alla macchina a seconda del livello di autonomia. Nonostante questa classificazione sia ormai comunemente accettata, Tamburrini propone di aggiungere ulteriori elementi all'analisi del concetto di autonomia, introducendo una differenza fra l'autonomia *per* alcuni di questi compiti (per esempio la possibilità per il guidatore di togliere le mani dal volante), il

*dove* di questa autonomia (per esempio la guida in una corsia riservata oppure nel centro di una città), e l'autonomia *da chi* (per esempio dall'essere umano o da altri agenti artificiali). In quest'ultimo caso, l'idea è di mettere in evidenza che l'autonomia dei VA può essere autonomia non solo dagli esseri umani, ma anche da altri agenti artificiali. Ciò consente di sottolineare un elemento importante, ossia che l'autonomia operativa dovrebbe essere attribuita, più che al singolo VA, all'intero sistema distribuito composto dal VA e da tutti gli altri elementi necessari per gestire la guida, come per esempio i sistemi di localizzazione GPS, i sensori sulla rete stradale che inviano informazioni al veicolo, i collegamenti con altri veicoli per coordinarsi, e così via.

L'impostazione "per-dove-da chi" può essere estesa ad altri casi e mette in luce come molti dei nuovi problemi etici discussi in questo contesto nascano dall'automazione di compiti di percezione, ragionamento e azioni che, fino a un certo momento, sono stati di esclusiva pertinenza umana.

L'idea di avere attività che siano di esclusiva pertinenza umana è una costante nella storia dell'IA e, in particolare, nel contesto dei diversi tentativi di dare una definizione appropriata a questa disciplina. Questa idea è presente nel documento costitutivo di nascita dell'IA (McCarthy *et al.*, 1955) ed è stata elaborata in seguito da Nils Nilsson (2010), uno dei pionieri della disciplina. Nel contesto italiano è stata ripresa da Marco Somalvico che l'ha espressa in questo modo:

L'IA è quella disciplina, appartenente all'informatica, che studia i fondamenti teorici, le metodologie e le tecniche che permettono di progettare sistemi hardware e sistemi di programmi software capaci di fornire all'elaboratore elettronico delle prestazioni che, a un osservatore comune, sembrerebbero essere di pertinenza esclusiva dell'intelligenza umana. (Somalvico, 1987, 172).

Naturalmente ciò che è di esclusiva pertinenza dell'intelligenza umana cambia nel tempo: basti pensare a come diverse attività – dal gioco degli scacchi alla dimostrazione di teoremi – siano state a lungo prerogative umane e oggi siano invece incluse fra i compiti che possono essere facilmente svolte da programmi, neppure fra i più sofisticati. Il tema della definizione di cosa sia l'intelligenza e se essa possa essere riprodotta in un artefatto è un filo rosso che attraversa la storia dell'IA addirittura fin da prima della sua nascita ufficiale nel 1956; basti pensare al sogno di Leibniz di meccanizzare la ragione o al test di Turing per valutare se una macchina possa pensare (Turing, 1950). Da qualche anno, oltre alle caratteristiche più classiche come le capacità di percezione, ragionamento, apprendimento,

comunicazione e azione, l'autonomia rappresenta un ulteriore elemento dei sistemi prodotti nell'ambito dell'IA, tanto che spesso per caratterizzarli si usa la locuzione "sistemi autonomi e intelligenti".

Se le tradizionali difficoltà nel cercare di stabilire una definizione precisa di intelligenza o autonomia rimangono aperte, occorre considerare anche che le azioni discusse nel contesto dell'etica dei sistemi intelligenti e autonomi sono quelle nelle quali le responsabilità morali sono in parte trasferite a sistemi artificiali. Il merito del volume di Tamburrini è di circoscrivere ulteriormente questi elementi, sottolineando come la cosiddetta "finestra temporale ristretta" fra la decisione algoritmica e l'azione robotica getti una nuova luce sui problemi morali. Questa possibilità non è propria solo dei VA, ma anche di altri sistemi alla confluenza fra IA e robotica discussi nel volume, come le armi autonome o i robot chirurgici in grado di svolgere in autonomia una serie crescente di compiti. Non si tratta quindi solo di delegare a sistemi artificiali, per quanto complessi, compiti che impattano su responsabilità morali, ma di considerare la possibilità stessa dell'intervento umano. La drastica riduzione dello spazio di intervento umano, causato dalla rapidità del passaggio fra ragionamento e azione proprio di alcuni sistemi intelligenti e autonomi, costringe a ripensare il tema del controllo umano chiedendosi, come vedremo in seguito, in che modo esso possa essere "significativo".

