

I VALORI E IL MONDO NUOVO

«La tecnologia non è quasi mai neutra Serve un codice per regolamentare l'IA»

di CIRO DE FLORIO e E GIUSEPPE RIVA



PROFESSORE ALLA CATTOLICA

Cesare De Florio è professore di Logica e Filosofia della scienza e docente di Logica, razionalità e decisioni presso la facoltà di Economia alla Cattolica

Ogni tecnologia ridisegna il mondo. Proprio per questo, la tecnologia non è quasi mai neutra da un punto di vista valoriale: essa incide sulla realtà e ogni azione causale non può che essere accompagnata da una certa rilevanza morale. L'etica della tecnologia ha a cuore proprio l'analisi normativa della progettazione, dell'implementazione e dell'uso dei manufatti tecnologici. E non c'è dubbio che, da circa una ventina d'anni, l'agenda degli esperti in etica della tecnologia sia dominata da due parole: Intelligenza Artificiale (IA). Per questo, negli ultimi mesi sia negli Usa ([whitehouse.gov/ostp/ai-bill-of-rights](https://www.whitehouse.gov/ostp/ai-bill-of-rights)), che da parte della Santa Sede (romecall.org/the-call) e dall'Europa (artificialintelligenceact.eu) è stata formulata una serie di principi etici con cui provare a rispondere alle molte domande portate dalla diffusione di questa tecnologia. E non mancano i tentativi di mettere ordine enucleando, al di là delle varie e differenti formulazioni, i *core principles* dell'etica dell'IA.

In generale all'IA sono associate due grandi narrazioni che presentano talvolta aspetti polarizzanti. Da un lato, ed è la narrazione «positiva», l'IA è intesa come il trionfo della ragione umana, la creazione

della ricerca etica in IA è costituito dalla trasparenza degli algoritmi. Più i software sono intelligenti, autonomi, adattivi, più diventa difficile (quando non impossibile) capire «dall'esterno» i meccanismi in base ai quali le informazioni vengono analizzate. Gli esempi sono ben noti e discussi: ci fideremo e accetteremo la sentenza di un giudice artificiale senza che ci spieghi come è arrivato alla sua decisione?

Il riflesso mediatico dell'importanza dell'etica nell'IA riguarda, in gran parte, la rilevanza normativa di sistemi di software e della gestione delle informazioni: profilare un imputato, selezionare un candidato sulla base di una predizione circa la sua produttività, diagnosticare una certa malattia o la propensione ad ammalarsi sono tutte azioni che riguardano l'elaborazione di informazione. I sistemi di IA possono però fare molto di più, possono agire concretamente nel nostro mondo, danneggiando o salvando vite umane, sollevando dalla fatica fisica o relegando gli esseri umani a spettatori. È l'unione della robotica con l'IA che apre campi in buona parte inesplorati e le cui conseguenze di carattere etico, economico, politico e sociale potrebbero essere dirompenti. L'interazione con i robot, infatti, introduce una serie di problemi che i software non hanno: l'agency di un robot e l'agency di un umano nei confronti di un robot è mediata in maniera irrinunciabile da un'interazione fisica. Possiamo toccare ed essere toccati dai nostri androidi.

Questo impone un'analisi etica a grana più fine, in cui tener conto della possibilità di danno psico-fisico (sembra di sentire la prima legge della robotica di Asimov: «Un robot non può recar danno a un essere umano e non può permettere che, a causa di un suo mancato intervento, un essere umano riceva danno»), del legame fiduciario che si viene a creare con un robot, dei sentimenti di avversione che potrebbero sorgere in una società pesantemente robotizzata (schiavi o rivali?).

Ma non tutti i robot, ovvero non tutti i dispositivi tecnologici dotati di sistemi di IA, assomigliano agli esseri umani. Ci sono all'orizzonte macchine (semi)automatiche il cui funzionamento è già sotto la lente degli esperti di etica di IA. Pensiamo, come esempi, alle auto a guida autonoma e alle armi automatiche; anche in questo caso ciò che è rilevante per la considerazione etica sono diversi (nuovi?) concetti di agency, di controllo, di autonomia.

Come dicevamo, il dibattito pubblico e, in certa misura, anche l'agenda politica vive di narrazioni episodiche: una diagnosi geniale ottenuta con sistemi di IA che ha salvato la vita di un bambino contrapposta, o forse solo giustapposta, all'incidente che ha visto protagonista un'auto a guida autonoma. Ma quello di cui la rivoluzione digitale e l'avvento dell'IA hanno invece bisogno non sono racconti ma sguardi razionali sul mondo, basati su un approccio «umano», integrato e multidisciplinare, che unisca la conoscenza degli aspetti tecnologici con quella dei processi e dei contesti in cui l'IA e la robotica verranno utilizzati. Senza questo doppio sguardo, il rischio di perdere la sfida etica è purtroppo molto elevato.



DIRIGE LO HUMANE TECHNOLOGY LAB

Giuseppe Riva è professore ordinario di Psicologia generale all'Università Cattolica del Sacro Cuore dove dirige lo Humane Technology Lab

L'ETICA



Più i software sono intelligenti, autonomi, adattivi, più diventa difficile capire dall'esterno i meccanismi in base ai quali le informazioni vengono analizzate

ne di ciò che ci distingue dal resto del mondo naturale: l'IA conquista territori che erano, fino a ieri, appannaggio esclusivo dell'intuito, dell'ispirazione, del pensiero umano. Esiste però un'altra narrazione, diremmo più umanocentrica, che si interroga non tanto sul destino ingegneristico dell'IA quanto, piuttosto, sull'impatto di questa tecnologia sulle varie dimensioni della nostra esperienza: dal lavoro alle relazioni interpersonali.

Queste due narrazioni sono largamente indipendenti e restituiscono immagini differenti della rivoluzione digitale. Quello che sembra però essenziale per governare questo fenomeno di portata così ampia è proprio la riunificazione di queste prospettive. Da qui, l'esigenza di una riflessione plurale sulla IA e sui suoi effetti.

Tuttavia, l'obiettivo non è semplice perché l'IA agisce con un grado di autonomia e indipendenza mai osservato in precedenza, facendo emergere una nuova categoria di attori: gli «agenti artificiali».

Le principali differenze tra le carte di principi circa l'IA elaborate negli ultimi anni stanno proprio qui, nella differente relazione di controllo che abbiamo nei confronti di dispositivi che apprendono, elaborano e agiscono in maniera slegata (o parzialmente slegata) dai nostri intenti particolari. E conseguentemente diventa prioritario poter ispezionare, tracciare, indagare i processi in base ai quali le macchine «decidono» di fare determinate cose.

Infatti, insieme all'agency, l'altro grande pilastro



L'unione della robotica con l'IA apre campi inesplorati, le cui conseguenze di carattere etico, economico, politico e sociale potrebbero essere dirompenti

COME BUSSOLA