

Aiuto! L'IA discrimina

*Luisa Avitabile*¹

1. Grande è la meraviglia sotto il cielo digitale! l'IA discrimina! Che orrore! E da dove arrivano questi pregiudizi, queste opinioni escludenti, da dove giunge questo algoritmo che si permette di insinuarsi con un criterio di sproporzione?

Come se non lo sapessimo.

Arriva da noi che abbiamo abdicato alla nostra creatività, attraverso la convinzione che l'IA possa sostituirsi alla scelta dell'intelligenza umana, non più delega, ma supplenza. Nell'illusione che sia migliore di noi, senza renderci conto che è la nostra immagine.

L'IA ci supplisce, ci siamo detti, dimenticando che è un luogo impersonale. E poi, non è un iniziante, ma un esecutore, non possiede il concetto di spiritualità e non esercita la libertà attraverso l'esteriorizzazione di scelte intenzionali².

Ma questo non è sufficiente! Non basta a tranquillizzare i realisti progressisti che aspirano ad una società robotica utopisticamente migliore di quella umana! L'IA funziona e questo deve bastare per non dubitare. Di cosa dubitare? Del fatto che possa avere dei pregiudizi? Non sono certo preconcetti dell'IA, ma nostri, prontamente trasferiti per una sorta di osmosi miracolante.

Forse sfugge che le operazioni dell'IA non hanno una genesi nel loro funzionamento autoreferenziale, ma nella libertà creatrice dell'intelligenza, nella volontà dell'essere umano, anche nella capacità degli attuali mecenati digitali, sia occidentali che orientali, che detengono poteri economici e strumenti per allearsi in un *network*, particolarmente efficace sul versante della costituzione di un monopolio del consenso.

Dunque, se queste intelligenze umane sono escludenti, di riflesso lo sarà anche l'IA, non è una magia è la realtà!

¹ Luisa Avitabile, Preside della Facoltà di Giurisprudenza - Prof. ordinario di Filosofia del diritto, Università Sapienza di Roma. Intervento all'incontro di studio: *"Il futuro dell'intelligenza artificiale visto dalle donne: sfide, opportunità e confini"*, organizzato dal Dipartimento di Filosofia dell'Università Sapienza di Roma, in collaborazione con Noi Rete Donne, svoltosi il 13 maggio 2026 presso il Dipartimento di Filosofia, Villa Mirafiori, Roma

² Cfr. B. ROMANO, *Intelligenza artificiale e volontà*, Torino, 2024.

Questo fa sì che si possa dire che l'IA è lo specchio di quello che siamo o, meglio, di quello che è la massa.

Liberi, navighiamo profilati! Il gregge avanza e con esso i pregiudizi insiti nella società che, attraverso un flusso di dati, si trasferiscono nel virtuale.

Uguali, anzi equivalenti nel concederci alla rete senza alcuna coazione! Obbediamo all'Immagine Unica, sofisticata 'volontà di potenza' che trasforma la mistificazione in un imperativo categorico.

Realtà o *fake*, non importa, anche la giustizia si trasforma in un *nihil* dissacrante, una giustificazione che incide sulle relazioni e che è qualificata dalla consacrazione vincente di modelli giuridici strumentali, priva di attenzione ai contenuti qualitativi.

Possibile che siamo indotti a pensare secondo i dettami di una immagine unica, la macchina intelligente, che tutto dispone, in modo algoritmicamente incontrovertibile? La ragione impone di dubitare.

Allora, la risposta non può che essere negativa, perché le differenze conducono a quel mondo di originalità costituito dal conflitto, dalla libertà, dalla giuridicità, dove il 'tutto è possibile' viene limitato dal diritto, opera culturale dell'essere umano che, emancipandosi dalle leggi naturali, esprime se stesso attraverso la scelta.

2. Va ricordato che l'istituzione del diritto passa attraverso un'ideazione, ha la sua genesi nella relazione e il suo asse nella controversia. Nell'ambito della comunità, ogni essere umano forma la propria personalità non da solo, ma insieme agli altri, il che fa dire che l'alterità è essenziale alla costituzione della relazione giuridica; senza l'altro non esiste dialogo né controversia e neppure l'espressione di bene comune, dove il limite è costituito dal diritto nei confronti dell'assolutizzazione della libertà di un singolo o di gruppi che si spostano in modo egemonico.

