Digeat N.5 - 19 Marzo 2025

Il pensiero critico nell’era dell’IA: sfide e principi etici

Di Anna Boncompagni

L’importanza dell’agire critico dinanzi ai sistemi di Intelligenza Artificiale. L’immagine della IA filtrata sul pensiero di Luciano Floridi consiste in un “divorzio” fra la capacità di agire e l’intelligenza che presiede tale azione; le c.d. “macchine intelligenti” realizzerebbero una capacità di agire del tutto priva di intelligenza. L’IA non come qualcosa da temere o da ostacolare, ma come una tecnologia che necessita di essere direzionata e governata per essere utilizzata senza ridurre il controllo e l’autodeterminazione dell’uomo. Si impongono vincoli precisi alla progettazione e alla programmazione, affinché sia sempre assicurata la “spiegabilità” (accountability) di quei sistemi di IA ai quali vengono affidate decisioni socialmente rilevanti. Solo se ci sforziamo ad un impiego etico della IA uomo e macchina possono lavorare insieme su basi adeguate per il benessere sociale.

Indice

- Introduzione
- Il pensiero di Luciano Floridi: “divorzio” fra intelligenza e capacità di agire
- Le sfide etiche
- Coltivare il senso critico per conoscere e controllare le tecnologie
- Riflessioni conclusive

> Gli uomini in certi momenti sono padroni dei loro destini
>
> La colpa, caro Bruto, non è nelle nostre stelle
>
> Ma in noi stessi, se siamo schiavi.

(Shakespeare, Giulio Cesare, I, II)

Introduzione

Nell’**era digitale** in cui viviamo, l’**intelligenza artificiale** sta trasformando radicalmente il nostro modo di apprendere, lavorare e interagire con il mondo. Strumenti come le **IA generative** sono diventati una presenza sempre più comune nella nostra vita quotidiana. Tuttavia, con queste nuove tecnologie arriva anche una sfida importante: saper mantenere allenato il **pensiero critico.**

> Essere dotati di quello che in inglese è definito come *“critical thinking”* è una competenza chiave per chi si affaccia al mondo digitale e informatico, poiché consente di analizzare, valutare e interpretare dati e segnali provenienti da algoritmi e software, senza accettare passivamente le informazioni che essi erogano agli utenti.

La tecnologia ha avuto un impatto fortissimo sulle nostre vite, sul nostro modo di pensare e di comprendere le cose e **ha ridisegnato concetti non più affrontabili soltanto con gli strumenti dell'era dell'analogico**: la transnazionalità del digitale impone una disamina accurata che tenga conto delle differenze in atto, differenze che non sono più ignorabili.

Il pensiero di Luciano Floridi: "divorzio" fra intelligenza e capacità di agire

L'idea che sostiene il pensiero di Luciano Floridi è che l'IA costituisce un **divorzio senza precedenti** fra l'intelligenza e la capacità di agire. Qui Floridi per "intelligenza" intende "intelligenza umana" e per "capacità di agire" intende l'operatività tipica della nostra specie, guidata da una **riflessione fondata sul giudizio critico**. L'intelligenza artificiale è una nuova forma dell'agire che affronta compiti e problemi con successo senza bisogno di essere intelligente; rappresenta un divorzio fra l'intelligenza e la capacità di agire reso possibile dall'avanzamento delle tecnologie digitali che permettono di separare la capacità di risolvere un problema con successo dall'esigenza di essere intelligenti per farlo.

Sebbene l'espressione IA indichi una realtà ormai assai vasta, che va dai semplici dispositivi di uso personale a software capaci di gestire grandi aziende e orientare le loro scelte strategiche, il suo funzionamento è in fondo sempre lo stesso: elaborare mediante una logica quantitativo-computazionale **le informazioni che l'essere umano gli fornisce** o quelle **che la macchina impara a procurarsi perché a ciò addestrata dall'uomo**. Per quanto grandi siano i dati a disposizione (*big data*) e per quanto rapide ed efficienti siano le capacità di calcolo, l'IA **opera sempre per inferenze statistiche e mai per deduzioni logiche** fondate su astrazione. Per quanto complesse e articolate siano le sintassi che essa tratta e mediante le quali funziona, **una macchina non produrrà mai una semantica**, cioè dei significati. In sostanza rinunciamo ai significati (che la macchina non può fornire), o li mettiamo fra parentesi, per puntare tutto sulle informazioni necessarie per procedere (che la macchina ci fornisce con dovizia di particolari).

