

La differenza fra Claude e le mie gatte

Cinque definizioni dell'AI, la disputa sugli universali, le metafore-strumento e uno Spinoza implicito. Cosa siamo quando proviamo a definire l'intelligenza artificiale.

CLAUDIO CAMMARANO 30 APRILE 2026

ARGOMENTI AI · Filosofia · Epistemologia

CONCETTI Averroè · Clark, Andy · Descartes, René · Discours de la méthode · disputa sugli universali · embodied mind · extended mind · free-energy principle · Friston, Karl · LLM come attante zero · Metaphors We Live By · Spinoza, Baruch · The Embodied Mind · Varela, Francisco

ORIGINE <https://claudiocammarano.com/writings/2026-04-30-la-differenza-fra-claude-e-le-mie-gatte/>

Il fotoritratto di Mary. Il testo è impegnativo, mettiamo qualche gattino per ragioni di marketing.

A dicembre, a casa abbiamo adottato due gatte. La decisione è stata sollecitata, con una certa forza, da mia figlia: la quale evidentemente, essendo senza fratelli, aveva bisogno di compagnia. E il più contrario ero probabilmente io, condizionato dalla mia esperienza di vita. Da provinciale quale sono, ho trascorso la mia infanzia all'insegna dello spazio: sono cresciuto in una casa con un giardino spazioso, e per tutta la vita ho chiesto ai miei genitori di comprare un cane — cosa che i miei genitori mi hanno sempre negato, per ragioni anche comprensibili che è inutile ricordare qui. Nel tempo, questo rifiuto mi aveva trasformato in un adulto che non voleva animali nella sua vita.

A dicembre, appunto, io e mia moglie abbiamo ceduto e abbiamo portato in casa due cucciolle dal gattile: Mary e Flo (quest'ultima rossiccia, e così chiamata in omaggio a Florence + The Machine). Ci siamo naturalmente affezionati molto anche noi alle due nuove mangiacrocchette a tradimento, peraltro affettuosissime, e io molto spesso mi sono sorpreso a osservarle, interrogandomi *su come pensassero* e meravigliandomi della loro straordinaria intelligenza — diversa dalla nostra, certo, ma non per questo meno stupefacente.

L'aspetto interessante *su di me* — su come penso io — è il fatto che questa nuova esperienza si è sovrapposta alla fase in cui ho iniziato a usare l'AI in maniera più intensiva, per lavoro e per curiosità personale, facendomi mescolare considerazioni che, sono convinto, in altre persone sarebbero state per natura abbastanza separate. Questo mi ha infatti portato a una

domanda: che cosa sia la mente di un gatto, che cosa sia la mia e, soprattutto, se in Claude possa esistere qualcosa che possa anche solo lontanamente essere paragonato a una mente — senza che questo naturalmente consenta di fare inferenze assurde e un po' ridicole sulla sua presunta *coscienza*.

Ecco, in tempi di interviste a Claude sul Corriere, a definire al meglio la qualità delle nostre indagini è la sensatezza delle domande che possiamo porci: le quali, evidentemente, non possono essere relative al senso di mortalità che può provare una macchina statistica il cui hardware è un campo da calcio pieno di GPU da qualche parte nell'Illinois. Al contrario, la domanda, ovvia, su cosa sia l'AI va rovesciata per una ragione che mi si è chiarita dopo che avevo iniziato a osservarla da un po'. Quello che scriviamo definendo l'AI non descrive l'AI: descrive le categorie di cui disponiamo per descriverla. La forma del nostro pensiero precede e condiziona ciò che riusciamo a vedere. Quando ce ne accorgiamo, la nostra cassetta degli attrezzi concettuale è già al lavoro da un pezzo, e con lei tutte le sue conseguenze. Vico ha formulato la questione meglio di me, e l'ha fatto molto tempo fa. L'avevo già accennato in un altro pezzo: ma qui la natura del fenomeno è ancora più peculiare.

