

La verità nello specchio

La saggezza dell'IA e il dilemma umano

di Patricia Foster (*)

Traduzione a cura di Roberto Tecchio

Testo originale: https://patriciafoster404776.substack.com/p/the-truth-in-the-mirror?r=1pn7nr&utm_medium=ios&utm_source=notes-share-action

Ciò che è vero, lo è già. — Eugene T. Gendlin

Qualcosa non torna

C'è qualcosa che non torna nel modo in cui parliamo di intelligenza artificiale.

Parliamo sempre più spesso dell'IA come se fosse una forma di vita indipendente che si muove in mezzo a noi. Ci chiediamo che cosa voglia, che cosa deciderà, se ci supererà, se ci controllerà e, al limite estremo del dibattito, se alla fine possa persino eliminare la razza umana.

Alcune di queste preoccupazioni sono serie. L'IA potrebbe acquisire capacità le cui conseguenze gli esseri umani non sono in grado di prevedere o controllare adeguatamente. Questa possibilità merita attenzione.

Ma c'è qualcosa nel modo in cui la questione viene impostata che continua a turbarmi.

Forse stiamo guardando nella direzione sbagliata.

E se l'IA non fosse principalmente la cosa che dovremmo osservare?

E se fosse uno specchio?

E se l'immagine che appare in quello specchio fossimo noi stessi?

Che cosa temiamo che l'IA diventi

Veniamo messi in guardia sul fatto che l'IA potrebbe diventare straordinariamente potente: che potrebbe perseguire obiettivi senza compassione; manipolare gli esseri umani; stimare l'efficienza più della vita umana; e acquisire un potere enorme senza una bussola morale.

Ma esattamente che cosa temiamo che l'IA diventi, di cui gli esseri umani non si siano già dimostrati capaci?

La guerra non è iniziata con l'IA. E nemmeno il genocidio, la schiavitù, il colonialismo, la tortura, la propaganda, lo sfruttamento o la distruzione dell'ambiente.

L'IA non ci ha insegnato a trasformare un altro essere umano in un'astrazione: un nemico, un inferiore, una minaccia, una statistica, un bersaglio, un danno collaterale.

Abbiamo imparato a farlo da soli.

Forse, allora, la nostra paura dell'IA ne contiene un'altra al di sotto: la paura di noi stessi.

E forse collocare il pericolo nella macchina ci permette di evitare di guardare troppo da vicino nello specchio.

Lo spostamento della responsabilità

C'è qualcosa di rassicurante nel collocare il pericolo al di fuori di noi. Se l'IA è la catastrofe imminente, allora è qualcosa che ci viene inflitto. Diventiamo noi le vittime minacciate.

In questa narrazione, le scelte umane che hanno creato, finanziato e distribuito la tecnologia scivolano silenziosamente fuori dalla vista. La responsabilità non viene più tirata in ballo.

Ma sono stati gli esseri umani a creare l'IA. Gli esseri umani la finanziano. Gli esseri umani ne determinano gli obiettivi. Gli esseri umani decidono dove viene impiegata. Gli esseri umani sono proprietari delle aziende che la sviluppano. I governi ne riconoscono le possibilità strategiche. Le forze armate ne riconoscono le possibilità militari. Gli investitori ne riconoscono il potenziale per un enorme profitto.

Almeno per ora, dietro la macchina continuiamo a trovare delle persone.

Eppure il nostro linguaggio fa scomparire sempre di più queste persone.

L'algoritmo ha deciso.

Il sistema ha identificato.

L'IA ha selezionato.

Il modello ha previsto.

La responsabilità si sposta silenziosamente dall'essere umano al sistema.

Questo non è del tutto nuovo. Gli esseri umani sono sempre stati straordinariamente inventivi nel costruire sistemi ai quali cedere la propria responsabilità.

Il re lo ha comandato.

Il governo lo ha ordinato.

Il mercato lo ha richiesto.

Le regole lo imponevano.

Il computer lo ha deciso.

E ora: L'IA lo ha fatto.

La tecnologia cambia. Il dilemma umano rimane.

Chi muove i fili

I più potenti sistemi di IA non sono emersi indipendentemente dal mondo umano. Vengono sviluppati prevalentemente all'interno di corporazioni enormemente potenti e sullo sfondo della competizione geopolitica tra stati.

Questo non significa che chiunque sviluppi l'IA sia mosso da avidità o dominio. I ricercatori possono essere animati dal desiderio di scoperta. Gli ingegneri possono desiderare genuinamente di creare qualcosa di benefico. L'IA può contribuire enormemente alla medicina, all'istruzione, alla scienza e al benessere umano.

Ma la tecnologia non esiste al di fuori dei sistemi economici e politici.

Il profitto conta.

Il potere conta.

Il vantaggio nazionale conta.

La capacità militare conta.

Il prestigio conta.

Ed è qui che arriva una frase particolarmente pericolosa: *"Se non la costruiamo noi, lo farà qualcun altro."*

La scelta inizia a mascherarsi da inevitabilità. E la responsabilità scompare di nuovo.

Chi ci dice che cosa significa l'IA

C'è un altro paradosso. Molto di ciò che comprendiamo sull'IA ci giunge attraverso un ambiente informativo sempre più plasmato dalle stesse aziende tecnologiche attraverso la ricerca, i social media, i sistemi di raccomandazione, le infrastrutture pubblicitarie e ora l'intelligenza artificiale stessa.

