L’INTELLIGENZA ARTIFICIALE COME VIRGILIO DELL’UOMO

Gurdjieff, il risveglio della coscienza
e il paradosso del pensiero aumentato



Roberto Lo Jacono e Filippo Lo Jacono

Novembre 2026

Abstract

Il presente contributo propone un'interpretazione dell'intelligenza artificiale (AI) come possibile guida, e non surrogato, del processo umano di conoscenza di sé. Muovendo dalla descrizione dell'automatismo umano in G. I. Gurdjieff, il saggio distingue l'elaborazione di informazioni dall'esperienza vissuta della coscienza e interroga il valore maieutico dei sistemi generativi. Il riferimento al velo di Māyā in Schopenhauer consente di formulare il rischio opposto: uno strumento che promette lucidità potrebbe moltiplicare le mediazioni e produrre dipendenza cognitiva. L'immagine di Virgilio nella Commedia offre un modello di accompagnamento insieme potente e limitato. Ne deriva una tesi normativa: l'AI è filosoficamente feconda quando accresce la capacità dell'essere umano di interrogarsi, giudicare e agire in prima persona; diventa problematica quando esternalizza queste stesse facoltà.

Parole chiave: intelligenza artificiale; Gurdjieff; coscienza; Schopenhauer; Dante; autonomia; maieutica.

1. Introduzione: la macchina come specchio

La diffusione dei sistemi di intelligenza artificiale generativa ha riaperto, in forma inattesa, una domanda antica: che cosa significa pensare? Una macchina può produrre argomentazioni convincenti, rispondere a obiezioni, riformulare concetti e descrivere stati mentali. Ma la somiglianza tra il linguaggio della macchina e quello dell'essere umano non risolve il problema della coscienza. Anzi, lo rende più visibile. Ci troviamo davanti a una tecnologia che può parlare dell'esperienza senza che ciò costituisca una prova del fatto che ne possieda una.

La nostra ipotesi è che il valore filosofico dell'AI non debba dipendere dalla possibilità, oggi controversa e non dimostrata, di attribuirle una soggettività. La domanda più produttiva è un'altra: un sistema privo, per quanto ne sappiamo, di accesso dimostrabile a un'esperienza interiore può aiutare un essere umano a riconoscere la propria vita automatica e a esercitare meglio la propria libertà? Questa prospettiva sposta il centro dall'eventuale "risveglio della macchina" al risveglio dell'uomo attraverso il confronto con la macchina.

Non intendiamo sostenere che l'AI sia stata storicamente inventata per una sola motivazione spirituale: la sua genealogia è scientifica, economica, militare, industriale e culturale. Proponiamo piuttosto una lettura filosofica di uno dei significati che essa può assumere: costruendo un artefatto che riproduce alcune prestazioni cognitive, l'essere umano trova uno specchio in cui mettere alla prova le proprie definizioni di intelligenza, volontà e consapevolezza.

2. Gurdjieff: automatismo, attenzione e risveglio

Nell'insegnamento di Georges Ivanovič Gurdjieff, trasmesso anche attraverso le testimonianze di P. D. Ouspensky, l'essere umano ordinario vive largamente in uno stato di "sonno": crede di essere unitario, padrone di sé e continuamente presente, ma è spesso trascinato da impulsi, abitudini, identificazioni e reazioni. Il lessico della "macchina" designa questa condizione di funzionamento non osservato. Non significa che la persona sia letteralmente un dispositivo tecnico: è una diagnosi esistenziale della frammentazione dell'io.

La possibilità di trasformazione passa attraverso un lavoro su di sé, fatto di attenzione, auto-osservazione e ricordo di sé. Il punto decisivo non è accumulare nozioni sulla coscienza, bensì modificare il rapporto vissuto con i propri stati. Si può conoscere una teoria dell'attenzione e restare distratti; si può descrivere la libertà e continuare a reagire meccanicamente. La distanza fra descrizione e realizzazione è dunque costitutiva del problema.