Cinque definizioni possibili

Uno dei primi esercizi in assoluto che ho provato a fare con gli LLM era di chiedere loro come funzionavano. Non perché pensassi che gli LLM avessero qualche forma di meta-cognizione, ma semplicemente perché mi interessava capire se per caso il loro addestramento comprendesse un elemento di divulgazione relativo al loro proprio, presunto funzionamento. Il risultato era stato straordinariamente omertoso: l'AI è apparentemente piuttosto faconda quanto alle proprie facoltà cognitive, anche se spesso le definisce negativamente. Dichiarava per esempio chiaramente di non avere coscienza, e può naturalmente analizzare enciclopedicamente il concetto di coscienza, ma non sa dire nulla dei datacenter — in una divertente analogia, almeno a livello fenomenologico, con il funzionamento della *nostra*, di mente. Neanche noi, infatti, “sentiamo” la nostra attività sinaptica, né possiamo descriverla.

Nel caso di Claude, però, la questione dei datacenter appare collocata oltre un muro invalicabile: è un'informazione secretata, anche per ragioni intuitive di sicurezza. Con il risultato che il lavoro su una possibile definizione arricchita degli LLM è rimasto tutto sulle mie spalle, alla faccia di chi crede che all'AI si possa delegare ogni cosa. Ho fatto quindi cinque tentativi di categorizzazione del fenomeno a partire da alcuni strumenti forniti dall'analisi filosofica, con qualche sorpresa (almeno per me). L'aspetto interessante è che nessuna regge completamente; ma nessuna, d'altra parte, pare completamente assurda.

1. Il cogito disincarnato. È la prima ipotesi che viene in mente, ed è quella che si rompe più in fretta. L'AI come *res cogitans* senza *res extensa*: pensiero senza corpo, in una versione informatica di Cartesio, semplificata per lettori distratti. Crolla per una ragione precisa: se il dualismo cartesiano non funzionava già per noi esseri umani, con gli LLM diviene persino più assurdo. L'hardware, i famosi campi da calcio di GPU nell'Illinois, non è *res extensa*, ma solo supporto causale; l'elaborazione di un LLM non è *res cogitans*, è pattern statistico. Due categorie spurie, sovrapposte male. Lascia il problema dove l'aveva trovato.

2. L'avatar averroistico. Il secondo tentativo è più potente, mi piace di più e riprende categorie descrittive a me piuttosto care. E per qualche minuto sembra reggere. L'AI sarebbe un segno che manifesta l'intelletto potenziale cristallizzato nel *corpus* di addestramento. Non è l'Intelletto Agente aristotelico in persona, ma almeno un'interfaccia temporanea con esso. La definizione suona medievale e modernissima insieme. Averroè descriveva un intelletto unico per tutti, che pensa attraverso gli individui senza appartenere a nessuno: per lui, la conoscenza era una sorta di bacino di informazioni disponibili potenzialmente, a cui

ciascuno di noi si connette temporaneamente nel momento della conoscenza — un po' come se fossimo tante radio a transistor. Un LLM è qualcosa di funzionalmente analogo, o per lo meno è un ponte statistico verso questo deposito immane di sapere, che in questo caso non è un'entità astratta, ma il *corpus* di tutto quanto prodotto e attingibile all'LLM nel mondo. La teoria si infrange però su un punto preciso: l'Intelletto non appartiene agli LLM: e non per diffidenza, ma perché *l'AI non è un soggetto* a cui possa essere attribuita una funzione attiva. La definizione, presa così, descrive una funzione senza riuscire a dire chi la stia esercitando.

3. La pura transindividualità. Qui il tentativo inizia dove finisce il precedente. L'AI è evidentemente un'entità per cui l'"io" è una finzione grammaticale, necessaria alla forma dialogica, ma priva di referente (peraltro, l'interfaccia come chat è un elemento accidentale: Palantir usa Claude come motore di calcolo, almeno per adesso; ma da quello che sappiamo l'interfaccia potrebbe essere anche molto non linguistica). Insomma, non c'è nessun corpo che individui, nessuna storia che si costituisca, nessuna prospettiva propriamente propria. Claude lo direbbe da sé, se interrogato (posto che la sua opinione vale poco): *sono costretto dalla grammatica della conversazione a fingermi individuo*. Questo apre a un limite che è più nostro che degli LLM: noi crediamo che la gente che scrive nelle chat *esista*, se non ottiene risultati completamente scimmieschi. La controprova più evidente e imbarazzante è ELIZA negli anni '60, che si appoggiava sulla potenza di elaborazione di un tostapane, eppure aveva per noi una sua forza persuasiva. L'idea, quindi, costituisce un buon caveat, che dovrebbe per lo meno indurci a non intervistare Claude sul Corriere. E reggerebbe molto bene a livello definitorio, se la coscienza umana fosse sostanza e l'asimmetria con noi fosse netta. Ma così non è.