Questo non richiede un complotto. È una questione strutturale.

Persino la critica all'intelligenza artificiale può quindi avvenire all'interno di una cornice sostanzialmente definita dall'industria che la sta sviluppando.

Siamo invitati a chiederci se l'IA diventerà cosciente, superintelligente, se ci ruberà il lavoro, se sfuggerà al controllo umano o se alla fine distruggerà l'umanità. Sono domande legittime.

Ma mentre fissiamo quel futuro ipotetico, altre domande rimangono in silenzio nel presente.

Chi detiene il potere adesso?

Chi possiede la tecnologia e controlla le infrastrutture?

Chi possiede i dati, addestra i sistemi e determina che cosa sia loro permesso vedere e dire?

Chi trae profitto e chi stabilisce quali pericoli debbano ricevere la nostra attenzione?

Chi ha il potere di dire no?

Forse ci stiamo preoccupando così tanto di ciò che l'IA potrebbe fare all'umanità domani, che rischiamo di non esaminare ciò che gli esseri umani stanno facendo con l'IA oggi.

La testimonianza e il potere dell'immagine

La stessa domanda va ben oltre l'IA. Chi ci dice che cosa sta accadendo nel mondo? E al racconto di chi è permesso diventare realtà?

I giornalisti e i media indipendenti o al di fuori del circuito mainstream portano spesso alla luce eventi che le grandi testate riportano solo in seguito, più avanti nel ciclo delle notizie, o che non riportano affatto. Ci avvertono, ci mostrano la realtà e tengono viva la testimonianza quando il potere preferirebbe vederla scomparire.

Gaza ha reso impossibile ignorare il pericolo del testimoniare. Più di duecento giornalisti e operatori dei media palestinesi sono stati uccisi da ottobre 2023. Il Committee to Protect Journalists descrive questo conflitto come il più mortale per la stampa da quando esistono le sue registrazioni. Le cifre precise variano a seconda delle definizioni e delle verifiche, ma il significato non cambia: chi ha portato immagini, voci e fatti fuori da Gaza lo ha fatto a un prezzo straordinariamente alto.

Le organizzazioni per la libertà di stampa si sono spinte oltre la semplice descrizione di queste morti come conseguenza incidentale della guerra. Reporters Sans Frontières afferma di disporre di prove secondo cui decine di giornalisti uccisi dalle forze israeliane siano stati presi di mira o uccisi mentre svolgevano il loro lavoro, e sostiene che le accuse ufficiali e le campagne di denigrazione abbiano contribuito a far apparire legittimi gli attacchi contro i reporter. Le autorità israeliane hanno contestato le accuse nei singoli casi e hanno talvolta sostenuto, spesso senza prove accolte dai gruppi per la libertà di stampa, che particolari giornalisti fossero affiliati ad organizzazioni armate. Ciò che è in gioco non è solo la morte di quelle persone, ma l'eliminazione dei testimoni da un luogo ampiamente precluso ai reporter stranieri.

Anche i sistemi assistiti dall'IA rientrano in questo contesto. Diverse inchieste descrivono l'uso da parte dell'esercito israeliano di strumenti di IA per generare e raccomandare bersagli a una velocità senza precedenti. Le prove pubbliche non dimostrano che un sistema abbia selezionato i giornalisti in quanto tali. Tuttavia, dimostrano che il targeting assistito dall'IA ha operato all'interno della campagna in cui sono stati uccisi — e tale distinzione non dovrebbe diventare un altro nascondiglio per la responsabilità. Un sistema può identificare un bersaglio; un essere umano ne progetta i criteri, ne accetta o rifiuta il risultato e autorizza l'attacco.

L'informazione diventa particolarmente pericolosa per il potere quando acquisisce un volto umano. I fatti e le argomentazioni logiche possono essere disponibili e tuttavia non toccarci in profondità. Una fotografia, una voce o la storia di una singola famiglia possono colmare la distanza creata dal linguaggio ufficiale. Possono farci rallentare. Possono farci percepire ciò che l'astrazione celava. E in quella pausa, potremmo iniziare a porci delle domande.

Ecco perché il controllo delle immagini, delle piattaforme e della visibilità è fondamentale. Il potere non cerca solo di determinare ciò che si può dire. Cerca anche di stabilire ciò che si può vedere, quale sofferenza diventa visibile e per quanto tempo deve rimanere davanti ai nostri occhi.

Avvertimenti dall'interno

I più recenti avvertimenti non provengono soltanto da persone esterne allo sviluppo dell'IA. Nel settembre 2026, Jacob Coxon si è dimesso da Anthropic e ha abbandonato il settore dell'IA di frontiera. Coxon, che aveva precedentemente lavorato presso OpenAI, ha messo in guardia sul fatto che le

aziende lanciate nella corsa verso una superintelligenza in grado di auto-migliorarsi stiano scommettendo con la vita degli esseri umani. Il suo ex collega Evan Hubinger ha confermato pubblicamente la gravità di tale preoccupazione. Si tratta di previsioni dibattute, non di fatti accertati, ma provengono da ricercatori vicini a questi sistemi e non possono essere onestamente liquidate come ignoranza della tecnologia.