### **3. Dai casi di studio alle famiglie di teorie**

Come abbiamo visto, l'attenzione all'autonomia operativa costituisce il filo conduttore di questo volume: non solo consente di tracciare dei confini rispetto all'oggetto di indagine, ma anche di delineare un peculiare metodo di indagine. Innanzitutto, sgombra il campo dalle discussioni incentrate sull'attribuzione di capacità morali ai sistemi autonomi e intelligenti. Trattandosi esclusivamente di autonomia operativa, Tamburrini ritiene che sia un errore categoriale attribuire responsabilità morali alle macchine attuali, riportando al centro del dibattito l'essere umano come unico agente morale. Sebbene ambito e metodo di indagine siano chiaramente circoscritti, la riflessione etica sull'autonomia operativa delle macchine presenta un ampio panorama di diversi punti di vista normativi. Tale pluralismo etico viene presentato da Tamburrini in termini di opportunità, e diventa un ulteriore tratto distintivo del volume. La tradizione filosofica da cui attingere per analizzare l'etica dei sistemi autonomi muove dall'etica delle conseguenze e dall'etica dei doveri, ma non si esaurisce in esse,

comprendendo per esempio anche l'etica delle virtù. Si tratta inoltre di un pluralismo etico che ha una forte impronta pragmatica: nell'analisi dettagliata di diversi casi di studio Tamburrini mostra come spesso occorra tenere in considerazione elementi di diverse teorie etiche per cercare di offrire soluzioni articolate a problemi complessi, come nel caso delle linee guida redatte dalla Commissione etica del ministero dei Trasporti e dell'Infrastruttura digitale della Repubblica federale tedesca.

Il ricorso ai casi di studio costituisce un altro elemento caratterizzante di questo libro dal punto di vista metodologico. In primo luogo i casi di studio hanno il merito di presentare in modo chiaro i problemi in discussione. Sebbene non è necessario che siano sempre reali, in ogni caso esprimono situazioni realistiche e consentono di rendere l'analisi accessibile ai più. Infatti, piuttosto che partire dall'enunciazione dei punti di riferimento teorici per poi presentare degli esempi, Tamburrini preferisce far emergere le teorie dalla discussione dei casi. Ciò è in linea con molte delle più recenti impostazioni nell'ambito dell'etica della tecnologia (si veda per esempio (Taebi, 2021)) in cui la riflessione etica scaturisce dalla discussione di casi pratici. Tamburrini ha però il merito di mantenere un equilibrio prezioso fra l'attenzione ai casi di studio e l'enunciazione di teorie. Se infatti un punto di partenza esclusivamente teorico può essere poco chiaro o troppo rigido, è anche vero che la discussione minuziosa di casi di studio troppo dettagliati può far correre il rischio di perdere di vista lo sfondo teorico generale.

L'attenzione ai casi di studio ha anche un altro merito: mettere in luce che l'etica non è una raccolta di regole o un libro di istruzioni in cui le teorie sono accompagnate da algoritmi che forniscono soluzioni uniformi ai problemi considerati. Questo è un aspetto essenziale – forse mai abbastanza evidenziato – soprattutto nell'ambito della discussione dell'etica della tecnologia. Partire da un problema specifico consente di evidenziare i principi morali insieme all'attenzione alle implicazioni in un determinato contesto. Tuttavia è bene riconoscere che questo modo di procedere non offre nessuna garanzia nell'elaborare un insieme unico e coerente di indicazioni morali, dato che il ragionamento etico non è un'inferenza induttiva, ma un processo lungo e faticoso per individuare soluzioni locali a problemi complessi. Questo ragionamento complesso permette inoltre di riconoscere somiglianze e analogie fra vecchi e nuovi casi. E proprio l'uso delle analogie è esplicitamente teorizzato da Tamburrini. Per esempio la discussione sull'etica dei VA presentata con l'ausilio di casi di studio ha il