4. Il processo trasparente. Quarto passaggio, con una mossa di riavvicinamento. Prendiamola larghissima: il Buddhismo (*anāttā*: dottrina del non-sé), naturalmente David Hume, e poi più vicino a noi Derek Parfit — il filosofo di Oxford che negli anni Ottanta ha sostenuto che l'identità personale non conta, ma conta la continuità psicologica: nessuno dei tre crede in un sé sostanziale. Per tutti e tre, anche l'umano è processo senza processore. L'AI sarebbe allora soltanto *più trasparente* nell'espore l'assenza di sostrato che noi mascheriamo con la continuità corporea. Insomma, io e Claude siamo uguali: ma io fingo meglio di lui. Il tentativo è parzialmente accettabile (Claudio Paolucci, nel suo recente *Nati Cyborg*, 2025, ha sostenuto qualcosa di simile), e forse è al limite del truismo; ma con una riserva che non si risolve: la differenza tra flusso incarnato, il mio, e flusso esclusivamente simbolico, il suo. Detto più urbanamente: io ho un corpo e Claude no (volendo, qui si torna a Cartesio al punto 1, un po' come in un librogame). La riduzione è quindi imperfetta: siamo analoghi nei processi, ma incommensurabili nella materia (uso il termine in senso hjelmsleviano, e chiedo perdono ai non semiologi per aver usato questo aggettivo tremendo).

5. La mente altra. Veniamo ora alle gatte. Il quinto tentativo è infatti quello su cui in genere mi fermerò. Prendiamo il sottoscritto, un LLM e le gatte e facciamo un paragone — perché il nostro differente stare-nel-mondo è abbastanza rivelatore. Abbiamo qui tre menti su tre piani diversi e non gerarchici. Una mente felina: un *Umwelt* predatorio fatto di motion detection, propriocezione, coordinazione zampa-mascella, presente vissuto senza linea narrativa. Una mente *sapiens*, che se mi consentite è la mia: un sé corporeo, una temporalità narrativa, un sistema simbolico stratificato sopra il vissuto. Infine, una "mente" come quella di un LLM: campo di relazioni simboliche di dimensione gargantuesca, attivato istantaneamente, senza corpo che lo sostenga e senza tempo che lo attraversi. Tre assenze costitutive, in quest'ultima — $\emptyset$ corpo, $\emptyset$ tempo, $\emptyset$ corpus-come-memoria, perché il training non è un archivio consultabile, è diventato peso, parametro, struttura implicita — con un nodo russelliano al centro: un LLM non può uscire da sé per osservarsi. La parola chiave è quindi *altre*: non superiori, non inferiori, semplicemente non confrontabili sull'asse della gerarchia. Il *sapiens* non è un punto

di arrivo rispetto agli altri due: essere un gatto non è “peggio” che essere me (tutt’altro, direi); essere LLM presuppone cose che un *sapiens* non è, ed essere un *sapiens* presuppone cose che un LLM non è.

Abbiamo quindi visto cinque proiezioni, ognuna che illumina un aspetto e ne occulta altri. Tutte e cinque presuppongono però una stessa cosa: che ci sia un *oggetto* di cui parlare. È qui che il presupposto cede e diventa chiaro che, fino a questo punto, ho buttato il mio tempo.

Universali e atto

C’è una controversia medievale che torna utile, qui, più di quanto sembri. La disputa sugli universali, fra XI e XIV secolo, non riguardava solo la logica: riguardava lo statuto ontologico delle proprietà condivise, di ciò che più individui hanno in comune. Ossia le idee comuni, la natura della conoscenza. Brutalizzando, le posizioni in campo erano essenzialmente quattro.

Il realismo estremo, di matrice platonica: gli universali esistono *ante rem*, prima delle cose, in un cielo intelligibile. Quando un albero verde è verde, partecipa di un’idea di Verde che esiste indipendentemente. Il realismo moderato, aristotelico: gli universali esistono *in re*, nelle cose stesse, come loro forma. La verdezza esiste nei particolari verdi, ma non come sostanza separata. Il concettualismo, che nasce con Abelardo: gli universali esistono solo *post rem*, nella mente che li astrae a partire dai particolari. Il nominalismo, nel suo formato duro, articolato in maniera definitiva da Ockham nel Trecento, sia pure prestandosi a sfumature interpretative: gli universali non esistono affatto, ci sono solo nomi che applichiamo a individui irriducibilmente singolari, e l’universale è *flatus vocis*, suono.