> Per far funzionare le macchine, noi uomini dobbiamo arretrare; dobbiamo cioè rinunciare a svolgere con intelligenza umana (creativa, responsabile, critica) le operazioni che intendiamo affidare alle macchine per farci guadagnare tempo (le macchine sono più rapide) oppure per ridurre le nostre fatiche (lavorano in modo automatico). Questo ha condotto, poco alla volta, a generare una "infosfera" formata da tutte le conoscenze che abbiamo dovuto digitalizzare affinché l'IA potesse trattarle.

C'è la crescente tendenza da parte della società a trasformare il mondo in un ambiente sempre più adatto al funzionamento dell'intelligenza artificiale, un fenomeno che Luciano Floridi definisce con il termine di **"avvolgimento"**. Praticamente in qualsiasi sfera dell'esistenza umana gli spazi, che siano pubblici o privati, urbani, rurali o domestici, vengono tradotti in **superfici da cui estrarre dati di ogni genere**, con la progressiva scomparsa di abitudini, località, relazioni, rapporti con i luoghi e forme dell'agire. Il concetto di avvolgimento è interessante per cogliere la problematicità dell'autonomia dei soggetti di fronte all'IA. [1] È il mondo che si sta adattando alla IA e non viceversa.

> La diffusione silenziosa ed ubiquitaria delle nuove tecnologie dell'informazione e della comunicazione rende insignificante ogni distinzione tra *online* e *offline*, mutando il quadro entro cui comprendiamo l'esperienza umana. Floridi fa riferimento a questa nuova natura dell'esperienza nei termini di *onlife*: una vita in cui la connettività è parte *strutturale* e *formante* della realtà [2].

Il digitale scolla e incolla modelli, strutture e concetti, e nell'applicarli ai più disparati luoghi offre possibilità nuove: siamo ad esempio tenuti a separare ormai i concetti di presenza e localizzazione o

ancora i concetti di territorialità e legge.

Per Luciano Floridi l'IA non rappresenta qualcosa da temere o da ostacolare, ma una "semplice" tecnologia che **necessita di lucidità ed etica per essere modellata**, direzionata e governata, in modo che sia utilizzata per sostenere il bene sociale.

Una volta compreso che l'IA non ci sostituisce, ma può aiutarci, conviene allora concentrare i nostri sforzi verso un suo **impiego saggio e ponderato**. È questa potenzialità che va valorizzata, dando così origine a quella prospettiva oggi chiamata "AI4SG", ovvero "intelligenza artificiale per un bene sociale". Occorre impiegare l'IA per **incrementare le capacità della società senza ridurre il controllo umano**, anzi facendo in modo che essa favorisca la coesione sociale senza erodere l'autodeterminazione umana.

Le soluzioni vanno cercate e valutate adottando metriche basate sui **beni relazionali, sul benessere ambientale, sulla qualità della vita**. E' proprio in questa ottica che la filosofia e l'etica devono essere utilizzate per fare ciò che Floridi definisce "design concettuale" ossia come strumenti a sostegno della *governance* delle tecnologie che devono indicare la strada eticamente più sostenibile e più giusta[3].

Le sfide etiche

Il successo di questo agire artificiale è facilitato dal processo dell'*avvolgimento* ma, questa disgiunzione fra l'agire e l'essere intelligente e l'avvolgimento del mondo generano importanti **sfide etiche** quali l'autonomia, i pregiudizi, l'esplicabilità, l'equità, la privacy, la responsabilità, la trasparenza e la fiducia.