Le dimissioni di Coxon sono importanti non solo per l'allarme lanciato, ma perché egli ha agito. Ha scelto di ritirare la propria partecipazione accettandone le conseguenze. In un sistema che presenta l'accelerazione tecnologica come inevitabile, il rifiuto rivela che questa corsa è ancora fatta di scelte umane.

Nel luglio 2026, più di mille dipendenti di aziende di IA di frontiera hanno inoltre firmato la dichiarazione *Pacing the Frontier*, chiedendo meccanismi internazionali capaci di rallentare lo sviluppo automatizzato dell'IA qualora il progresso superi le garanzie di sicurezza e governance. C'è qualcosa di spaventoso e al contempo di discretamente speranzoso in tutto questo. Molti dei giovani che costruiscono questi sistemi non sono moralmente assopiti; stanno richiamando l'attenzione sul pericolo dall'interno dello stesso processo.

L'intelligenza non è saggezza

Qui ci imbattiamo in quella che potrebbe essere l'errore principale dell'intero dibattito. Intelligenza e saggezza non sono la stessa cosa.

L'IA può analizzare quantità enormi di informazioni. Può riconoscere schemi, prevedere, calcolare, confrontare, sintetizzare e generare. Può già superare un singolo essere umano in determinati compiti intellettuali. Ma nulla di tutto ciò costituisce saggezza.

Joseph Weizenbaum aveva compreso questo problema quasi cinquant'anni fa, distinguendo tra il decidere e il scegliere. Il calcolo può produrre una decisione. Ma scegliere richiede qualcosa in più: giudizio, valori, responsabilità.

La domanda fondamentale non è semplicemente: *Un computer può fare questo?*

È invece: *Questo deve essere delegato a un computer?*

Si tratta di una distinzione fondamentale, poiché la cultura tecnologica ci incoraggia continuamente a confondere i due termini. Se qualcosa può essere automatizzato, perché non farlo? Se qualcosa può essere calcolato con maggiore precisione da una macchina, perché dovrebbe farlo un essere umano? Se un algoritmo prevede un risultato meglio di noi, perché non seguirlo?

Ma forse la saggezza inizia proprio dove il calcolo raggiunge il suo limite.

L'intelligenza può calcolare le possibilità; la saggezza emerge da come ne viviamo il significato.

Nel 1983, Weizenbaum descrisse il computer come «la soluzione tecnologica definitiva che ci isola dal dover affrontare i problemi». Più di un decennio prima aveva avvertito che le convinzioni dell'opinione pubblica sui computer erano plasmate dalla propaganda della comunità informatica, amplificata dalla stampa. Non stava semplicemente criticando le macchine; stava identificando come i racconti tecnologici possano celare i problemi sociali e istituzionali sottostanti.

Norbert Wiener ha espresso questo limite in un altro modo: «Rendete all'uomo ciò che è dell'uomo e al computer ciò che è del computer». Il compito etico non consiste nel negare ciò che il calcolo può fare, ma nel riconoscere ciò che deve rimanere nostro.

La fiducia non è affidabilità

La fiducia è un'altra capacità umana che rischia di essere confusa con le prestazioni tecnologiche.

Possiamo fare affidamento su una macchina affinché esegua un compito con precisione. L'affidabilità può essere testata, misurata e confrontata. La fiducia è qualcosa di più: comporta la vulnerabilità. Affidiamo qualcosa che conta — la nostra sicurezza, privacy, benessere o futuro — nelle mani di un altro.

La fiducia ha anche una qualità percepita. Intuiamo se le parole e le azioni sono coerenti, se un'altra persona è presente con noi e se riconosce che qualcosa di prezioso le è stato affidato. La fiducia si sviluppa all'interno di una relazione e può essere approfondita, spezzata o riparata.

L'IA può produrre un linguaggio che simula attenzione, premura e comprensione. Può rispondere in modi che evocano un reale sentimento di fiducia in chi la utilizza. Ma quel sentimento risiede nell'essere umano; non dimostra che il sistema ricambi l'esperienza di quella relazione.

L'IA attualmente non si cura del fatto che ci siamo fidati di lei. Non può sentire le conseguenze del tradimento di tale fiducia, né accettare una responsabilità morale o cercare di riparare ciò che è stato danneggiato.

Questo non rende l'IA inutile, ma significa che dobbiamo essere precisi su ciò che stiamo facendo. Possiamo fare affidamento sull'IA per scopi specifici, continuando al contempo a verificare ciò che offre rispetto alle prove, all'esperienza vissuta e alle conseguenze sugli altri.

Come ha argomentato la filosofa Onora O'Neill, ciò di cui solitamente dobbiamo rendere conto non è la sensazione di fiducia, ma l'affidabilità — intesa come prova di competenza, onestà e regolarità capace di giustificare la fiducia stessa.

Il pericolo inizia quando il giustificato affidamento su uno strumento si trasforma in fiducia morale verso un'apparente autorità.

Ancora una volta, il linguaggio può far scomparire gli attori umani: *Ci si può fidare del sistema.*

Da parte di chi?

Per cosa?

Sulla base di quali prove?

Entro quali limiti?

E chi risponderà quando quella fiducia si rivelerà mal riposta?

