

MARIAROSARIA TADDEO

COSTRUIRE L'ETICA DELL'INTELLIGENZA ARTIFICIALE

Altro che bufale! Mentre ci interroghiamo sull'impatto che una comunicazione fuori controllo può avere sulla nostra visione del mondo e sulle decisioni che prendiamo ogni giorno, la nostra società è davanti ad una sfida ben più grande che potrebbe trasformarla per sempre. L'uso sempre più pervasivo in tutti i settori dell'Intelligenza Artificiale (IA) cambia la nostra percezione di quello che facciamo, delle nostre scelte, dal lavoro alla guerra, dalla ricerca all'aiuto domestico. Di robot, algoritmi e nuovi sistemi tecnologici e delle sfide non solo scientifiche ma innanzitutto etiche si occupa da anni Mariarosaria Taddeo, filosofa che lavora all'Università di Oxford dove è vicedirettrice del Digital Ethics Lab e professore associato all'Oxford Internet Institute.

Partiamo da una cornice generale per chi non ha le idee molto chiare sull'Intelligenza artificiale, prima di affrontare i temi etici che ne riguardano lo sviluppo: che cosa è IA e che cosa non è?

L'espressione 'intelligenza artificiale' trae in inganno, perché evoca scenari fantascientifici, come quelli di *Blade Runner* o *Her*. Ma questa interpretazione non ha spazio al di fuori della letteratura o della cinematografia fantascientifica. Chi progetta sistemi di intelligenza artificiale e chi ne studia le implicazioni concettuali e legali ed etiche sa che si tratta di tutt'altra cosa, meno avveniristica ma non meno complessa ed interessante. Per Intelligenza artificiale non intendiamo altro che macchine che si comportano come se

fossero intelligenti. Per essere più precisi, seguendo la definizione che ne dà il filosofo Luciano Floridi: «L'IA è una risorsa di agenti autonomi, capaci di interagire e di imparare e che possono essere usati per eseguire compiti che altrimenti richiederebbero l'intelligenza umana per essere eseguiti con successo». Dunque l'IA non è davvero intelligente, nel senso in cui lo è un essere umano, ma è la prima tecnologia nella storia dell'umanità che impara dalle sue interazioni con l'ambiente e interagisce con esso in modo 'autonomo'. Questa autonomia e la capacità di imparare sono all'origine delle opportunità e dei rischi che l'IA pone.

Leggendo documenti e commenti sulle nuove prospettive che l'IA apre, non è sempre evidente come dobbiamo considerarla: un'opportunità o una sfida incerta e preoccupante?

L'IA è uno strumento potentissimo, che offre tantissime opportunità e pone rischi seri. L'impatto sarà più positivo o negativo a seconda dell'uso che come individui e società ne faremo. Direi che forse la domanda da porsi è come facciamo ad usare l'IA in modo da averne più vantaggi che svantaggi.

Come facciamo, qual è la risposta?

Credo che una governance etica dell'IA offra una parte importante della risposta; l'altra parte importante riguarda il progetto di società post-IA che vogliamo sviluppare. Guardiamo prima alla governance e partiamo di rischi. Mi viene in mente il famoso caso di Tay, il Twitter bot lanciato da Microsoft a marzo 2016 e ritirato 16 ore dopo il lancio, perché aveva iniziato a diffondere messaggi razzisti, nazisti e misogini. Tay è un esempio eclatante delle cosiddette 'conseguenze non previste'. Allo stesso tempo, la stessa autonomia dell'IA ci permette di identificare correlazioni tra grandi gruppi di dati che altrimenti sarebbe difficile trovare e che possono essere di grande aiuto a supporto di decisioni pre-

se in diversi ambiti, dalla scienza alla politica. Si pensi al recente successo dell'IA di DeepMind nel definire la forma delle proteine e dell'impatto che questo avrà sulla ricerca e le cure mediche. È chiaro che non possiamo passare la mano su una tecnologia che offre questo tipo di supporto alle attività umane, ma è altrettanto evidente che per usare con successo questa tecnologia abbiamo bisogno di regole e protocolli che ci permettano di mantenere il controllo su di essa e di intervenire quando le cose non vanno nella direzione giusta. Per questo ci serve una governance che sia in grado di guidare lo sviluppo e gli usi di questa tecnologia e che li faccia collimare con i principi etici fondanti della nostra società.