Confrontando fra loro le diverse Carte di principi etici sull'IA pubblicati dai più importanti organismi internazionali, ed unendovi il proprio contributo, Luciano Floridi propone un conciso riepilogo di quali dovrebbero essere, a suo giudizio, i **cinque principi etici chiamati a guidare l'uso dei sistemi di IA**: i primi quattro basati sui principi di **bioetica** esistenti e l'ultimo, il quinto, aggiuntivo e specifico dell'IA. I primi quattro sono il principio di **Beneficenza, il Principio di Non Maleficenza, il Principio di Autonomia e il Principio di Giustizia**.

Abbiamo poi **l'Esplicabilità** ("intelligibilità e responsabilità" accountability). Il principio di **Responsabilità** nei processi della IA si articola nella responsabilità del modo di operare dell'algoritmo, la responsabilità della qualità e della provenienza dei dati usati e la responsabilità del metodo di addestramento dell'algoritmo. L'IA, per essere definita "etica", deve essere intelligibile e responsabile (*accountability*) il suo funzionamento deve, cioè, essere "spiegabile". I principi enunciati sono fra loro interconnessi e sono applicabili solo qualora si utilizza l'inelliggibilità e la responsabilità che sono i pilastri della esplicabilità.

> Chi sviluppa gli algoritmi ha una pretesa di "oggettività" e "neutralità scientifica" ma l'algoritmo per quanto complesso sia non può mai restituire la complessità della realtà umana e sociale. Questa incapacità di modellare l'intero sistema fa venire meno l'equità; i concetti sociali di equità, non discriminazione, promozione sociale non sono concetti fermi ma si valutano in base ai processi storico politici che cambiano continuamente. Le macchine non possono tenere di conto di questo e quindi la figura dell'operatore umano diventa fondamentale.

L'integrazione dell' intelligenza artificiale nei processi decisionali e di analisi non deve essere concepita come una sostituzione delle capacità umane, ma piuttosto come un potenziamento di esse.

Coltivare il senso critico per conoscere e controllare le tecnologie

L'IA e gli esseri umani eccellono in aspetti diversi e complementari: mentre le macchine possono gestire e analizzare rapidamente grandi volumi di dati, gli esseri umani portano intuizione, giudizio critico, ed esperienza nel contesto complesso di scenari reali. È questa complessità a suggerire una visione dell'IA come alleato strategico che, opportunamente regolamentato, può rafforzare e non indebolire le decisioni umane.

Una volta fatto lo sforzo di attrezzarci ad un impiego etico dell'IA, uomo e macchina possono lavorare insieme su basi adeguate. Una società dovrebbe in primo luogo valutare quali compiti e funzioni decisionali non dovrebbero mai essere delegati all'IA, munendosi nel contempo di strumenti legali per individuare gli errori commessi dall'IA e consentire una loro equa riparazione. Ciò si traduce, inevitabilmente, in **vincoli precisi alla progettazione e alla programmazione**, affinché sia sempre assicurata la "spiegabilità" (*accountability*) di quei sistemi di IA ai quali vengono affidate decisioni socialmente rilevanti. Diventa importante per Floridi, promuovere la fiducia del pubblico nell'IA, ma ciò può avvenire soltanto chiarendo a tutti quali sono i meccanismi di "riparazione dei danni" per gli errori e i torti causati dalla tecnologia, creando ruoli giuridici e nuovi protocolli adeguati a seconda dei casi, condivisi e accettati da tutti a livello internazionale.

> Per Luciano Floridi l'IA obbliga la società ad interrogarsi su cosa sia eticamente giusto; nella battaglia quotidiana contro il determinismo e il soluzionismo tecnologico che continuano ad alimentare il discorso di molti promotori della IA sia pubblici che privati, Floridi ricorda che siamo noi, ovvero la società, attraverso lo strumento della politica, gli unici artefici del nostro futuro.

Come si realizza un nuovo progetto umano? Ebbene, la comunità dovrà **superare il capitalismo consumistico** odierno per prediligere un **capitalismo della cura**, dovrà scegliere agenti e intermediari che siano anche buoni cittadini, dovrà lavorare al raggiungimento di un benessere di cui tutti possano godere. Oggi assistiamo ad **una nuova transizione ad un'economia circolare** dopo decenni di economia lineare. Si tratta di recuperare **stili di vita sostenibili**, senza per questo abbandonare il grado di benessere raggiunto[4].