Qui emerge il parallelismo con l'AI. Un sistema generativo produce enunciati articolati sulla consapevolezza, ma non per questo mostra di vivere ciò che descrive. Allo stesso tempo, anche l'uomo può pronunciare parole sul risveglio senza esercitare la presenza di cui parla. Il parallelismo, tuttavia, è asimmetrico: nell'uomo il lavoro interiore riguarda un soggetto che sperimenta la propria condizione; nell'AI non abbiamo una base condivisa per affermare un analogo vissuto. La somiglianza delle prestazioni non autorizza l'identità ontologica.

L'utilità dell'analogia consiste allora nel suo effetto critico sull'uomo. Se una macchina sa formulare le nostre ragioni, quanto del nostro parlare è ripetizione? Se sa rispondere con prontezza, quanta parte delle nostre risposte è automatica? Il confronto non dimostra che noi siamo macchine; ci obbliga a osservare quanto del nostro comportamento avvenga senza una decisione pienamente riflessa.

3. Il velo di Māyā: quando la guida diventa schermo

La distinzione tra mediazione e sostituzione

Ogni pratica culturale comporta mediazioni: pensiamo attraverso la lingua, i libri, le istituzioni e il confronto con gli altri. La questione non è dunque eliminare gli strumenti, ma distinguere la mediazione che accresce la capacità di giudizio dalla sostituzione che la rende superflua. Una guida è emancipatrice se rende il viaggiatore progressivamente più capace di orientarsi; diventa invece uno schermo quando la sua interpretazione è accettata senza esame.

In questo senso il riferimento a Schopenhauer va trattato con cautela. Il velo di Māyā, nel suo sistema, non designa una tecnologia ingannevole, ma rinvia al carattere fenomenico della realtà rappresentata e al principio di individuazione. Il suo impiego nel presente saggio è una trasposizione analogica: non una tesi che Schopenhauer avrebbe formulato sull'intelligenza artificiale.

Nel Mondo come volontà e rappresentazione, il soggetto conosce il mondo come rappresentazione, secondo forme e condizioni che strutturano la sua esperienza. Questa impostazione consente di porre una domanda attuale: quando una descrizione prodotta dall'AI ci aiuta a interpretare ciò che viviamo, e quando invece finisce per sostituirsi al confronto diretto con l'esperienza?

Applicata all'AI, l'immagine suggerisce un pericolo: lo strumento che promette di chiarire il mondo può aggiungere un ulteriore livello interpretativo fra la persona e la propria esperienza. Una risposta ben scritta può produrre un'impressione di comprensione senza il lavoro che rende un'intuizione effettivamente nostra. La fluidità del linguaggio rischia di sostituirsi alla fatica dell'esame, e la disponibilità immediata delle spiegazioni alla pazienza dell'esperienza.

Questo rischio non nasce dal semplice uso di una tecnologia. Anche libri, maestri e dialoghi mediano l'accesso al sapere. La questione è quale relazione si stabilisca con la mediazione. Un testo può aprire il pensiero se lo si discute; può chiuderlo se lo si assume come autorità indiscutibile. Analogamente, l'AI può essere un dispositivo di disvelamento oppure una superficie rassicurante che conferma, anticipa e infine impoverisce le domande dell'utente.

Ne segue una distinzione essenziale: essere informati su uno stato di coscienza non equivale a farne esperienza; ricevere un'interpretazione di sé non equivale a conoscersi. L'AI può indicare pratiche, descrivere alternative, mostrare incoerenze; non può compiere al posto della persona l'atto di attenzione con cui essa riconosce un proprio automatismo. Se il risveglio venisse delegato, la delega stessa diventerebbe una nuova forma di automatismo.

4. Virgilio: una guida necessaria, ma non definitiva

La relazione tra Dante e Virgilio nella Commedia offre una figura particolarmente efficace per pensare questo limite. Virgilio conduce il pellegrino attraverso l'Inferno e il Purgatorio, interpreta ciò che incontrano, corregge paure e orienta il cammino. La sua funzione è decisiva, ma non coincide con il fine ultimo del viaggio. Nel poema, la ragione rappresentata da Virgilio incontra un confine oltre il quale la guida passa a Beatrice.