Poi c'è la questione più ampia e più difficile da immaginare, ovvero di quale sia la società che vogliamo costruire con l'aiuto dell'intelligenza artificiale.

Esattamente. La sfida più grande è ragionare sul nostro progetto di società in cui l'IA non sarà più un fattore innovativo, ma una tecnologia essenziale, totalmente inglobata nelle infrastrutture e negli ambienti e da cui dipenderà il funzionamento della società stessa. Sarà un po' com'è adesso per noi la corrente elettrica. Per questo il progetto va definito oggi. Oltre ad essere autonoma, l'IA è anche una tecnologia trasformativa: non solo cambia il modo in cui facciamo le cose, ma cambia il modo in cui concepiamo il mondo e le nostre interazioni con esso. Si pensi, per esempio, ai concetti di responsabilità e a quello di lavoro. Abbiamo per millenni pensato alla responsabilità morale in uno schema uno-da-uno, ad un'azione corrisponde un agente che è responsabile delle conseguenze di quell'azione. Pensiamo ora a una macchina a guida autonoma: ci sono decine e decine di sistemi di IA che seguono i diversi aspetti del funzionamento della vettura. I sistemi interagiscono con l'ambiente

e con gli altri sistemi della macchina e imparano da queste interazioni. Se uno di questi sistemi commette un errore che influenza tutti gli altri portando ad una conseguenza disastrosa, a chi attribuiamo la responsabilità? Al programmatore, alla sua azienda, a tutte le aziende coinvolte, alla proprietaria della macchina? Allo stesso modo, si parla molto dell'impatto dell'IA sul lavoro. Non ci sono ancora dati che ci permettano di fare stime concrete, ma le nostre società oggi sono fondate sul lavoro: se le macchine lavoreranno per noi, come vorremmo che cambiassero le nostre società? La risposta a queste domande prevede una visione strategica, *politica* oserei dire, che riguarda il tipo di società che vogliamo sviluppare e come usare l'IA per arrivarci.

L'avvento dell'IA nella vita di tutti i giorni è ineludibile per tutti i vantaggi che può portare, ma come ci difendiamo dai rischi che ci sfugga di mano? Non soltanto correggendo gli errori quando ce ne accorgiamo?

Il rischio che l'intelligenza artificiale ci sfugga di mano è spaventoso, ma non è realistico. Ne discutevano, in un famoso scambio su «Science», Wiener e Samuel già nel 1960. In quello scambio Wiener sosteneva che le macchine possono essere 'originali', sorpassare e sorprendere i loro creatori con conclusioni e comportamenti inaspettati, cioè che «le macchine ci sfuggano di mano». Samuel rifiuta la tesi di Wiener e inizia il suo articolo così: «La macchina non possiede una volontà, e le sue cosiddette 'conclusioni' sono solo le conseguenze logiche del suo input, come rivelato dal funzionamento meccanicistico di un assemblaggio in-animato di parti meccaniche ed elettriche». Io concordo con Samuel. Ma penso che nella sua analisi Wiener abbia evidenziato un problema molto importante che ancora oggi non è risolto: come mantenere il controllo sugli usi e le conseguenze di questi usi. È quello che in contesti di difesa

si chiama *meaningful control*. Il *meaningful control* richiede la presenza di un essere umano in grado di capire quando l'IA sbaglia ed intervenire prontamente. Il concetto può essere esteso fino ad includere la definizione di criteri per la scelta dei compiti che si delegano all'IA, dei rischi che si vogliono correre, delle misure di appello e risarcimento che si prevedono in caso di errore. In altre parole, il controllo dell'IA non passa solo per il controllo della tecnologia, ma riguarda anche il controllo dei processi attraverso cui l'IA è progettata e usata.

E i vantaggi, di cui parlava prima, non sono da sottovalutare.

Consideriamo, per esempio, l'uso dell'IA nella ricerca per il vaccino contro il Covid-19. In questo caso l'IA aiuta i ricercatori a comprendere il virus e la sua struttura e a prevedere quale dei suoi componenti provocherà una risposta immunitaria. Può aiutare gli scienziati a scegliere gli elementi di potenziali vaccini e capire meglio i dati sperimentali. Il successo recente dell'IA di DeepMind nel definire la forma 3D di una proteina dalla sua sequenza di amminoacidi è un altro esempio del grandissimo potenziale di questa tecnologia nel supportare gli sforzi dell'umanità nel comprendere la realtà che ci circonda e nel supportare i tentativi per migliorare la nostra condizione e proteggere l'ambiente. Dunque, non possiamo prescindere da una tecnologia che ci offre questi vantaggi ma abbiamo bisogno di regole e protocolli che ci permettano di mantenere il controllo su di essa e di intervenire quando le cose non vanno nella direzione giusta.