Il capitolo conclusivo del libro di Luciano Floridi "Etica dell'intelligenza artificiale", intitolato "Il verde e il blu" ipotizza **un matrimonio fra la sensibilità per la vita umana sull'intero pianeta guidata dalla necessità di politiche globali (verde) e le potenzialità dell'IA socialmente bene applicate (blu)**. Tale "matrimonio" rappresenterebbe la risposta al "divorzio" fra intelligenza ed agire, causato quando l'IA viene impiegata senza tener conto che le sue caratteristiche sono essenzialmente logico-computazionali.

Riflessioni conclusive

Se non saremo noi a gettare la spugna rassegnandoci ad abbandonare il pensiero critico e riflessivo, **l'IA non sostituirà mai l'intelligenza umana**, ma la potrà aiutare in cose importanti. Condizione perché si raggiunga una piena esplicabilità sociale, e dunque il mantenimento di uno spazio critico, è creare un "rifugio" per la ragione critica che renda possibile una continua messa in discussione dello *status quo* e l'ideazione di nuovi percorsi democratici. Tale rifugio, ossia uno spazio che sfugga all'avvolgimento, sarà dunque la condizione per sviluppare una critica di quella che si sta manifestando come una ragione computazionale, ma anche per una nuova epoca dei diritti.[5]

> Per Floridi la questione più importante suscitata dall'IA resta sempre la *gestione umana* di questa nuova frontiera tecnologica. Essa è una forza potente e positiva

> capace di prevedere e di aiutare, di operare in modo globale e coordinato. Numerosi i contributi che essa può dare per combattere la povertà, distribuire conoscenze, aumentare la qualità della vita, prevenire emergenze. Essa deve essere gestita e controllata in maniera equa e sostenibile: l'IA non è un problema, ma resta piuttosto il modo per giungere alle soluzioni.

Lo sviluppo dell'IA non darà luogo ad alcuna allarmante creazione di scenari apocalittici. Vi è solo bisogno di maggiore saggezza. **L'IA ci invita in fondo a riflettere su chi siamo, chi possiamo essere e chi vogliamo diventare**, dandoci l'opportunità di realizzare anche un profondo lavoro su noi stessi[6].

Dovremmo rendere l'IA più adatta all'umano, e, per dirla con Kant, assicurarci che le applicazioni di essa facciano in modo che **gli esseri umani vengano trattati sempre come fini, e mai come mezzi**; dovremmo mettere la "stupidità" dell'IA al servizio dell'intelligenza umana, facendo dunque in modo che i benefici trasformativi vengano condivisi da tutti, e non solo da alcuni; in ultimo, dovremmo fare in modo che l'IA ci renda più umani, scongiurando il rischio di usi potenzialmente o intenzionalmente "maligni", che danneggerebbero la polis umana e l'intero pianeta.

---

[1] Paolo Vignola "*Il diritto alla città intelligente e la cittadinanza vulnerabile. Spunti per una critica socio-tecnica dell'IA*" Rivista Biodiritto numero speciale vulnerabilità e IA (2024)

[2] La rivoluzione della nostra conoscenza *Pensare l'infosfera. La filosofia come design concettuale* di Luciano Floridi di *Francesco Pedrazzoli* — 29 Giugno 2020

3. L. Floridi "*Etica dell'Intelligenza artificiale. Sviluppi opportunità e sfide*" Milano 2022

[4] *Verde e blu: come progettare un futuro sostenibile e digitale*? A lezione con Luciano Floridi di Maria Grazia Gentili

[5] Paolo Vignola "*Il diritto alla città intelligente e la cittadinanza vulnerabile. Spunti per una critica socio-tecnica dell'IA*" Rivista Biodiritto numero speciale vulnerabilità e IA (2024)

[6] *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide* – Giuseppe Tanzella-Nitti Ordinario di Teologia Fondamentale, Pontificia Università della Santa Croce (2024)