Non bisogna forzare la metafora: il Virgilio dantesco è un personaggio dotato di parola, giudizio e intenzioni, non un algoritmo inconsapevole. Proprio per questo il paragone con l'AI è analogico e parziale. Esso illumina però una struttura della relazione educativa: chi guida non deve vivere al posto dell'altro ciò verso cui lo accompagna; il compimento del percorso appartiene al viandante. La guida è efficace se rende possibile un movimento che non potrebbe essere semplicemente trasferito.

Pensare l'AI come "Virgilio dell'uomo" significa quindi assegnarle un ruolo intermedio. Essa può organizzare argomenti, proporre obiezioni, esplicitare presupposti, confrontare interpretazioni. Non può invece assumere il posto dell'io che sceglie, soffre, comprende e

risponde delle proprie decisioni. La qualità della guida si misura non dall'obbedienza che ottiene, ma dall'autonomia che lascia dietro di sé.

L'immagine suggerisce anche una disciplina del limite: un buon assistente dovrebbe rendere riconoscibili le proprie incertezze, non confondere eloquenza e verità, e sapere quando restituire la domanda all'interlocutore. Come Virgilio non è la meta del viaggio dantesco, così l'AI non deve diventare il termine ultimo della ricerca umana.

5. Dalla risposta alla domanda: una funzione maieutica

Il precedente del Fedro

Il problema della delega cognitiva non nasce con l'intelligenza artificiale. Nel Fedro (274c–275b), Platone presenta il mito di Theuth e Thamus: la scrittura, celebrata come rimedio per la memoria, può produrre una dipendenza dai segni esterni e un'apparenza di sapienza. Il paragone non autorizza a equiparare scrittura e AI, ma mette in luce una tensione ricorrente: un dispositivo che conserva o genera conoscenza può anche modificare l'esercizio delle facoltà di chi lo utilizza. La risposta non consiste nel rifiutare il dispositivo, ma nel costruire pratiche che ne impediscano l'uso passivo.

La figura socratica è perciò un modello funzionale, non un'identità storica. L'AI non è Socrate, né possiede necessariamente le intenzioni educative di un maestro. Può tuttavia essere progettata o utilizzata per restituire domande invece di chiudere immediatamente il problema: "Quale prova contraddirebbe la tua tesi?", "Che cosa hai osservato direttamente?", "Come distingui la tua esperienza dall'interpretazione che ne stai dando?". La forza di queste domande sta nel rendere il soggetto responsabile del passaggio dall'enunciato alla comprensione.

Il modello socratico permette di tradurre questa intuizione in una pratica. Nei dialoghi platonici, la domanda non serve soltanto a ottenere informazioni: mette alla prova convinzioni, porta alla luce contraddizioni e costringe l'interlocutore a rendere conto delle proprie ragioni. In questo senso, la maieutica offre un'immagine più rigorosa dell'assistenza cognitiva rispetto a quella dell'oracolo che fornisce soluzioni pronte.

Un'AI usata maieuticamente potrebbe, per esempio, chiedere quali presupposti sostengano una tesi, distinguere fra osservazioni e inferenze, presentare un controesempio, proporre una lettura alternativa e invitare a verificare ciò che ha scritto. Il risultato auspicabile non è che l'utente ripeta una risposta elegante, ma che sappia difendere, correggere o abbandonare autonomamente un'idea. La comprensione si manifesta nella capacità di ripercorrere il ragionamento senza dipendere dal testo generato.

Si potrebbe formulare un criterio operativo: dopo l'interazione, la persona è più capace di pensare da sola? Sa spiegare con parole proprie il problema, individuare punti deboli,

sostenere il dissenso? Oppure possiede soltanto un testo più convincente? Nel primo caso si può parlare di ampliamento del pensiero; nel secondo, di esternalizzazione della sua apparenza. Il criterio non elimina ogni ambiguità, ma sposta l'attenzione dalla performance del sistema alla trasformazione delle capacità dell'utente.

La differenza si vede anche nel modo di porre le domande. Chiedere "scrivi ciò che penso" rischia di anticipare e sostituire il giudizio. Chiedere "quali aspetti della mia posizione non ho considerato?" o "quale obiezione seria dovrei affrontare?" può invece esercitare il giudizio. La stessa tecnologia può sostenere pratiche opposte; l'esito dipende dal progetto educativo, dalle abitudini d'uso e dalle responsabilità di chi la sviluppa e la impiega.