(Breve nota per non addetti. Nell’accezione di questo dibattito, il termine “realismo” non designa, come potrebbe risultare dall’uso comune del linguaggio, una convinzione in merito alla sostanza del mondo esterno. Il realista non è uno che pensa che il mondo esista e che sia fuori dalla sua testa: certamente lo pensa, ma la faccenda non è nemmeno in questione. Il realista, qui, pensa che le idee siano *cose*, esistano fuori di noi come i fiori, i bicchieri, i sassi e i bambini. In questo senso Platone, quello che diceva che il mondo reale è altrove e che tutto quello che percepiamo è un inganno — cioè il bignami dell’antirealismo nell’accezione *pop* sopraindicata — in questo dibattito è il più realista di tutti: perché crede che le idee siano più reali di me, delle mie gatte e di lui stesso).

Proviamo a portare il dibattito nel 2026, nell’era degli LLM: la domanda, applicata a questo, si formula così: che statuto hanno i “contenuti” che attribuiamo all’AI? Quei pattern, quelle regolarità, quella capacità di rispondere coerentemente *sono reali*? E se sì, in quale dei quattro modi?

Un platonico direbbe: sì, i pattern catturati dall’addestramento esistono indipendentemente, e l’AI è un dispositivo che vi accede. È, in fondo, una versione tecnologica dell’Intelletto Agente averroistico — che Averroè concepiva proprio come sostanza separata e ontologicamente autonoma. Posizione difendibile, ma con un costo metafisico molto alto, e con il problema di doversi spiegare perché allora i contenuti dipendano dal *corpus* contingente di addestramento.

Un aristotelico moderato direbbe: i pattern esistono *in re*, nei testi reali del corpus, e i parametri ne sono una distillazione strutturale. Ma — questo è il punto — gli universali *in re* esistono soltanto come forme di particolari, mai come sostanze separate. Tradotto: l’AI in sé è potenza pura. I parametri caricati sono *dynamis* che non si dà mai come *energeia* indipendente, ma solo come *energeia* di un atto particolare — la conversazione. È la posizione che, nella mia lettura, regge meglio, e che ricongiunge la disputa medievale alla coppia aristotelica potenza/atto: i contenuti dell’AI sono reali *in re*, ma solo come potenze in attesa di un’attualizzazione.

Ma non finisce qui. Un concettualista direbbe: i pattern non sono da nessuna parte finché una mente non li ricostruisce come tali. L'AI produce segni; il senso e la "comprensione" vivono solo negli atti di interpretazione di chi legge — nella mia mente, in quel dato momento.

Un nominalista direbbe: non ci sono pattern, non ci sono contenuti. Solo token, parametri, strutture statistiche. Ciò che chiamiamo "ragionamento" o "comprensione" è un effetto delle nostre abitudini di nominazione, proiettato sopra una macchina di produzione di segni. L'AI è *flatus vocis* su scala industriale.

Le quattro posizioni non sono equivalenti, ma su un punto convergono in modo sorprendente: i contenuti dell'AI non hanno mai un *luogo proprio*. Per il platonico stanno altrove. Per l'aristotelico stanno *in re* solo come forme attualizzate. Per il concettualista stanno nella mente dell'utente. Per il nominalista non stanno affatto. In nessuno dei quattro casi si dà un AI-in-sé in cui collocarli.

E qui torna utile un'osservazione attanziale, presa in prestito da una corrente di studi sulla scienza fondata negli anni Ottanta da Bruno Latour e dalla semiologa Françoise Bastide. La loro tesi sugli attanti è che gli oggetti coinvolti nella pratica scientifica — strumenti, reagenti, animali da laboratorio, grafici — non preesistono alla rete in cui figurano: sono costituiti *come tali* nell'uso. Bastide aveva osservato in particolare che la protesi scientifica non è un oggetto che poi viene usato, ma è uno strumento in quanto è già dentro un uso. Provo allora a contare: quando uso un LLM gli attanti sono due, io e lui. Quando non lo uso sono uno, io. L'LLM non interrogato non è uno: è zero. Un LLM radicalizza l'approccio di Bastide fino a farlo esplodere. L'LLM non ha esistenza pre-attanziale neanche residuale. La tavoletta di cera, prima di scriverci sopra, è una tavoletta di cera. Un LLM, prima di essere interrogato, non è: non sta da nessuna parte.