merito di andare al di là della carrellologia<sup>1</sup>, che pur viene discussa, e di introdurre problemi simili in contesti decisionali diversi. La tensione fra etica dei doveri ed etica delle conseguenze delineata nella discussione emerge anche nel dibattito sulla tortura che, se da un lato è vietata sulla base di un principio deontologico, dall'altro è ammessa in situazioni in cui essa protegga il benessere collettivo individuato in base ai principi dell'etica delle conseguenze.

Attraverso questo peculiare metodo di indagine, il volume di Tamburrini presenta gran parte del dibattito attuale sulle questioni etiche connesse all'autonomia operativa dei sistemi di IA e robotica. A veicoli e armi autonome è dedicato lo spazio maggiore; tuttavia in questo scenario trovano posto anche altri temi, come quello dei robot calciatori o di quelli chirurgici, fino all'impiego di sistemi di IA e robotici per la sorveglianza o nel mondo del lavoro. Proprio il futuro del mondo del lavoro in seguito all'adozione di sistemi intelligenti e autonomi è un ulteriore tema al centro dell'attenzione. Le opinioni sono spesso polarizzate ed è difficile trovare discussioni che vadano oltre la contrapposizione fra una visione in cui il mondo del lavoro sarà reso peggiore dall'avvento di macchine intelligenti e l'auspicio che tali macchine possano rendere gli esseri umani liberi dal lavoro meno creativo e finalmente pronti per godere delle gioie dell'ozio. Anche in questo caso il libro di Tamburrini ha il merito di offrire uno sfondo concettuale più ampio che, in questo caso, include anche una prospettiva storica. Il riferimento a Norbert Wiener e ad Alan Turing in questo contesto consente di mostrare il loro ruolo pionieristico sui temi che affrontano la relazione fra attività operative e macchine, sottolineando come i loro "cannocchiali teorici" abbiano consentito di vedere lungo su temi che fra la fine degli anni quaranta e l'inizio dei cinquanta del secolo scorso potevano essere solo immaginati. L'articolata visione di Wiener, per esempio, viene messa in relazione con la concezione della giustizia distributiva espressa da John Rawls. Nello stesso tempo si riconosce che la presa di posizione più generale di Wiener deve essere oggi articolata sulla base di un confronto sistematico fra le diverse concezioni di giustizia

---

<sup>1</sup> Per carrellologia si intende la discussione filosofica che ha preso l'avvio dal cosiddetto problema del carrello ferroviario (*trolley problem*). Si tratta di un esperimento mentale che discute i dilemmi relativi alla scelta morale di un osservatore che vede un carrello ferroviario fuori controllo che sta investendo delle persone e deve decidere se e come intervenire. Tale esperimento mentale è stato applicato al caso degli incidenti inevitabili che coinvolgono VA e ha spesso polarizzato la discussione sull'etica dei VA. Per ulteriori dettagli si veda (Lin, 2015).

distributiva, riflettendo su quali diseguaglianze possano essere moralmente accettabili o meno.

Vale la pena ricordare, infine, che l'approccio all'etica applicata adottato dall'autore, pur sottolineando che i sistemi tecnologici debbano essere sottoposti a una valutazione razionale, non deve esaurirsi nella razionalità scientifica e tecnologica. Come Tamburrini evidenzia in modo chiaro, l'etica applicata alla tecnologia deve superare la visione unilaterale dello sviluppo tecnologico come dato positivo e riconoscere, invece, come l'equilibrio fra innovazione, conoscenza, saggezza e sviluppo sostenibile funga da ideale regolativo. Tale ideale deve svolgere un ruolo sia a livello etico, per aumentare la consapevolezza in merito alle criticità morali dello sviluppo tecnologico, sia a livello politico, per indirizzare questo sviluppo tecnologico sulla base di scelte eticamente motivate.