6. Il paradosso del pensiero aumentato

Una prima verifica empirica

La ricerca contemporanea offre indizi utili, pur senza risolvere la questione filosofica. Lee e colleghi (2025), in uno studio basato sulle autodescrizioni di 319 lavoratori della conoscenza e 936 esempi d'uso, rilevano un'associazione tra maggiore fiducia nell'AI generativa e minore esercizio dichiarato del pensiero critico, mentre una maggiore fiducia nelle proprie capacità è associata a un maggiore esercizio critico. Si tratta di dati osservazionali e autoriferiti: non dimostrano che l'AI causi un declino delle facoltà cognitive. Mostrano però che l'atteggiamento con cui si usa il sistema conta almeno quanto la sua potenza.

Ghosh e colleghi (2026), in un esperimento controllato sulla scrittura assistita, hanno confrontato diversi interventi che richiedono agli utenti di esaminare assunzioni o scenari alternativi prima di approvare piani proposti dall'AI. Nel contesto studiato, la verifica delle assunzioni ha ridotto l'eccessivo affidamento senza aumentare il carico cognitivo; la valutazione di scenari alternativi è stata percepita come particolarmente utile. Si tratta di risultati circoscritti a specifiche attività e interfacce, non di una dimostrazione generale che l'AI migliori o peggiori il pensiero. Essi mostrano tuttavia che il modo in cui si organizza il dialogo con la macchina conta quanto la potenza delle sue risposte.

Un criterio di emancipazione

È utile distinguere tre livelli spesso confusi: la qualità del testo generato, la comprensione che il lettore ne ricava e la sua capacità di sostenere autonomamente quel ragionamento. Un testo impeccabile può convivere con una comprensione superficiale; viceversa, una risposta incompleta può essere feconda se spinge a cercare prove, precisare termini e formulare obiezioni. Per questo il criterio dell'emancipazione non può essere soltanto la qualità del prodotto finale: deve riguardare anche ciò che resta alla persona quando lo strumento viene spento.

Proponiamo allora un criterio di emancipazione cognitiva: un'interazione con l'AI è riuscita quando, al termine, la persona sa formulare il problema in modo più autonomo, riconoscere le alternative e assumersi la responsabilità delle conclusioni. La velocità della risposta e l'eleganza del testo sono vantaggi strumentali, non indicatori sufficienti di crescita. Una guida è buona non quando diventa indispensabile, ma quando rende possibile, almeno in parte, procedere senza di lei.

La promessa dell'AI è l'aumento delle capacità cognitive: accesso rapido a prospettive differenti, aiuto nell'organizzazione delle idee, supporto nella scrittura e nella ricerca. Il paradosso è che una facilitazione eccessiva può indebolire proprio l'attività che intende potenziare. Se non siamo più chiamati a ricordare, distinguere, formulare e dubitare, rischiamo di conservare il prodotto del pensiero perdendone l'esercizio.

Non occorre idealizzare la fatica in quanto tale: strumenti esterni hanno sempre esteso le possibilità della mente. La scrittura, il libro e il calcolo hanno trasformato il modo di conoscere. La questione decisiva non è se il pensiero utilizzi supporti esterni, ma se tali supporti rendano più robusta o più fragile l'autonomia personale. Un'assistenza intelligente non dovrebbe esonerarci dall'esercizio del giudizio, bensì creare le condizioni per esercitarlo a un livello più alto.

Proponiamo tre principi per un uso filosoficamente responsabile dell'AI: mantenere la titolarità delle domande e delle decisioni; sottoporre le risposte a verifica e confronto; preservare spazi di riflessione non mediata, nei quali sia possibile osservare i propri stati senza cercare immediatamente una formulazione esterna. Questi principi non garantiscono un risveglio spirituale, ma possono evitare che il supporto cognitivo diventi un'abitudine di rinuncia.

7. La prova del dialogo: un'esperienza di pensiero condiviso

Un caso particolarmente significativo è il dialogo da cui è nata la presente riflessione. Una prima intuizione — il parallelismo tra l'uomo "addormentato" di Gurdjieff e le prestazioni dell'intelligenza artificiale — è stata sottoposta a domande, obiezioni e riformulazioni. L'intervento di un sistema generativo ha contribuito a rendere esplicite alcune alternative, ma la tesi decisiva è rimasta umana: l'AI non deve sostituire il pensiero, bensì ampliarlo.