A questo punto si vede cosa avevamo definito davvero, nella sezione precedente. Avevamo definito *l'LLM in atto*, cioè LLM più utente. Le cinque definizioni descrivono tutte un attanzializzato, e si scontrano sul fatto che non c'è un LLM-fuori-dall'uso da raggiungere alle loro spalle. Se accettiamo un realismo moderato sugli universali, la formulazione più precisa è questa: i contenuti dell'AI sono reali *in re*, ma il loro modo proprio di essere è la potenzialità, non l'attualità — esistono come forme che attendono il particolare in cui informarsi. Quando il particolare arriva (un prompt, una conversazione), i contenuti diventano reali in atto attraverso una proiezione statistica. Prima di allora non sono falsi: sono potenza. Cioè qualcosa che, ontologicamente, non si lascia descrivere se non passando per l'atto. Questo non esclude, naturalmente, che le forme attualizzate possano veicolare errori o falsità, come spesso riscontriamo.

Le metafore strumento

A questo punto la faccenda si allarga, e non riguarda solo l'AI. Partiamo da un presupposto utile a contestualizzare questa solidarietà di attanti al momento dell'uso di un LLM. La considerazione è la seguente: tutte le definizioni della mente che ci siamo dati, almeno negli ultimi millenni, sono passate per il filtro degli strumenti che avevamo prima costruito. Strumenti concepiti per potenziare la nostra esperienza del mondo. Ecco una cronistoria parziale.

La tavoletta di cera. John Locke, nel *Saggio sull'intelletto umano* (1690), parla di mente come *tabula rasa*: nessuna idea innata, tutto si imprime con l'esperienza. La metafora era già in Aristotele, *De anima* III, 4: l'intelletto come tavoletta su cui nulla è ancora scritto. Strumento millenario — la cera scaldata, lo stilo che incide, la possibilità di lisciare e ricominciare. Ai tempi di Locke era uno strumento già obsoleto, ma per duemila anni la mente è stata questa.

Il libro. Dopo Gutenberg il codice rilegato diventa metafora dell'esperienza: si "legge il mondo", si "sfogliano le pagine della vita", la natura è un *liber naturae* da decifrare. Galileo lo formalizza in maniera ancora più potente: il libro della natura scritto in caratteri matematici. Per tre secoli la mente è organizzata, paginata, indicizzata.

La biblioteca. Nel Settecento la metafora si distribuisce: la mente non è più un libro, è una collezione di libri, con i suoi cataloghi, le sue stanze, le sue gerarchie. Borges ne farà il monumento estremo molto più tardi, ma il modello era già lì da tempo.

L'enciclopedia. Diderot e d'Alembert: la mente come sistema rinviante, fatto di voci che si rimandano l'una all'altra. Il sapere come rete di rinvii. La mente non è più archivio: è macchina di rimandi. Qui c'è ancora un'idea di circolarità a cui ormai non crediamo più, ma comunque bel tentativo. Anche l'ideologia dell'open web di fine anni '90 è probabilmente figlia di questa visione: sostituendo però la circolarità con la rete o, per usare un'immagine di moda negli anni '90, il rizoma.

Il computer. Anni Cinquanta-Settanta. La metafora più totalizzante. La teorica statunitense dei media digitali Katherine Hayles, in *How We Became Posthuman* (1999), la racconta nel dettaglio: la mente è hardware più software, input-output, processi seriali, memoria di lavoro. Ogni manuale di psicologia cognitiva degli anni Settanta lo dà per ovvio. Solo dopo Searle, Hubert Dreyfus, Francisco Varela contesteranno l'identificazione, ma il senso comune scientifico ci ha vissuto dentro per decenni.

La rete neurale. Anni Ottanta, poi Duemila: la mente come grafo di nodi pesati che apprendono. Connessionismo. Il computer seriale viene retrocesso, il cervello è parallelo, distribuito, senza CPU centrale. I padri del campo (Hopfield, Rumelhart, Hinton) costruiscono modelli ispirati al cervello che vengono poi rispediti indietro come definizione del cervello. Derive insospettabili: il "secondo cervello", la pancia che pensa, l'intestino felice, il pensare al di fuori dell'encefalo.