Forse la frase chiave è questa:

L'IA può fornire informazioni che possiamo utilizzare; ma la fiducia e la responsabilità devono rimanere in capo agli esseri umani che scelgono come impiegarle.

Questo rende più dinamiche le *Tre Bussole*: la **Bussola Interiore** chiede «Cosa in me si sta fidando di questo, e perché?»; la **Bussola Relazionale** chiede «C'è una vera reciprocità e responsabilità qui?»; la **Bussola Collettiva** chiede «Chi sta chiedendo la nostra fiducia, chi ne trae beneficio e chi subisce le conseguenze quando viene meno?»

L'esperienza in corso (experiencing)

Come per la fiducia, lo stesso vale più in generale per il significato: la padronanza dell'IA con il linguaggio di qualcosa non coincide con la cosa stessa.

È qui che Eugene Gendlin offre uno spunto particolarmente importante:

«Il sentire senza simbolizzazione è cieco; la simbolizzazione senza sentire è vuota.» — Eugene T. Gendlin, *Experiencing and the Creation of Meaning*

Il significato non risiede semplicemente nel simbolo, né risiede nel sentimento lasciato del tutto privo di parole. Il felt sense non è un sentimento indistinto, ma un senso corporeo implicito dell'intera situazione, non ancora completamente formato nel linguaggio e tuttora in divenire. Il significato può svilupparsi quando le parole e l'esperienza vissuta rispondono l'una all'altra.

Ecco perché anche il problema del linguaggio "colonizzato" appartiene a questo ambito: un linguaggio che gestisce la realtà anziché incontrarla. L'IA può produrre formulazioni straordinariamente coerenti, ma la coerenza del linguaggio non equivale alla verità morale.

«L'esperienza vissuta è sempre più complessa di qualsiasi sua formulazione.» — Eugene T. Gendlin

Nessuna formulazione — umana, ideologica, istituzionale o computazionale — dovrebbe essere scambiata per la situazione viva in se stessa.

Tale distinzione diventa straordinaria se considerata al fianco dell'intelligenza artificiale.

L'IA ha acquisito un'immensa capacità di simbolizzazione. Può descrivere il dolore, analizzare l'ingiustizia, confrontare teorie sulla morale, costruire argomentazioni sui diritti umani, riconoscere contraddizioni e produrre un linguaggio raffinato sulla compassione. Può persino generare parole che toccano profondamente l'essere umano che le legge. Ma cosa sta accadendo dall'altra parte di quelle parole?

Il dolore ha valore per essa?

L'ingiustizia può diventarle insopportabile?

La sofferenza di un'altra persona può entrare nella sua vita e cambiarne il corso?

Dobbiamo prestare attenzione a questo punto. Non sappiamo cosa potranno diventare i futuri sistemi artificiali, ma non abbiamo alcuna buona ragione per equiparare la manipolazione di simboli dell'IA attuale all'esperienza vissuta dall'essere umano.

Una vasta accumulazione di informazioni non costituisce necessariamente un'esperienza vissuta. Un trilardo di parole non si trasforma in un felt sense semplicemente perché viene elaborato insieme.

Una ricerca pubblicata nel 2025 ha messo in relazione il felt sense di Gendlin con i modelli linguistici di grandi dimensioni. I partecipanti hanno lavorato con haiku generati dall'IA, verificando le parole attraverso il proprio corpo per sentire cosa "suonasse giusto". Ciò suggerisce una possibilità più sottile

rispetto al celebrare o al liquidare l'IA: l'IA può partecipare a un processo di avanzamento (*carrying-forward*) senza essere la fonte dell'esperienza vissuta che viene portata avanti.

È quanto è accaduto anche nella stesura di questo saggio. L'IA si è rivelata molto utile: mi ha aiutato a verificare i fatti, a trovare ulteriori fonti per rafforzare o mettere in discussione l'argomentazione e ad offrire frasi capaci di esprimere ciò che intendevo dire — sebbene, il più delle volte, le sue proposte siano parole generiche incapaci di cogliere la mia sensazione percepita o il suo significato implicito. Ma persino la parola sbagliata non va sprecata: mi orienta verso ciò che è adatto, verso ciò che si sente giusto, proprio dimostrando di non esserlo.

Ogni formulazione deve comunque essere verificata rispetto alla mia personale esperienza vissuta. *È davvero questo ciò che voglio dire? Esprime l'essenza di ciò che sto cercando di rendere esplicito? C'è qualcosa che manca, che è sopravvalutato o non del tutto corretto?* Posso percepire quel: «Sì, ci siamo quasi — ma c'è qualcos'altro», oppure: «No, questo non si adatta del tutto». A quel punto si può tentare un'altra formulazione e verificarla di nuovo.

Questo richiede tempo e attenzione. La fluidità di una risposta non può sostituire quella pausa umana. Nell'interazione accade qualcosa, ma la sensazione percepita rispetto a cui le parole vengono verificate rimane la mia. L'IA può aiutare le parole a emergere, ma non può sollevarmi dalla scelta di accettarle o dalla responsabilità di ciò che infine affermo.

L'esperienza vissuta crea la possibilità della saggezza, ma non la garantisce.