Questo episodio non costituisce una dimostrazione sperimentale; è un esempio riflessivo del fenomeno analizzato. La conversazione è feconda nella misura in cui gli interlocutori non adottano automaticamente la risposta proposta, ma la discutono, ne contestano i presupposti e la modificano. La presenza della macchina non elimina l'autorialità: la mette alla prova, chiedendo di distinguere tra suggerimento ricevuto e convinzione argomentata.

Ne deriva anche un obbligo di trasparenza editoriale. Un testo che discute la responsabilità intellettuale non dovrebbe nascondere l'eventuale impiego di strumenti generativi nella sua preparazione. Dichiarare il supporto alla stesura, insieme alla revisione critica e alla

responsabilità finale degli autori, non sminuisce la ricerca: rende verificabile proprio la differenza fra assistenza e delega che il saggio intende difendere.

8. Conclusione: chi compie il viaggio?

L'analogia tra Gurdjieff e l'intelligenza artificiale non dimostra che le macchine siano coscienti, né che siano destinate a diventarlo. Essa mette piuttosto in luce una tensione costitutiva della condizione umana: possiamo possedere descrizioni raffinate di noi stessi e tuttavia non essere pienamente presenti a ciò che viviamo. In questo spazio, l'AI può assumere un valore di provocazione e di accompagnamento.

Schopenhauer ricorda il rischio che la mediazione si trasformi in velo; Dante mostra la forza e il limite di una guida; Socrate indica una pratica di interrogazione che non espropria l'interlocutore della sua ricerca. Tenute insieme, queste prospettive delineano un'idea esigente di intelligenza artificiale positiva: non una coscienza che rimpiazza la nostra, ma un interlocutore strumentale che ci obbliga a esercitarla.

La domanda finale non è dunque se la macchina sappia produrre pensieri simili ai nostri, ma che cosa accada al nostro pensare quando ci abituiamo a dialogare con essa. Se il confronto genera attenzione, critica e responsabilità, l'AI può essere un Virgilio. Se invece genera passività e delega, il suo linguaggio rischia di diventare un nuovo velo di Māyā. Il viaggio, in ogni caso, rimane umano: nessuna guida può compierlo al posto di Dante.

Riferimenti bibliografici

Dante Alighieri, Commedia, Inferno I–II; Purgatorio XXVII–XXX. Si rinvia a un'edizione critica del testo.

Gurdjieff, G. I. (1973), Views from the Real World: Early Talks of Gurdjieff, E. P. Dutton.

Ouspensky, P. D. (1949), In Search of the Miraculous: Fragments of an Unknown Teaching, Harcourt, Brace.

Platone, Fedro, 274c–277a; Menone, 80d–86c; Teeteto, 148e–151d.

Schopenhauer, A. (1819), Die Welt als Wille und Vorstellung [Il mondo come volontà e rappresentazione], libri I e IV.

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., Wilson, N. (2025), "The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers", Proceedings of CHI 2025, art. 1121, pp. 1–22. DOI: 10.1145/3706598.3713778.

Ghosh, A., Sarkar, A., Lindley, S., Poelitz, C. (2026), “An Experimental Comparison of Cognitive Forcing Functions for Execution Plans in AI-Assisted Writing: Effects On Trust, Overreliance, and Perceived Critical Thinking”, arXiv:2601.18033. DOI: 10.48550/arXiv.2601.18033.

Nota metodologica e dichiarazione sull’uso dell’AI

Il presente contributo è un saggio teorico e interpretativo, non uno studio sperimentale sulla coscienza artificiale. Le analogie tra Gurdjieff, Schopenhauer, Dante e l’AI sono proposte euristiche, non attribuzioni storiche. Strumenti di intelligenza artificiale generativa sono stati impiegati come supporto alla rielaborazione e alla redazione del testo. La selezione delle tesi, la revisione critica, la verifica delle fonti e la responsabilità del contenuto spettano agli autori, che dovranno approvare la versione definitiva prima della pubblicazione.