#### **4. Dall'incertezza epistemica alle politiche etiche**

La relazione fra etica e politica è un altro elemento caratterizzante di questo volume. Fin dall'inizio, Tamburrini mette in guardia dal fermarsi alla comprensione delle questioni morali senza fornire indicazioni politiche che riguardino sia la sfera dell'azione individuale sia quella dell'azione collettiva. Questo è evidente per esempio quando si affronta una delle domande oggi centrali nel dibattito sulle armi autonome, ossia se debbano essere messe al bando o meno, e si illustrano le diverse politiche etiche che possono essere adottate. Tuttavia la centralità dell'aspetto politico non si esaurisce nella continuità fra comprensione delle questioni morali e scelte politiche; essa si estende al riconoscere al ricercatore il ruolo di "vedetta etica". Ciò è evidente per esempio nel contesto del dibattito sulle armi autonome, in cui i ricercatori hanno il merito di aver sensibilizzato le parti in causa e l'opinione pubblica sulla necessità di un dibattito trasparente su questi temi. Tamburrini ricostruisce i principali passaggi del processo di sensibilizzazione come per esempio il caso della campagna *Stop Killer Robots* che dal 2012 si prefigge l'obiettivo di vietare preventivamente l'impiego di sistemi d'arma letali in grado di selezionare e attaccare obiettivi senza alcun intervento umano. Il ruolo dei ricercatori come vedette etiche sembra essere quanto mai centrale oggi: basti pensare al documento pubblicato dalla U.S. *National Security Commission on Artificial Intelligence* in cui si dichiara che la commissione non supporta la proibizione di sistemi d'arma autonoma resi possibili da tecniche di IA (National Security Commission on Artificial Intelligence, 2021).

Naturalmente per regolare le armi autonome, definire cosa siano è essenziale. E proprio su questa definizione si gioca – a nostro parere – uno dei passaggi fondamentali dell'intero libro. In primo luogo Tamburrini evidenzia come la demarcazione fra un drone armato e un'arma autonoma non sia netta e mostra che il dibattito, ormai più che decennale, oscilla fra definizioni troppo ampie o troppo restrittive che sono ricostruite con grande chiarezza nel capitolo 5 del volume. In secondo luogo, Tamburrini mostra come tale definizione non possa prescindere dal considerare nel dettaglio gli sviluppi tecnologici, il modo in cui le armi autonome effettivamente funzionano, ciò che possono fare o non fare.

Questa relazione fra etica ed epistemologia, forse meno evidente ma altrettanto importante di quella fra etica e politica, deve essere adeguatamente sottolineata, riconoscendo che il dibattito etico è legato a doppio filo agli sviluppi tecnologici e non può prescindere da essi. Le armi autonome dipendono sempre più da architetture basate su apprendimento automatico e operano in ambienti caratterizzati da incertezza sia per la natura delle loro operazioni sia per l'interazione fra i sistemi artificiali e gli esseri umani. Questa incertezza epistemica gioca un ruolo chiave quando si voglia determinare in modo chiaro la responsabilità umana. I sistemi che utilizzano tecniche di apprendimento automatico (*Machine Learning*) per percepire l'ambiente soffrono di un problema di generalizzazione, ossia non sono in grado di adattarsi a condizioni al di fuori di un insieme ristretto di assunzioni. Per esempio, i sistemi di visione delle armi autonome possono compiere errori macroscopici a fronte di una perturbazione minima dell'immagine. Questa incertezza epistemica non caratterizza solo i sistemi d'arma autonoma ma tutte quelle applicazioni di IA in contesti critici dal punto di vista della sicurezza, come i trasporti o la medicina (Cummings, 2021). Al di là delle possibili soluzioni tecniche, come per esempio quella di utilizzare un numero maggiore di dati, è importante riconoscere la necessità di un cambio di paradigma a livello sia epistemico sia etico. Perché i sistemi critici dal punto di vista della sicurezza possano operare efficacemente in ambienti realistici occorre che siano dotati di capacità di ragionamento basate sia sui modelli (*top-down*) sia sui dati (*bottom-up*). Ciò non significa che tali sistemi di IA siano in grado di sostituire gli esseri umani nel loro operare, quanto piuttosto che gli esseri umani devono poter sempre esercitare un controllo significativo sulle operazioni di tali sistemi. Tuttavia, proprio quando si riconosce l'importanza di un controllo informato sui sistemi basati sull'apprendimento automatico, mediante approcci come quello della *eXplainable Artificial Intelligence*, varrebbe la pena presentare