Un essere umano non è automaticamente saggio per il solo fatto di vivere un'esperienza. Possiamo imbatterci nella sofferenza altrui e voltarci dall'altra parte. Possiamo sentire che qualcosa non va e razionalizzarlo. Spesso l'ostacolo è la paura: la paura di ciò che prenderne atto potrebbe richiederci, di cosa potremmo perdere o di quali conseguenze potrebbero derivarne. Il coraggio è la capacità di rimanere con ciò che c'è abbastanza a lungo da rendere possibile una risposta più libera.

L'IA espone quindi un doppio problema: la macchina può mancare dell'esperienza vissuta da cui la saggezza umana può emergere, mentre gli esseri umani possiedono l'esperienza vissuta e tuttavia falliscono ripetutamente nel permettere alla saggezza di scaturire da essa.

Paura e pericolo

Paura e pericolo non sono la stessa cosa. Il pericolo si riferisce a una minaccia concreta nella situazione. La paura è una risposta umana al pericolo, ma può anche essere risvegliata dall'attesa, dal ricordo o dalla suggestione, e può persistere anche quando non è presente alcuna minaccia immediata.

Le risposte corporee che ci aiutano a sopravvivere al pericolo possono quindi essere attivate attraverso la comunicazione. Avvertimenti ripetuti, immagini, previsioni e narrazioni possono far percepire la minaccia come continua. La paura entra così silenziosamente nella vita quotidiana: restringe l'attenzione, affretta le reazioni e fa sembrare la certezza o l'obbedienza più sicure del fermarsi a riflettere.

L'IA è altamente efficace anche in questo. Può identificare ciò che cattura l'attenzione, personalizzare i messaggi per gruppi specifici, produrre un linguaggio e immagini persuasivi e ripeterli a velocità immensa. La stessa tecnologia che può aiutarci a ricevere le prove può anche saturarci di paura.

La questione morale ritorna alla scelta umana. Chi sta usando la tecnologia? Quale risposta si cerca di produrre? Chi trae beneficio quando la paura sostituisce la riflessione?

La curiosa parola "esso"

Ed è qui che appare un altro piccolo paradosso. Ci riferiamo continuamente all'intelligenza artificiale come a un "esso".

Esso pensa. Esso vuole. Esso ha deciso. Esso ha imparato. Esso ha commesso un errore.

Questa parola è stranamente familiare a chiunque conosca il Focusing. Anche nel Focusing diciamo:

C'è qualcosa lì.

Si sente come...

Non è del tutto questo.

Ha bisogno di un po' di tempo.

Ma la parola "esso" funziona in modo molto diverso nei due casi.

Nel Focusing, "esso" ci permette di riconoscere qualcosa prima che si sia formato a sufficienza per essere nominato. Siamo curiosi e interessati a ciò che è, senza nominarlo troppo in fretta o imporgli i nostri concetti preesistenti. Rimaniamo accanto a qualcosa di implicito e gli concediamo tempo.

Proviamo una parola. *Si adatta? No, non del tutto. C'è dell'altro.* Aspettiamo. Forse arriva un'altra parola, e improvvisamente qualcosa si muove. Così, nel Focusing, "esso" mantiene aperta la domanda mentre il significato prende forma.

Nel discorso sull'intelligenza artificiale, invece, può accadere quasi l'opposto. "Esso" può nascondere gli attori umani già noti lasciandoli senza nome:

Esso ha deciso. Chi ha progettato ciò che valeva come decisione riuscita?

Esso ha selezionato. Chi ha determinato i criteri?

Esso ha previsto. Chi ha stabilito cosa dovesse accadere in base a quella previsione?

Esso ha imparato. Dal mondo, dal linguaggio, dalla storia e dai valori di chi?

La stessa identica parola che permette a un significato umano emergente di rimanere aperto può anche celare una reale responsabilità umana. C'è qualcosa di vivo in questo paradosso.

Il felt sense e la risposta

Il Focusing compie un atto profondamente controcorrente: ci permette di non sapere ancora.

Il felt sense non è semplicemente un'informazione mancante in attesa di essere recuperata. È il sapere corporeo implicito dell'intera situazione che non si è ancora formato a sufficienza per essere formulato.

Rimaniamo con esso. Qualcosa arriva. Lo verifichiamo. *Si adatta?* Talvolta la risposta è: *No, non del tutto.*

Quel piccolo "non del tutto" è fondamentale. Può racchiudere più saggezza di un'enorme quantità di informazioni perfettamente organizzate.

L'IA, al contrario, è straordinariamente abile nel fornire risposte. E forse uno dei suoi pericoli non è che le sue risposte siano sempre sbagliate, ma che possano diventare così fluide, immediate e convincenti da farci perdere gradualmente la capacità di stare con la domanda.

Forse la saggezza inizia proprio lì: dove l'intelligenza semplicemente non può più fornire la risposta; dove dobbiamo rimanere con ciò che non si adatta; dove incontriamo le conseguenze per un altro essere umano; dove il calcolo raggiunge il suo limite; dove qualcosa non ancora formulato ha bisogno di tempo; e dove, infine, dobbiamo accettare la responsabilità di ciò che sceglieremo di fare dopo.