A proposito di società e di influenza dell'Intelligenza artificiale sui comportamenti dei singoli o di gruppi ci sono esperimenti impressionanti in cui l'IA – attraverso informazioni false e pressioni psicologiche – potrebbe arrivare

a indurre addirittura a comportamenti violenti o autodistruttivi, senza che la vittima se ne accorga. Magari sono ipotesi fantascientifiche, ma pongono il tema delle possibilità di manipolazione dell'IA su di noi.

L'IA è uno strumento potentissimo e va usato con cautela. Esistono gruppi pro-anorexia o pro-suicidio sui social media e algoritmi di IA comunemente usati dai social media che possono attrarre utenti verso questi gruppi. È un chiaro caso di uso non etico dell'IA. L'IA è un facilitatore invisibile che media le nostre interazioni con gli altri e con l'ambiente, in un modo conveniente e appena percettibile. Più questa tecnologia è diffusa, più diventa una parte strutturale delle nostre attività e degli ambienti in cui viviamo, più la sua mediazione si fa invisibile ed estesa. Qui è la nostra autodeterminazione che rischia di essere erosa. Per questo è importante capire quali scelte deleghiamo all'IA come individui e come società e come facciamo a preservare la nostra autodeterminazione e sfruttare i vantaggi che l'IA ci offre.

Il cammino delle scoperte legate all'IA è pieno anche di inciampi: Cambridge Analytica, per citarne uno su tutti. Non è un prezzo un po' troppo alto?

Cambridge Analytica non è un inciampo nello sviluppo dell'IA, bensì un esempio di una cattiva governance a cattivo uso di questa tecnologia e dei danni che questo tipo di uso possono determinare. Tutte le tecnologie hanno un potenziale positivo ed uno negativo (ma attenzione, non è vero che la tecnologia è eticamente neutra), e i due potenziali sono tanto maggiori quando la tecnologia è trasformativa. Si pensi al nucleare. Con il nucleare abbiamo capito subito (purtroppo anche sperimentato) il potenziale negativo e abbiamo creato regole per arginarne i rischi; allo stesso tempo abbiamo sviluppato una ricerca che utilizza la fisica subatomica per sviluppare nuove forme di energia e miglio-

rare la diagnostica. Con l'IA ci siamo distratti guardando ai rischi fantascientifici e non ci siamo accorti di problemi meno eclatanti ma con impatto concreto sulle nostre società e, come nel caso di Cambridge Analytica, sulle nostre democrazie. Pensavamo a Terminator e non ci siamo accorti che gli smartphone che portiamo nelle nostre tasche venivano usati per manipolare il nostro voto. Direi, quindi, che per certi versi Cambridge Analytica è il prezzo della nostra distrazione.

Fin dove ci si può spingere nel campo della medicina e della salute senza contravvenire ad un'etica condivisa in tema di scoperte e di utilizzo dell'IA sugli esseri umani?

Una recente analisi comparativa dei principi etici per l'uso dell'IA ha mostrato che esiste una convergenza sui principi fondamentali della bioetica: i principi di beneficenza, di non-maleficenza, di autonomia e di giustizia. È importante notarlo, perché se questi principi ci forniscono indicazioni rispetto agli usi etici dell'IA in generale, sono ancora più determinanti nel campo della medicina. Non possiamo delegare all'IA decisioni che riguardino la salute o la vita delle persone, ma possiamo (si potrebbe anche argomentare che dobbiamo) fare leva sul potenziale dell'IA come strumento per migliorare il modo in cui queste decisioni sono prese da esseri umani. Credo che sia chiaro che un sistema di intelligenza artificiale non possa né debba sostituire un medico. La questione è come integrare i sistemi IA nei processi clinici, per esempio per fare in modo di ottimizzare le cure offerte. Questa integrazione deve garantire autonomia agli esperti ed è – quindi – importante preservare la possibilità degli esperti di dissentire rispetto ad una decisione suggerita dall'IA. Allo stesso tempo, è necessario chiarire quali siano i compiti da delegare all'IA; come si può assicurare che la fiducia dei pazienti nei medici non sia tradita; e come

garantire che l'autonomia e la dignità dei pazienti non siano violate. Senza considerare le questioni relative alla trasparenza e all'accountability.

