

*Ripensare l'umano nell'era delle tecnologie digitali*

Milano, 18 febbraio 2025

---

*Perché abbiamo bisogno della filosofia in un'epoca dominata dalla IA*

Prof.ssa Viola Schiaffonati

Recentemente l'intelligenza artificiale (IA) è entrata a fare parte in maniera sempre più profonda e silenziosa del tessuto delle nostre società: come facciamo affari, le amicizie che abbiamo, il modo in cui vediamo il mondo, il modo in cui siamo spinti – anche se gentilmente – a comprare quel prodotto, vedere quel film, ascoltare quella canzone, fare quella scelta. In altre parole, l'IA sta invisibilmente scivolando in ogni ambito delle nostre vite. Tuttavia, se vogliamo costruire un futuro in cui il grande potere dell'IA sia davvero a beneficio dell'umanità e condiviso da tutti, occorre smettere di guardare all'IA come esclusivamente a un tema tecnico e considerare una prospettiva più ampia, quella della filosofia per dare forma al tipo di società in cui vogliamo vivere. Per far ciò, il primo passaggio è di mettere in luce alcune peculiarità di queste tecnologie; il secondo è di allargare lo sguardo per includere non solo l'etica ma la filosofia più in generale.

Cercare di esaurire la domanda relativa a cosa l'IA ha di peculiare è un compito ambizioso. In questo contesto possiamo limitarci ad alcuni aspetti, oltre a quelli più noti e generalmente discussi come il fatto che i sistemi di IA, o almeno alcuni di essi, imparano dall'esperienza, sono autonomi, intelligenti e così via. La prima peculiarità ha a che fare con il *fattore di invisibilità*, termine coniato dall'eticista Jim Moore, in uno scritto del 1985, in cui cercava di capire perché i computer sono così rivoluzionari per giungere alla conclusione che un aspetto davvero distintivo riguarda l'invisibilità delle loro operazioni. Il fattore di invisibilità non è quindi peculiarità dei sistemi di IA. Tuttavia se consideriamo i tipi di invisibilità descritti da Moore – invisibilità di

*abuso*, invisibilità dei *valori di programmazione* e invisibilità di *calcolo* – possiamo chiaramente vedere che questi tre fattori non solo sono evidenti oggi nelle attuali tecnologie di IA ma sono anche esacerbati e aumentati da queste.

La seconda peculiarità è che le tecnologie di IA, o almeno alcune di esse, sono *tecnologie sperimentali*, così come lo sono per esempio le nanotecnologie. Le tecnologie sperimentali possono essere definite come tecnologie i cui benefici e rischi non sono facilmente valutabili sulla base dell'esperienza passata. In altre parole, benefici e rischi emergono solo quando queste tecnologie sono inserite nel loro contesto d'uso. Ciò significa che può essere difficile, o addirittura impossibile, modificarle e addirittura decidere di non usarle più. Qualche esempio. I sistemi di supporto o di sostituzione alle decisioni umane, da quelli impiegati nell'ambito delle assunzioni lavorative fino ai sistemi di polizia predittiva con i loro pregiudizi nascosti, le possibili discriminazioni e le ingiustizie. I sistemi di sorveglianza potenziati da tecniche di IA come la sorveglianza facciale o il riconoscimento biometrico. I sistemi d'arma autonomi con la licenza di uccidere esseri umani. Per arrivare alla rivoluzione nel mondo del lavoro e delle professioni con la massiccia introduzione dell'IA generativa per supportare e sostituire mansioni e processi fino a ora squisitamente umani. Questioni assai rilevanti che – chiaramente – sono etiche e politiche e, solo in misura minore, tecniche.

Dato questo panorama occorre quindi ripensare al concetto di responsabilità morale, e in par-

ticolare, alla nozione tradizionale di responsabilità *passiva* (ricostruire la catena della responsabilità dopo che si è verificato un evento negativo inaspettato), tradizionalmente impiegata nell'etica della tecnologia, per muoversi verso la responsabilità *attiva*: evitare effetti negativi ma anche promuovere effetti positivi, anticipando per esempio la responsabilità già nella fase di progettazione di una tecnologia. Si tratta di una prospettiva con numerose potenzialità in particolare se applicata alla progettazione tecnologica. Tuttavia, proprio a causa di queste potenzialità, deve essere maneggiata con cura.

Per chiarire meglio questi punti – specificamente la responsabilità attiva – è utile fare degli esempi che mettano chiaramente in luce come alcune scelte di progettazione possano evitare effetti negativi ma anche promuoverne di positivi. La progettazione di sistemi che incorporino la *privacy by design* è sicuramente uno di questi. Tuttavia, questi esempi non sono sufficienti per evidenziare alcune criticità di cui occorre tenere conto. È necessario dunque chiamare in causa la riflessione filosofica, e in particolare uno dei suoi strumenti più consueti: gli esperimenti mentali. Solo in questo modo è possibile fare emergere le sfumature, le complessità e gli effetti inattesi che possono emergere anche nel quadro della responsabilità attiva. Il contributo della filosofia in questo contesto è proprio quello di mostrare che non è sufficiente progettare sistemi di IA con l'intenzione di evitare conseguenze negative e promuovere effetti positivi, ma che è possibile farlo solo allargando lo sguardo, ampliando l'orizzonte e avendo sempre ben chiaro che i problemi etici, sociali, politici non possono essere risolti solo da soluzioni tecniche.

L'IA sta profondamente cambiando la nostra società: più di tutto sta ridefinendo i parametri di cosa significhi essere umano oggi. Per questa impresa abbiamo bisogno di filosofi in grado di valutare le questioni morali ed etiche che questa ridefinizione comporta. Abbiamo bisogno anche di

scienziati, di ingegneri, di informatici che siano in grado di apprezzare le discipline umanistiche – la filosofia nel caso in questione – perché saranno proprio questi ad essere meglio preparati ad evadere le trappole del potere tecnocratico e dei problemi politici scambiati per problemi tecnici. Saranno loro a doversi confrontare con le sfide del presente e del futuro: il cambiamento climatico, lo sviluppo dell'IA, le diseguaglianze crescenti. Ma proprio perché si tratta di sfide così significative, le buone intenzioni progettuali non bastano: per questo che abbiamo bisogno anche della riflessione critica della filosofia.