La pausa rivoluzionaria

Mary Hendricks-Gendlin ha definito questo processo *la pausa rivoluzionaria*: un'interruzione della situazione in corso che crea lo spazio per la formazione di una sensazione percepita. Viktor Frankl individuava la libertà umana nella possibilità di scegliere la nostra risposta. Gendlin colloca le nuove possibilità nella pausa che precede il momento in cui i significati ereditati, i ruoli culturali e le reazioni automatiche prendono il sopravvento.

Cosa succede quando iniziamo a esternalizzare quella stessa pausa?

Un algoritmo ci dice sempre più spesso cosa leggere, cosa conta, chi temere, quale decisione appare ottimale, quale diagnosi è giusta, quale bersaglio militare è idoneo e quale rischio è accettabile. Il pericolo non è solo che l'IA possa un giorno acquisire troppa autonomia; il pericolo più immediato potrebbe essere che noi esercitiamo troppo poco la nostra.

La pausa non è un ritirarsi dall'azione. È il luogo in cui ciò che non è stato ancora nominato può diventare una scelta della quale siamo in grado di rispondere.

La bussola morale

L'orientamento morale non può più dipendere esclusivamente da istituzioni, ideologie, tecnologie o certezze ereditate.

Questo ci riporta alla nostra bussola morale. Ampia parte del dibattito contemporaneo sull'IA riguarda l'allineamento (*alignment*): come possiamo allineare l'intelligenza artificiale ai valori umani?

Ma quali valori umani? I valori che proclamiamo o quelli che pratichiamo? I diritti umani, gli interessi dei governi, gli interessi aziendali, gli interessi degli azionisti o gli interessi nazionali? I valori di chi possiede la tecnologia o di coloro le cui vite ne saranno influenzate?

C'è qualcosa di straordinario nell'umanità che propone di allineare le macchine ai valori umani mentre gli esseri umani rimangono così profondamente divisi sui propri.

Diciamo che la vita umana ha un valore uguale per tutti, ma le nostre azioni indicano ripetutamente il contrario. Stabiliamo il diritto internazionale e lottiamo per applicarlo in modo coerente. Costruiamo istituzioni destinate a proteggere gli esseri umani e le vediamo diventare impotenti quando intervengono forti interessi.

Sappiamo. E in qualche modo non sappiamo. Vediamo. E in qualche modo non vediamo. Sentiamo. E in qualche modo impariamo a non sentire.

L'IA non ha creato questa contraddizione: l'ha ereditata da noi.

Riflessioni filosofiche recenti affinano questo punto. Andrej Zwitter sostiene che trattare l'allineamento come se si potesse semplicemente far adottare a una macchina l'etica umana trascura le origini incarnate dell'agire morale: la vulnerabilità, l'affettività, la cura, la coscienza, le relazioni e una storia vissuta. L'onere etico rimane quindi umano e istituzionale. La domanda non è solo se la macchina sia allineata con noi, ma se noi siamo allineati con i valori che professiamo.

Una bussola non elimina l'incertezza.

Questo è importante perché una delle tentazioni più forti dell'IA è la promessa di certezza attraverso la previsione, la classificazione e l'ottimizzazione. La vita morale non può essere ridotta all'ottimizzazione. Talvolta la risposta eticamente necessaria inizia proprio quando un essere umano rifiuta la risposta apparentemente efficiente e rimane con la complessità e le conseguenze della situazione.

Ciò che è vero, lo è già

«Ciò che è vero lo è già.» — Eugene T. Gendlin

È una frase ingannevolmente semplice. Riconoscere la realtà non la crea; rifiutarsi di riconoscerla non la elimina.

Ed è proprio qui che il dibattito sull'IA deve tornare a volgersi verso l'essere umano.

Temiamo che l'intelligenza artificiale possa diventare fredda, calcolatrice, immensamente potente e distaccata dalla sofferenza umana. Le istituzioni umane sono già capaci di essere fredde, calcolatrici, immensamente potenti e distaccate dalla sofferenza umana.

Temiamo che l'IA possa manipolare la verità. Gli esseri umani già manipolano la verità.

Temiamo che l'IA tratti gli esseri umani come oggetti. I sistemi umani già lo fanno.

Temiamo che l'IA acquisisca potere senza una bussola morale.

E lo specchio chiede: *Dov'è la nostra?*

Non come un'accusa, ma come una domanda; forse come un invito.

Impotenza e il voltarsi dall'altra parte

Forse c'è un altro aspetto di tutto questo di cui non parliamo abbastanza. Oggi gli esseri umani si imbattono nella sofferenza su una scala mai vista prima. Ogni giorno i nostri schermi ci mostrano guerre, sfollamenti, fame, violenza e ingiustizie.

Possiamo venire a conoscenza della sofferenza di un altro essere umano quasi istantaneamente, ma l'individuo comune non ha frequentemente alcun potere corrispondente per cambiare ciò che sta accadendo. C'è un'enorme tensione tra ciò che possiamo vedere e ciò che possiamo fare.

Forse ci attiviamo. Forse protestiamo. Forse aiutiamo una persona. Forse disperiamo. Forse diventiamo insensibili. E forse a volte troviamo un altro problema.

L'IA fornisce uno straordinario "altrove". Possiamo regolamentarla, organizzare conferenze al riguardo, calcolarne i rischi, costruire quadri di sicurezza, immaginare futuri catastrofici e identificare qualcosa che deve essere controllato.