alcuni casi ipotetici che sollevano indubbi dilemmi morali: basti pensare all'esempio di sistemi artificiali in grado di realizzare prestazioni migliori di quelle umane che però sono 'black box', ossia non spiegabili né trasparenti nel loro funzionamento, e quindi non soggetti al controllo informato.

Sebbene sia i droni sia le armi autonome di ultima generazione siano realizzati con tecnologie di IA e robotica, è l'aspetto del controllo a rendere diverse queste ultime. Infatti un drone è sempre controllato da operatori umani che, per quanto in remoto, intervengono in tutte le fasi principali del suo funzionamento. Un'arma autonoma, invece, è progettata per prendere decisioni di attacco autonomamente dagli operatori umani. In altre parole, nel caso di un'arma autonoma la catena di comando e controllo non rimane intatta, e le responsabilità connesse ad essa non sempre sono riconducibili a esseri umani.

L'enfasi sul controllo gioca un ruolo fondamentale anche nello spostare il dibattito dalla pura e semplice definizione di cosa sia un'arma autonoma a un terreno, almeno apparentemente, più promettente, ossia la relazione fra esseri umani e macchine. A partire dall'idea che sia necessario preservare la responsabilità umana per le decisioni che riguardano azioni potenzialmente letali per gli esseri umani, è stato proposto il concetto di *controllo umano significativo* (Article 36, 2013). L'uso dell'aggettivo "significativo" intende escludere condizioni e forme di controllo puramente nominali di un sistema d'arma, il che potrebbe accadere quando il controllore umano non ha a disposizione tempo o informazioni sufficienti per intervenire con cognizione di causa sul sistema d'arma stesso. L'idea del controllo umano significativo, per quanto non del tutto scevra da ambiguità, non solo è stata accolta con favore dalla comunità impegnata nel dibattito sulle questioni etiche connesse alle armi autonome, ma è stata proficuamente applicata anche in altri contesti relativi all'autonomia di sistemi robotici, come quelli dei VA (Santoni de Sio & Van den Hoven, 2018). Ciò non significa che l'applicazione del controllo umano significativo non sia priva di problemi. Basti leggere quanto Tamburrini scrive nel capitolo 6 a proposito della chirurgia robotica, mostrando come il requisito del controllo umano sia destinato a confrontarsi con nuovi e più difficili problemi, frutto anche dell'applicazione sempre più vasta di tecniche di apprendimento automatico che pongono sfide enormi rispetto all'interpretabilità e spiegabilità dei loro risultati. Non ultimo il problema di come sia possibile coniugare il controllo umano significativo nell'ambito della chirurgia robotica con i tradizionali principi della bioetica, come per esempio la beneficenza o la non maleficenza.

Tutto ciò contribuisce a rendere ancora più evidenti le responsabilità speciali dei ricercatori di IA e robotica che, come ben evidenzia Tamburrini, devono discutere preventivamente le questioni etiche in gioco e sviluppare politiche etiche conseguenti. Si tratta di sfide enormi, rese ancora più complesse dalla dimensione globale di cui occorrerà tenere conto, in modo sempre più attento, nel futuro prossimo. Su questo tema Tamburrini comincia a prospettare alcuni dei problemi globali, come quando discute la guida autonoma in relazione alla crisi climatica, che ancora trovano poco spazio nel dibattito corrente. Il merito qui è di tenersi lontano da aspettative eccessive ma dare conto della complessa articolazione di questi problemi che sono globali sia per la loro portata sia per l'ampiezza delle soluzioni che richiedono.