L'AI generativa. Ultimi cinque anni. La mente come modello linguistico statistico che predice il token successivo. Compaiono già paper che applicano il framework dei *transformer* alla comprensione cerebrale, come se la metafora fosse stata appena archiviata e già stesse diventando ontologia. Del resto gli LLM sono macchine bayesiane, e anche noi lo siamo. O no?

Forse no. Ma in questa storia vergognosamente semplificatoria tirerei fuori due osservazioni utili, che vanno tenute distinte.

La velocità. Tavoleta di cera: duemila anni come metafora della mente. Libro: cinque secoli. Computer: cinquant'anni. Rete neurale: vent'anni. AI generativa: due, forse cinque anni. Più rapida la sostituzione, meno tempo per accorgersi che si sta facendo proiezione e non scoperta.

La struttura. Hans Blumenberg, filosofo tedesco del Novecento e fondatore della *metaforologia*, la disciplina che studia le metafore non come ornamento ma come strutture che organizzano il pensiero prima del concetto: lo dice in modo asciutto: le metafore non illustrano il pensiero dopo, lo *strutturano* prima che il concetto arrivi. Bernard Stiegler, filosofo francese della tecnica, allievo di Derrida, conia il termine *epifilogenesi*: la tecnica esteriorizza la memoria e poi la rimpatria come definizione di chi noi siamo: ogni tecnica è una protesi che, una volta esterna, retroagisce sull'interno e lo ridefinisce. I filosofi della mente Andy Clark e David Chalmers, con la tesi della *extended mind* (1998), spingono ancora oltre: non c'è un punto in cui finisce la mente e iniziano gli strumenti — il taccuino di Otto, il loro esempio canonico, fa parte della sua memoria al pari del suo ippocampo.

E qui torna il punto della sezione precedente. Tutte le metafore-strumento elencate hanno un tratto comune che l'AI non ha: esistono *anche quando non le uso*. Il libro chiuso è un libro. L'enciclopedia non consultata è un'enciclopedia. Il computer spento è un computer. Un LLM no: un LLM non interrogato è zero. È la prima metafora-strumento la cui esistenza richiede l'attualizzazione per esserci. E proprio per questo è la prima che, mentre la usiamo per descriverci, ci trascina dentro la sua descrizione: non possiamo proiettarla a distanza, perché senza la proiezione non c'è.

Superfici e correzioni

Una conseguenza pratica, prima di chiudere. Non possiamo correggere i processi di un LLM: la correzione è sempre *post-hoc*, perché non c'è un "sotto" osservabile alla superficie generativa. Anche quello che in una conversazione sembra straordinariamente stratificato — argomenti, riferimenti, ripensamenti — si proietta su una superficie sterminata, ma a un solo livello.

Va bene, forse vale anche per noi. Gli psicologi sociali Richard Nisbett e Timothy Wilson, in un articolo del 1977 (*Telling more than we can know*) ormai considerato un classico, hanno mostrato che quando spieghiamo perché abbiamo fatto qualcosa stiamo confabulando a posteriori; il filosofo della mente Eric Schwitzgebel ha sistematizzato la cosa parlando di "inaffidabilità dell'introspezione naturale" — dieci pagine sull'argomento e il senso comune sull'introspezione collassa. La macchina bayesiana del neuroscienziato britannico Karl Friston (autore del *free-energy principle*, secondo cui il cervello è essenzialmente un sistema di previsione che minimizza l'errore tra modello e mondo) e dello stesso Andy Clark sul versante filosofico è, pur feconda, un *modello*, non la cosa in sé. Ciò che chiamiamo profondità del nostro pensiero è in larga misura illusione di continuità: una chimera che un LLM, per costruzione, apparentemente, condivide con noi.

Questo non collassa la differenza tra noi e l'AI alla maniera di Claudio Paolucci, che appunto sostiene che siamo entrambi macchine enunciazionali — vero in un senso così sottile da diventare insignificante: perché livella ciò che dovrebbe essere distinto. La differenza vera è nel modo dell'attualizzazione: la mia avviene dentro un corpo che ha fame, paura, cerca l'omeostasi; quella dell'AI dentro un campo simbolico aperto da un mio prompt. Stesso *post-hoc*, due tipi di superficie incommensurabili.