Con l'IA si può fare – ha detto lei di recente – anche una guerra senza spargimento di sangue. È chiaro che lì per lì è meno cruenta, ma sul lungo periodo serve davvero a risparmiare vite umane?

Ci sono due modi in cui l'IA può supportare operazioni di difesa nazionale: nel contesto puramente cyber (tecnicamente: uso non cinetico) e contesto operativo del campo di battaglia (uso cinetico). Partiamo dagli usi non-cinetici. L'intelligenza artificiale è pronta a rivoluzionare questa attività. Luciano Floridi spiega che il digitale separa concetti che prima erano uniti e unisce concetti che prima erano separati. I conflitti cyber (conflitti non-cinetici che usano le tecnologie del digitale) offrono un esempio chiaro di questo 'taglia e incolla'. Per millenni le guerre sono state concepite come attività aggressive veicolate con la violenza. Ora con l'IA come ultima innovazione digitale abbiamo tagliato questi due concetti: la guerra non è più violenta. Ma non per questo, la guerra ha smesso di essere aggressiva, visto che può ancora essere ingiusta e tornare facilmente ad essere violenta. Come ho evidenziato nella mia ricerca, la guerra cibernetica pone tre categorie di problemi: *risks*, *rights*, e *responsibilities*, che chiamo le 3R. I rischi riguardano il potenziale aumento del numero di conflitti e vittime. Il vantaggio dei conflitti basati sulle tecnologie digitali di essere incruenti ha l'inconveniente di rendere la guerra meno problematica per dispiegamento di forze e quindi più facile per gli Stati impegnarsi in conflitti basati sul digitale. Paradossalmente, se non regolamentata, la guerra cibernetica potrebbe – come per certi versi ha fatto finora – portare ad una crescita del numero dei conflitti e quindi minare la stabili-

tà internazionale. Oramai molti Paesi prevedono di implementare l'IA per la difesa informatica nazionale, generando una corsa agli armamenti informatici che potrebbe creare un circolo vizioso che rischierebbe di portare anche ad un attacco fisico. Per quanto riguarda i diritti, i conflitti basati sulle tecnologie digitali sono pervasivi perché non solo possono attaccare infrastrutture civili, ma possono anche essere lanciati attraverso computer e siti web civili. Ciò potrebbe portare a livelli più elevati di controllo imposto dai governi al fine di difendere i propri cittadini da possibili attacchi. In questo caso, i diritti della libertà individuale, della privacy e dell'anonimato possono essere sottoposti a forti pressioni.

C'è poi il problema dei problemi, quello della responsabilità: chi risponde delle 'scelte' compiute dall'IA?

Quando si utilizzano armi (semi) autonome per attacchi informatici, è difficile sia identificare e punire eventuali aggressori, sia attribuire responsabilità per l'uso improprio dei sistemi digitali. Servono regole chiare e applicate severamente. Il cyberspazio è un dominio di guerra e l'IA è una nuova capacità di difesa. I regolamenti sono quindi necessari per l'uso statale dell'IA, come lo sono per altri settori militari: aria, mare, terra e spazio. Sono altresì necessari criteri per determinare risposte proporzionate agli attacchi, nonché stabilire soglie chiare o 'linee rosse' per distinguere gli attacchi informatici legali e illegali e applicare sanzioni adeguate per gli atti illegali. Credo che adesso spetti a forum regionali, come la Nato e l'Unione Europea ma anche l'Asean, rilanciare gli sforzi (dopo il fallimento del gruppo di esperti delle Nazioni Unite) e preparare il terreno per un'iniziativa guidata dalle Nazioni Unite.

Le armi intelligenti si usano anche sul campo di battaglia, come ricordava lei prima.