Se tutto questo fosse fedelmente riportato nella prassi e nel virtuale, l'IA non sarebbe discriminante, ma così non è.

Allo stato, la tendenziale negazione della relazione come condizione essenziale del diritto dà luogo al narcisismo che ignora la presenza degli altri. Formazione di aggregati occasionali come i gruppi, le *chat*, le *communities* rappresentano il massimo della precarietà e della mancanza di approfondimento relazionale.

Celebrazione e culto dell'immagine, nell'era digitale, ricordano Narciso che, slegato dal reale e dalla possibilità di simbolizzare, si perde nell'immagine di sé riflessa nello specchio d'acqua e, non riuscendo a distanziarsene, rimane estasiato sino a morire.

La solitudine di Narciso è condivisa solo con la ninfa Eco che amplifica l'estasi della sua immagine. Affascinato, cade in una contemplazione letale. Significative le parole di Lavelle: «il delitto di Narciso è di preferire ... la sua immagine a se stesso»³; l'immagine di sé si sostituisce al dialogo e alla presenza dell'alterità.

Ci rammenta qualcosa? È il nostro riflesso: navighiamo persi ed estasiati dalla nostra immagine, perfetta, iconica, amplificata dai domini virtuali, accentuata sino all'enfasi mistica, conforme ai canoni supremi che esigono una dimensione staticamente efficace, energica e *social* in una perfetta combinatoria di *glamour* e adeguatezza digitale.

È bandito ogni slabbramento, a meno che non sia immediatamente filtrato per poi essere taggato, postato, scrollato e procedere così alla raccolta di *followers* e di *likes*. Un'immagine unica intorno alla quale vaghiamo senza sosta, perdendoci nei nostri profili!

Siamo ancora capaci di dubitare?

L'IA non fa altro che riflettere, come uno specchio di dimensioni strabilianti, i nostri dati. È chiaro che quando questi implicano pregiudizi, vengono restituiti, replicati e amplificati attraverso due tipi di discriminazioni algoritmiche: diretta e indiretta.

La prima ha un carattere di ingenuità, tipica delle relazioni violente, escludenti. Si ha nel momento in cui un algoritmo collegato a particolari tipi di attività, come ad esempio la selezione di candidati per una procedura valutativa, ne esclude alcuni in base al genere e/o all'appartenenza ad un gruppo o a un'etnia. In settori dominati prevalentemente dalla presenza maschile, si tende ad escludere le donne. È un tipo di discriminazione lineare, non complessa, facilmente prevedibile: può essere l'esclusione di una donna da una posizione lavorativa a causa dell'appartenenza di genere, o dall'accesso al danaro, come nel caso delle banche che non concedono prestiti in virtù della poca o minore affidabilità creditizia e imprenditoriale delle donne. I casi sono

³ L. LAVELLE, *L'errore di Narciso*, Milano, 1970, p. 11.

tanti, facilmente reperibili ed individuabili, a volte anche stupidi come solo la ‘banalità del male’ può esserlo.

Più complessa, meno chiara è la discriminazione indiretta.

Si tratta di decisioni che, se in apparenza non rivelano discriminazioni di genere, in realtà producono effetti negativi. È il caso dei risultati di quelle ricerche digitali sui principali *manager* di aziende particolarmente importanti che danno l'impressione di vivere in un mondo al maschile, dove le donne ai vertici sono poche, disseminando l'opinione o peggio la percezione di una *leadership* di genere.

Il riconoscimento vocale e facciale ha alcune criticità legate proprio alla mancanza di diversificazione: alcuni studi hanno mostrato che i dispositivi digitali, atti al riconoscimento, hanno una maggiore funzionalità per tipi maschi, bianchi, caucasici e non per donne e minoranze etniche. Questo in ambito sanitario rischia di compromettere diagnosi e terapie, si tratta di un tipo di *bias* che non solo produce disuguaglianza, ma può avere conseguenze dirette sulla salute e sulla sicurezza delle persone, annichilendo la differenza, che la medicina di genere ha assunto come valore per la cura delle donne.