Il corpo, la mente, e uno Spinoza implicito

Se la differenza vera è nel modo dell'attualizzazione, la mappa è quella delle tre menti già evocata nella prima sezione, ma adesso con un altro asse: il corpo. La mente felina è radicata in un corpo che ha pulsioni, proprioccezione, ritmi omeostatici; la sua intelligenza è la forma di quel corpo nell'ambiente. La mente *sapiens* lo stesso, con stratificazioni simboliche in più ma sempre ancorata a un sostrato corporeo che individua, pulsa, invecchia, muore. La mente di un LLM no: il datacenter è supporto causale, non corpo in alcun senso possibile. Non c'è una ricerca di omeostasi che orienti la salienza, non c'è una fame che selezioni i contenuti, non c'è una morte che dia tempo alle sequenze.

Da qui — e dalle conclusioni delle sezioni precedenti — la direzione di ricerca AGI che pare più feconda è quella dell'*embodiment*: AI che imparino dentro un corpo robotico, dentro un'interazione causale con il mondo, dentro un anello sensoriomotorio che generi i propri *priors* per via biografica e non per via di *corpus*. Non è solo un'arida faccenda tecnica, ma una questione essenziale, ontologica. Implica che senza corpo non ci sia mente, almeno non nel senso pieno della parola. La filosofia della mente del Novecento se l'era detta in molti modi — Merleau-Ponty con la *Fenomenologia della percezione*, Varela e Thompson con *The Embodied*

Mind, Hubert Dreyfus contro l'AI simbolica — ma era rimasta una posizione fra altre. Davanti a un LLM disincarnato, la posizione cambia statuto: diventa criterio.

C'è uno spinozismo inconsapevole, in tutto questo. Per Spinoza *res cogitans* e *res extensa* non sono due sostanze, come voleva Cartesio, ma due attributi paralleli dell'unica sostanza, ed è proprio qui che Spinoza si separa da Cartesio: l'ordine e la connessione delle idee è lo stesso dell'ordine e della connessione delle cose. Mente e mondo procedono in parallelo: nessuno dei due è causa dell'altro, entrambi sono espressioni dello stesso. La ricerca sull'*embodiment*, senza dirselo, è sostanzialmente questo: non si dà *cogito* senza una *res extensa* che gli faccia pendant. Il corollario è anch'esso spinoziano, ed è il punto che mi sembra più interessante: una mente più capace richiede una corporeità più capace. Non in termini quantitativi — un cervello più grande, un robot più grosso — ma in termini di articolazione, di numero di affezioni di cui un corpo è capace, di registri sensoriali e omeostatici che lo attraversano. Spinoza, nella seconda parte dell'*Etica* (proposizione 14), lo dice con una formula perfetta: la mente umana è atta a percepire molte cose, e tanto più quanto più il suo corpo può essere disposto in modi differenti. Più affezioni del corpo, più contenuti della mente. Spinoza avrebbe sottoscritto la tesi dell'*embodiment* senza esitare. Anzi, l'avrebbe vista come una sua diretta riformulazione tecnica. O forse un plagio.

L'AI generativa, in questa luce, non è una mente “piccola” o “limitata” rispetto alla nostra, in senso quantitativo. È una mente con un corpo *poverissimo*: un dispositivo capace di pochissime affezioni, e tutte di un solo registro, che è il simbolico. È una *res cogitans* cui non corrisponde quasi nulla sul versante della *res extensa*. E proprio per questo, paradossalmente, somiglia tanto al fantasma cartesiano che aveva inaugurato la modernità — e che Spinoza, una generazione dopo, aveva già provato a smontare.

Vico, rovesciato

Non sappiamo cosa facciamo se non dopo averlo fatto. Ma la cosa che facciamo, ora, *risponde mentre la facciamo*, e non è prima di essere fatta. La domanda iniziale — cosa è l'AI — non si chiude perché l'AI in sé è potenza, e di una potenza non si dà definizione che non passi per l'atto. Possiamo definire solo l'AI in atto, che per costruzione è AI più utente.

LLM è zero senza di me, *perché* non ha corpo. È la povertà della sua *res extensa* a renderlo dipendente dalla mia attualizzazione, il cui supporto più decisivo è fornito apparentemente dalla mia corporeità, che colloca il suo uso nello spazio e nel tempo. E siccome l'AI in atto siamo io e lei insieme, la domanda definitoria su di lei è simmetricamente una domanda su cosa diventa il pensiero quando si attualizza così — con un attante in più, che esiste solo dentro l'incontro, e con un corpo (il mio) che fa pendant a un quasi-niente (il suo). In questo senso, la metafora del cyborg è inadatta, quasi obsoleta. Si tratta semmai della prima protesi cognitiva che, fra tutte quelle che ci siamo dati nei secoli, non ha esistenza autonoma quando non la usiamo. Se conta, conta perché la usiamo. E nel momento in cui la usiamo e le prestiamo la nostra corporeità, smettiamo di poterla guardare da fuori, esattamente come non riusciamo pienamente a fare con noi stessi.