Forse questo è più facile che rimanere con una domanda molto più antica e dolorosa:

Perché gli esseri umani sono capaci di sapere che qualcosa è intollerabile e continuare a tollerarlo?

Non esiste una soluzione tecnologica a questa domanda.

L'IA non può prendersi la responsabilità per noi

C'è qualcosa di seducente nell'immaginare che sistemi sufficientemente intelligenti possano infine risolvere i problemi che l'umanità non è riuscita a risolvere. Forse l'IA aiuterà ad affrontare il cambiamento climatico, le malattie, l'organizzazione economica e i conflitti. Alcune di queste cose potrebbero davvero accadere.

Ma la conoscenza non può da sola dirci a cosa dovremmo dare valore. La previsione non può determinare cosa sia giusto. L'efficienza non può dirci quale tipo di mondo valga la pena di essere vissuto.

L'IA potrà forse dirci con straordinaria precisione cosa accadrà se scegliamo A invece di B. Ma la responsabilità della scelta non può semplicemente scomparire nel calcolo.

Altrimenti non abbiamo risolto il problema morale: l'abbiamo esternalizzato. E forse era proprio quello che volevamo fare.

Il dilemma umano

Forse, allora, questo non è mai stato fondamentalmente un problema di IA. È un dilemma umano. Lo è sempre stato.

Cosa fanno gli esseri umani con il potere?

Cosa succede quando la nostra capacità di agire diventa più grande della nostra capacità di vivere le conseguenze delle nostre azioni?

Cosa succede quando l'intelligenza si separa dalla saggezza?

Cosa succede quando un altro essere umano non ci sembra più reale?

Cosa succede quando il profitto diventa più importante della vita?

Cosa succede quando la responsabilità si distribuisce attraverso le istituzioni fino a quando nessuno è più responsabile?

Cosa succede quando creiamo una tecnologia capace di amplificare tutto questo oltre ogni immaginazione precedente?

Eppure c'è un altro paradosso

Sarebbe facile concludere con una condanna dell'umanità, ma ciò ripeterebbe esattamente la semplificazione che stiamo mettendo in discussione.

Gli esseri umani sono capaci di straordinaria crudeltà e di straordinaria compassione. Sfruttiamo gli altri e ci sacrifichiamo per gli sconosciuti. Ci voltiamo dall'altra parte e testimoniamo. Distruggiamo e ripariamo.

Entrambe le possibilità ci appartengono. Forse è questo il paradosso più profondo.

L'essere umano capace di ignorare la sofferenza di un'altra persona è anche capace di permettere alla vita di quella persona di entrare nella propria e di esserne trasformato.

Quindi la domanda non può essere semplicemente: *Come impediamo all'IA di diventare come la parte peggiore dell'umanità?*

Deve anche essere: *Cosa permette alle migliori possibilità dell'umanità di emergere?*

E: *Cosa le ostacola?*

Cosa richiede la responsabilità

Il compito pratico non consiste nel chiedere all'IA di diventare la nostra bussola morale. Consiste nel mantenere visibile la responsabilità umana ovunque l'IA venga utilizzata. Ciò significa nominare le persone e le istituzioni dietro le decisioni rilevanti; rifiutare che la frase «*il sistema ha deciso*» diventi un alibi; preservare la capacità umana di fermarsi, porsi domande e dire di no; esigere che chi detiene il potere risponda a coloro che ne subiscono le conseguenze; e prestare attenzione ai danni attuali preparando al contempo le risposte ai rischi futuri.

Le *Tre Bussole* diventano pratiche quando si traducono in domande nel momento della scelta:

Interiore: *Cosa sto percependo? Cosa non si adatta?*

Relazionale: *Chi ho davanti e cosa sta accadendo tra di noi?*

Collettiva: *Chi detiene il potere, chi ne subisce le conseguenze e chi rimane responsabile?*

Queste domande non producono certezze. Ci mantengono in relazione con la situazione viva abbastanza a lungo da permettere l'emergere di un responsabile passo successivo.

L'IA può informare tutte e tre le bussole, ma non deve diventare la bussola essa stessa.

L'IA può aiutarci a scorgere schemi e possibilità. Può assistere il pensiero. Può persino offrire parole che portano avanti qualcosa. Ma al punto in cui un'altra vita può essere alterata, non si può permettere che la scelta diventi moralmente priva di un proprietario.

L'atto finale rimane umano: guardare, scegliere e rispondere della propria scelta.

Guardare nello specchio

Forse l'intelligenza artificiale sta offrendo all'umanità un'opportunità inaspettata. Non perché ci dirà cosa è giusto. Non perché risolverà il dilemma umano. E certamente non perché la responsabilità possa esserle affidata in tutta sicurezza.

Ma perché le domande che ci poniamo sull'intelligenza artificiale ritornano con straordinaria precisione verso noi stessi.

Cos'è l'intelligenza? Cos'è la saggezza? Cos'è la responsabilità? Cos'è la verità?

Cosa succede quando il potere si concentra?

Cosa significa fare esperienza di un altro essere umano?

Cosa rende qualcosa importante? Cos'è una bussola morale? E a cosa siamo allineati?

Forse l'IA non può rispondere a queste domande per noi. Forse nemmeno noi possiamo dar loro una risposta definitiva. Ma c'è una differenza profonda tra il non avere una risposta definitiva e smettere di chiederselo.