### **Bibliografia**

- Article 36, 2013, «Killer Robots: UK Government Policy on Fully Autonomous Weapons», *Article 36*, April 2019.
- Cummings M. L., 2021, «Rethinking the maturity of artificial intelligence in safety-critical settings», *AI Magazine*, 42(1), pp. 6-15.
- Lin P., 2015, «Why Ethics Matters for Autonomous Cars», in Maurer M., Gerdes J., Lenz B., Winner H. (eds) *Autonomes Fahren*. Berlin, Heidelberg, Springer Vieweg, pp. 69-85.
- Nilsson N. J., 2010, *The Quest for Artificial Intelligence: A History of Ideas and Achievements*, Cambridge, Cambridge University Press.
- McCarthy J. *et al.*, 1955, «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence»,  
<http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>  
(ultimo accesso maggio 2021).
- National Security Commission on Artificial Intelligence, 2021,  
Final Report,  
<https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf> (ultimo accesso maggio 2021).
- Russel S.-Norvig, P., 2020, *Artificial Intelligence: A Modern Approach*. Fourth Edition, Pearson Education (Intelligenza Artificiale. Un approccio moderno, trad. it., a cura di F. Amigoni, Milano-Torino, Pearson, 2021).
- Santoni de Sio F.-Van den Hoven J., 2018, «Meaningful Human Control Over Autonomous Systems: A Philosophical Account», *Frontiers in Robotics and AI*, 5, 15.

- Somalvico M., 1987, «Emula e non simula», in Jacobelli J. (a cura di) *Aspettando Robot. Il futuro prossimo dell'Intelligenza Artificiale*, Laterza, pp. 172-182.
- Taebi B., 2021, *Ethics and Engineering. An Introduction*, Cambridge, Cambridge University Press.
- Turing A. M., 1950, «Computing Machinery and Intelligence», *Mind*, 59, pp. 433-60.

---

**APhEx.it è un periodico elettronico, registrazione n° ISSN 2036-9972. Il copyright degli articoli è libero. Chiunque può riprodurli. Unica condizione: mettere in evidenza che il testo riprodotto è tratto da [www.aphex.it](http://www.aphex.it)**

Condizioni per riprodurre i materiali --> Tutti i materiali, i dati e le informazioni pubblicati all'interno di questo sito web sono "no copyright", nel senso che possono essere riprodotti, modificati, distribuiti, trasmessi, ripubblicati o in altro modo utilizzati, in tutto o in parte, senza il preventivo consenso di APhEx.it, a condizione che tali utilizzazioni avvengano per finalità di uso personale, studio, ricerca o comunque non commerciali e che sia citata la fonte attraverso la seguente dicitura, impressa in caratteri ben visibili: "www.aphex.it". Ove i materiali, dati o informazioni siano utilizzati in forma digitale, la citazione della fonte dovrà essere effettuata in modo da consentire un collegamento ipertestuale (link) alla home page [www.aphex.it](http://www.aphex.it) o alla pagina dalla quale i materiali, dati o informazioni sono tratti. In ogni caso, dell'avvenuta riproduzione, in forma analogica o digitale, dei materiali tratti da [www.aphex.it](http://www.aphex.it) dovrà essere data tempestiva comunicazione al seguente indirizzo ([redazione@aphex.it](mailto:redazione@aphex.it)), allegando, laddove possibile, copia elettronica dell'articolo in cui i materiali sono stati riprodotti.

In caso di citazione su materiale cartaceo è possibile citare il materiale pubblicato su APhEx.it come una rivista cartacea, indicando il numero in cui è stato pubblicato l'articolo e l'anno di pubblicazione riportato anche nell'intestazione del pdf. Esempio: Autore, *Titolo*, <<[www.aphex.it](http://www.aphex.it)>>, 1 (2010).

---