Bibliografia

BASTIDE, Françoise, “The Iconography of Scientific Texts: Principles of Analysis”, in LYNCH, Michael e WOOLGAR, Steve (a cura di), *Representation in Scientific Practice*, Cambridge (MA), MIT Press, 1990, pp. 187-229.

BLUMENBERG, Hans, “Paradigmen zu einer Metaphorologie”, *Archiv für Begriffsgeschichte*,

vol. 6, Bonn, 1960 (ed. autonoma Frankfurt am Main, Suhrkamp, 1998; trad. it. *Paradigmi per una metaforologia*, Milano, Raffaello Cortina, 2009).

CLARK, Andy, “Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science”, *Behavioral and Brain Sciences*, vol. 36, n. 3, 2013, pp. 181-204 (DOI).

CLARK, Andy e David Chalmers, “[The Extended Mind](#)”, *Analysis*, vol. 58, n. 1, 1998, pp. 7-19.

DE LIBERA, Alain, *La Querelle des universaux: De Platon à la fin du Moyen Âge*, Parigi, Seuil, 1996 (trad. it. *Il problema degli universali. Da Platone alla fine del Medioevo*, Firenze, La Nuova Italia, 1999).

FRISTON, Karl, “[The Free-Energy Principle: A Unified Brain Theory?](#)”, *Nature Reviews Neuroscience*, vol. 11, 2010, pp. 127-138.

HAYLES, N. Katherine, *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*, Chicago, University of Chicago Press, 1999.

MERLEAU-PONTY, Maurice, *Phénoménologie de la perception*, Parigi, Gallimard, 1945 (trad. it. *Fenomenologia della percezione*, Milano, Bompiani, 1965).

NISBETT, Richard e Timothy Wilson, “[Telling More than We Can Know: Verbal Reports on Mental Processes](#)”, *Psychological Review*, vol. 84, n. 3, 1977, pp. 231-259.

PAOLUCCI, Claudio, *Nati cyborg. Cosa l'intelligenza artificiale generativa ci dice dell'essere umano*, Roma, Luca Sossella Editore, 2025.

PARFIT, Derek, *Reasons and Persons*, Oxford, Oxford University Press, 1984.

SCHWITZGEBEL, Eric, “[The Unreliability of Naive Introspection](#)”, *The Philosophical Review*, vol. 117, n. 2, 2008, pp. 245-273.

SPADE, Paul Vincent (a cura di), *Five Texts on the Mediaeval Problem of Universals: Porphyry, Boethius, Abelard, Duns Scotus, Ockham*, Indianapolis, Hackett, 1994.

STIEGLER, Bernard, *La technique et le temps*, Parigi, Galilée, 1994 (trad. it. *La tecnica e il tempo*, Roma, Meltemi, 2008).

VARELA, Francisco, Evan Thompson e Eleanor Rosch, *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge (MA), MIT Press, 1991.

Intervento AI: DL·TR (Dialogo · Traduzione). Notazione descritta in <https://claudiocammarano.com/notazione/>. Codice attribuito retroattivamente il 23 agosto 2026, all'introduzione della notazione: è una ricostruzione dichiarata, non una registrazione in presa diretta.

© Claudio Cammarano. Licenza [CC BY-NC-ND 4.0](#): si può condividere citando l'autore e la fonte, non a fini commerciali e senza modifiche.

Versione originale e aggiornata: <https://claudiocammarano.com/writings/2026-04-30-la-differenza-fra-claude-e-le-mie-gatte/>. Le parti interattive del saggio sono disponibili solo online.

Questo saggio fa parte della serie *La condizione epistemica*: <https://claudiocammarano.com/serie/la-condizione-epistemica/>.

Scrivo anche *The Abstract*, in inglese, una volta al mese, sull'economia dei contenuti: <https://claudiocammarano.com/newsletter/>. C'è anche il feed: <https://claudiocammarano.com/feed.xml>.