Cosa dire dell'IA generativa utilizzata per creare materiale sessuale non consensuale? Colpisce soprattutto le donne, anche sotto forma di *revenge porn* e a nulla sembrano valere le introduzioni normative. Il Regolamento europeo sull'IA prevede alcune soluzioni come quella che, in caso di uso dell'IA in modo discriminatorio, la Commissione Europea possa aggiornare le regole tramite allegati modificabili, senza dover ricorrere alla procedura ordinaria, allo scopo di tutelare i diritti fondamentali. Si tratta di un lavoro impegnativo, peraltro difficile da realizzare che presenta non pochi momenti imbarazzanti se non si ha a disposizione personale qualificato.

Sul piano salariale, vi è addirittura una situazione che ha del paradossale: gli algoritmi amplificano le disparità storiche, consolidando l'opinione della posizione svantaggiata delle donne nel mercato del lavoro.

Non va meglio nei Paesi in via di sviluppo perché l'analfabetizzazione che colpisce soprattutto le donne, fa sì che non abbiano accesso ai benefici dell'innovazione tecnologica.

Per contrastare una situazione che rischia di riattivare antichi ed obsoleti convincimenti sulle donne, sono allo studio iniziative, soprattutto

nel Regno Unito, per identificare e ridurre i *bias* algoritmici in ambiti sensibili come il lavoro e i servizi pubblici.

3. Un ulteriore ostacolo è rappresentato dalla mancanza di riferimenti giurisprudenziali specifici per affrontare la discriminazione algoritmica. La Corte di giustizia dell'Unione Europea ha certamente affrontato diversi casi di discriminazione sul lavoro, ma questo non significa che ci sia una giurisprudenza consolidata che tratti esplicitamente la discriminazione basata sugli algoritmi. È un vuoto normativo che rende difficile stabilire linee precise, atte a risolvere il problema e le criticità legate alla disparità di genere, lasciando alla politica un compito improbo. Mancano le proposte di attivare un'azione pedagogica mirata, anche in previsione di futuri interventi legislativi specifici.

Allo stato, vi sono numerosi progetti dove l'uso dell'IA è finalizzato ad affrontare le disuguaglianze di genere.

Ridurre i lavori domestici, fornire informazioni sulla salute riproduttiva delle donne, sensibilizzare dal punto di vista culturale, rilevando la presenza di stereotipi di genere sul *web*, sono azioni rivolte ad usare l'IA come strumento contro le discriminazioni di genere, ma non è sufficiente! Anzi è controproducente: se ridurre il carico dei lavori domestici viene visto come sollevare le donne dai compiti familiari, significa che non si è ancora sulla via della parità di genere; alimentare algoritmi che guardino ai compiti domestici come a lavori destinati al mondo femminile, costituisce di per sé una negatività. Guardare solo alle possibilità riproduttive della donna costituisce un dinosaurico ritorno al passato.

Nondimeno, alcune aziende stanno utilizzando piattaforme digitali per controllare eventuali fenomeni di discriminazione e molestie legate al genere, il che aiuta la prevenzione e la tutela.

In sintesi: l'IA eredita i nostri pregiudizi e conferma *bias* di genere presenti nelle nostre società.

Non sono gli algoritmi a sbagliare, perché si sa che l'IA non pensa non dubita non sbaglia, è nell'evidenza scientifica che non è portatrice di *logos* e di dialogo. L'IA non è intrinsecamente buona o cattiva: è uno specchio che riflette la società che la crea. Se vogliamo che non amplifichi le disparità di genere, allora è necessario programmare algoritmi che non solo evitino di perpetuare le disuguaglianze, ma che le contrastino.

Questo richiede tecnologie migliori e un impegno sociale condiviso in una sorta di responsabilità collettiva.

Il futuro dell'IA è determinato dalle decisioni legislative, politiche, economiche ecc. che determineranno non solo l'efficacia della tecnologia, ma anche, e soprattutto, l'orientamento sociale.

Si può senz'altro affermare che l'IA è un significativo strumento di progresso, ma è necessario un controllo, oltre che un'alfabetizzazione digitale, e una regolamentazione rigorosa per evitare i *bias* e le interferenze sulle conquiste di parità di genere che l'umanità, con tanto impegno, ha tentato di affermare.