Si tratta del cosiddetto uso cinetico, che ci fa pensare ad armi autonome che decidono se e come colpire bersagli fisici anche umani. Nel dibattito che si è sviluppato negli ultimi vent'anni si sono cristallizzate due posizioni: quella di coloro che definiscono immorale l'uso di armi autonome perché non si può delegare la decisione di colpire bersagli umani ad una macchina: è una scelta che lede la dignità umana, il valore militare di coloro che combattano e pone problemi relativi a principi della Teoria della Guerra Giusta (la teoria etica che sta a fondamento delle leggi internazionali che regolano i conflitti), senza contare i problemi relativi all'attribuzione di responsabilità e al controllo limitato dell'IA, che abbiamo già discusso. Ma c'è anche chi sostiene che l'uso di armi autonome permetterebbe di salvare le vite delle donne e degli uomini impegnati nei conflitti e favorirebbe guerre brevi. Le ragioni di questa posizione sono valide, ma deboli rispetto a quelle di chi spinge per vietare le armi autonome. Anche se il dibattito è ancora aperto, dal 2013 già trenta Paesi hanno chiesto il divieto di armi completamente autonome: Algeria, Argentina, Austria, Bolivia, Brasile, Cile, Cina, Colombia, Costa Rica, Cuba, Gibuti, Ecuador, Egitto, El Salvador, Ghana, Guatemala, Santa Sede, Iraq, Giordania, Messico, Marocco, Namibia, Nicaragua, Pakistan, Panama, Perù, Stato di Palestina, Uganda, Venezuela e Zimbabwe. La Cina preferirebbe un trattato per vietare l'uso ma non la loro produzione, il che non sorprende dato che è tra le nazioni più avanzate nello sviluppo di tali armi.

Molti esperti condividono la sua l'analisi dei rischi e delle opportunità dell'IA e suggeriscono che molti di questi problemi potrebbero essere risolti usando 'etica-by-design'. Che cos'è l'etica-by-design? Può funzionare?

È un'area di ricerca che si propone di identificare, già nella fase di progettazione – by design, appunto –, i principi etici

secondo cui le macchine dovrebbero agire per fare in modo che siano disegnate per rispettare questi principi di default. Se pensiamo alle App sviluppate durante la pandemia per tracciare i contagi da Covid-19, noteremo che usano l'Api fornita da Google e Apple e adottano il protocollo DP-3T, cioè proteggono la privacy degli utenti by design. Sono state progettate in modo che i rischi della violazione della privacy siano bassissimi e che le App stesse non siano facilmente manipolabili per invadere la privacy degli utenti. Dalla bicicletta alla forma delle posate, il design ha forti implicazioni etiche, perché ci permette di fare o non fare certe cose con gli artefatti tecnologici. Ma rischia di essere un approccio molto limitato quando si considera l'IA. Per due ragioni: la prima è che l'IA impara dalle sue interazioni con l'ambiente e modifica i suoi comportamenti in base a quello che impara (ricordiamo il caso di Tay). Dunque è facile pensare che sistemi IA, che sono stati progettati con principi di etica by design, possano comunque mostrare comportamenti problematici dal punto di vista etico una volta messi in uso. La seconda ragione è che l'IA opera in contesti complessi, prendendo decisioni potenzialmente di grande impatto sugli individui e sulle società, e lo fa elaborando (senza molta trasparenza) grandi quantità di dati. Usare principi etici in fase di progettazione è necessario ma non sufficiente a garantire che il sistema sviluppi necessariamente soluzioni etiche.

A chi spetta di occuparsi dell'etica dell'IA, agli scienziati o ai filosofi? Come se ne costruisce una?

Ho sempre sostenuto che l'etica del digitale, non solo dell'IA, sia un ambito di ricerca interdisciplinare. Quindi se penso al tavolo sul quale si scrive l'etica del digitale, lo vedo popolato di esperte ed esperti di diverse discipline. Dell'etica si occupano i filosofi, e a ragione. L'etica, senza un'opportuna riflessione filosofica, rischia di diventare mero buon senso.

Direi che i filosofi guidano la conversazione. Ma accanto ai filosofi servono anche i tech providers (ingegneri, sviluppatori, informatici), perché è importante capire quali sono le capacità tecniche di cui stiamo parlando, i limiti che si possono valicare e quali no. Ma non basta, servono anche i policy-makers. L'etica del digitale non è fatta per essere stampata in un libro e dimenticata, ma serve a dare forma ai progetti di governance delle società del digitale. I policy-makers hanno il dovere di ascoltare filosofi e scienziati, che devono saper identificare rischi e opportunità ed indicare anche gli approcci etici per risolverli. Servono anche i politici per definire strategie che facciano leva sul digitale per migliorare le nostre società – farne società democratiche, pluraliste, aperte, giuste – e proteggere l'ambiente. Ovviamente, il club degli esperti in digital ethics non è esclusivo, possiamo immaginare altri partecipanti.