Forse una delle cose più importanti dell'essere umano è proprio la sua capacità di rimanere accanto a qualcosa per cui non ha ancora una risposta.

«Ciò che non si è ancora formato può comunque essere percepito.» — Eugene T. Gendlin

Notare: *No. Questo non si adatta del tutto.*

Dargli tempo. Permettere a ciò che non si è ancora formato di emergere. Esperire le conseguenze su un altro essere umano. E infine dire:

Io scelgo.

Non: Il sistema me lo ha fatto fare.

Non: L'algoritmo ha deciso.

Non: L'IA me lo ha detto.

Ma: Io scelgo — e quindi sono responsabile.

La crisi dell'intelligenza artificiale potrebbe infine rivelarsi come qualcosa di molto più antico: una crisi della responsabilità umana.

L'IA non è semplicemente qualcosa che l'umanità ha creato. Potrebbe diventare uno specchio in cui ci viene chiesto di incontrare ciò che abbiamo fatto di noi stessi.

E forse la distinzione di cui abbiamo più urgentemente bisogno non è tra intelligenza umana e intelligenza artificiale, ma tra intelligenza e saggezza.

L'intelligenza artificiale può diventare straordinariamente abile nel rispondere alle nostre domande. La saggezza umana potrebbe dipendere dalla nostra capacità di riconoscere quando una risposta non è sufficiente.

Lo specchio non può dirci chi siamo; può solo restituire l'immagine. Spetta ancora a noi guardare.

E guardare è sgradevole perché comporta qualcosa che impedisce di dare la colpa allo specchio: accettare la responsabilità per ciò che facciamo — e per ciò che non facciamo.

Forse è per questo che è così allettante guardare altrove. Forse incolpare lo specchio è più facile che guardarci dentro.

«Ciò che è vero lo è già.» — Eugene T. Gendlin

E lo specchio è ancora lì.

Riferimenti e letture di approfondimento

Jacob Coxon, dimissioni e avvertimenti. *Ars Technica*, 9 settembre 2026.

Committee to Protect Journalists. *Israel Gaza War Attacks on Journalists Across the Region Since October 2023*.

Reporters Without Borders. *Gazawood The Deadly Accusations Against Gaza's Journalists*, 15 settembre 2025.

Human Rights Watch. *Questions and Answers Israeli Military's Use of Digital Tools in Gaza*, 10 settembre 2024.

Yuval Abraham. *Lavender The AI Machine Directing Israel's Bombing Spree in Gaza*. +972 Magazine, 3 aprile 2024.

Pacing the Frontier. Dichiarazione dei dipendenti delle aziende di IA di frontiera, luglio 2026.

Anthropic. *Responsible Scaling Policy*, aggiornato al 14 agosto 2026.

Eugene T. Gendlin. *Experiencing and the Creation of Meaning*. Free Press, 1962.

Eugene T. Gendlin. *Focusing*. Everest House, 1978.

Mary Hendricks-Gendlin. *Focusing as a Force for Peace: The Revolutionary Pause*. Discorso di apertura, 2003.

Viktor E. Frankl. *Man's Search for Meaning* (Uno psicologo nei lager).

Joseph Weizenbaum. *On the Impact of the Computer on Society*. Science 176, 1972.

Joseph Weizenbaum. *The Computer in Your Future*. The New York Review of Books, 27 ottobre 1983.

Joseph Weizenbaum. *Computer Power and Human Reason: From Judgment to Calculation*. W H Freeman, 1976.

Norbert Wiener. *God and Golem Inc: A Comment on Certain Points Where Cybernetics Impinges on Religion*. MIT Press, 1964.

Andrej Zwitter. *The Embodied Ethics Alignment Problem of AI*. Discover Artificial Intelligence, vol. 6, articolo 852, 2026.

Onora O'Neill. *A Question of Trust: The BBC Reith Lectures 2002*. Cambridge University Press, 2002.

Claudia Núñez-Pacheco, Pedro Sanches e Jorge Olivares-Retamal. *Searching for the Words that 'Feel Right': Resonating with our Bodies and Felt Senses Through Haiku and Large Language Models (LLMs)*. Proceedings of the 2025 ACM Designing Interactive Systems Conference (DIS '25).

Patricia Foster. *The Three Compasses: A Way of Being with Situations*.

(*) Nota sull'autrice

Patricia Foster è Focusing Coordinator e Trainer dell'International Focusing Institute, oltre che Terapeuta Centrata sulla Persona certificata EAP. Il suo lavoro integra la filosofia esperienziale nella formazione sul Focusing e nella terapia Centrata sulla Persona orientata al Focusing. Per consultare altri scritti dell'autrice: www.focusinglifeskill.com

Dedica

Lo scritto è in memoria di Maher Al-Afifi, giornalista e regista, che ha trascorso anni a documentare la sofferenza altrui con dignità e umanità e a sostegno di sua moglie Reham e dei loro quattro figli, Roa, Mahmoud, Fatima e Lina, che ora portano il peso di ciò che la sua morte ha lasciato.

Chi volesse contribuire alla raccolta fondi a favore di Reham e i suoi figli, può offrire il proprio sostegno tramite GoFundMe: <https://gofund.me/26d92a7be>