Qualche tentativo è stato fatto, come il codice etico dell'Unione europea del 2019: è una prima risposta sufficiente o in che cosa andrebbe modificato o aggiornato? È l'unico esempio di codice etico della materia?

Due anni fa, due miei colleghi del Digital Ethics Lab (University of Oxford) hanno fatto un'analisi comparativa dei principi etici per l'IA. L'analisi identificò sei documenti principali. Un'analisi più recente ne ha identificati 83. In parte, questo è un successo. Negli ultimi due o tre anni c'è stata una forte sensibilizzazione ai problemi dell'IA, si è capito che senza una governance etica sarà impossibile adottare questa tecnologia con successo. Le linee guida europee sullo sviluppo ed uso dell'IA vanno in questo senso. Hanno uno scopo simile anche i principi proposti dell'Ocse, ma, come questi, rimangono abbastanza astratti. Ho criticato in passato la scelta dell'Unione europea di focalizzarsi sulla 'fiducia' nell'IA come elemento chiave per farne usi posi-

tivi. Indubbiamente la fiducia degli utenti nella tecnologia è importante per favorirne l'adozione. Tuttavia, la definizione e lo sviluppo di linee guida, standard e procedure di certificazione con l'obiettivo di sviluppare un'IA affidabile è concettualmente fuorviante e può comportare gravi rischi? per le nostre società. Possiamo definire la fiducia come la decisione di delegare un compito, senza alcuna forma di controllo o supervisione sul modo in cui il compito viene eseguito. Se si considerano la mancanza di trasparenza e le capacità di apprendimento dei sistemi IA, è chiaro che delegare senza controllare può essere molto rischioso e portare anche ad un rifiuto della tecnologia, qualora dovessero manifestarsi incidenti gravi. Questo non vuol dire che non dovremmo delegare compiti all'IA, soprattutto quando l'IA dimostra di essere in grado di eseguirli in modo efficiente ed efficace. Tuttavia, alcune forme di controllo sono necessarie per mitigare i rischi. Io credo che piuttosto che promuovere lo sviluppo di un'IA affidabile, ci si debba concentrare sullo sviluppo di prassi affidabili, che garantiscano trasparenza, accountability, misure per ricorsi in caso di errori, e ottimizzazione delle risorse.

Come si può usare L'IA per costruire una società più giusta? Può far alcuni esempi? Quali possono essere gli usi affidabili a patire dai quali si può pensare lo sviluppo dell'IA nel futuro prossimo?

Per esempio, potremmo incentivare iniziative che usano l'IA per il bene comune. *Al for Good* è oramai un ambito di sviluppo dell'IA abbastanza diffuso. Lo sviluppo delle tecnologie dell'IA è stato guidato principalmente da interessi commerciali. Tuttavia, il numero di progetti non commerciali che sfruttano l'IA in tutto il mondo è proliferato. Si va da modelli di machine learning, che prevedono il rischio di suicidio, a quelli che calcolano il rischio d'incarcerazione, come ci rac-

conta Bauman. L'IA offre opportunità senza precedenti in molti domini e potrebbe rivelarsi di grande importanza in un momento in cui i problemi sono sempre più globali, complessi e interconnessi. L'IA può fornire un supporto chiave per migliorare i risultati di salute e mitigare i rischi ambientali. Ma anche questi usi rischiano di essere sminuiti se prima non si saranno risolti i problemi etici legati all'IA e non si sarà compreso come regolamentarne gli usi per il bene comune. Non possiamo immaginare di usare l'IA a fin di bene, per esempio per garantire cure mediche a tutti i cittadini di Paesi con infrastrutture insufficienti, se non siamo sicuri di aver sradicato il pregiudizio di genere o aver garantito la trasparenza di questa tecnologia. Il rischio è altrimenti di usare – seppur a fin di bene – tecnologie che possono creare danni e problemi agli utenti. Questo è stato il caso scoperto nel 2019 di un algoritmo ampiamente utilizzato negli ospedali statunitensi per allocare l'assistenza sanitaria ai pazienti e che discriminava sistematicamente le persone afroamericane. L'IA rispecchia le nostre società e da sola non può migliorarle. Sta agli esseri umani usarla come strumento per sviluppare società più tolleranti, aperte e giuste.