

Abstract

Il lavoro pedagogico nell'era dell'IA generativa: per una farmacologia critica dell'educazione algoritmica

Pedagogical work in the era of Generative AI: toward a critical pharmacology of algorithmic education

Salvatore Paone

Docente a contratto Università degli Studi di Firenze

salvatore.paone@unifi.it

This contribution examines the impact of generative artificial intelligence on pedagogical work through the concept of *technological pharmakon* (Stiegler). The analysis highlights the operational characteristics of LLMs based on statistical prediction and their implications for educational practice. By integrating the *capabilities approach* (Nussbaum) with the ethics of *craftsmanship* (Sennett), an interpretative framework is proposed that considers both the potential and limitations of cognitive automation. The work suggests guiding principles to preserve the relational dimensions of education, proposing a critical reflection on the competencies required of teachers in the current technological context.

Keywords: work pedagogy, technological pharmakon, Large Language Models, solutionism, capabilities approach, cognitive automation

Il contributo esamina l'impatto dell'intelligenza artificiale generativa sul lavoro pedagogico attraverso il concetto di *pharmakon tecnologico* (Stiegler). L'analisi evidenzia le caratteristiche operative degli LLM basati su predizione statistica e le implicazioni per la pratica educativa. Integrando il *capabilities approach* (Nussbaum) con l'*etica artigianale* (Sennett), si propone un framework interpretativo che considera potenzialità e limiti dell'automazione cognitiva. Il lavoro suggerisce principi orientativi per preservare le dimensioni relazionali dell'educazione, proponendo una riflessione critica sulle competenze richieste ai docenti nell'attuale contesto tecnologico.

Parole chiave: pedagogia del lavoro, pharmakon tecnologico, Large Language Models, soluzionismo, capabilities approach, automazione cognitiva

Citation: Paone S. (2025). Il lavoro pedagogico nell'era dell'IA generativa: per una farmacologia critica dell'educazione algoritmica. *Pampaedia, Bollettino As.Pe.I*, 199(2), 105-120.

Copyright: © 2025 Author(s). | **License:** Attribution 4.0 International (CC BY 4.0).

Conflicts of interest: The Author(s) declare(s) no conflicts of interest.

DOI: <https://doi.org/10.7346/aspei-022025-08>



1. L'intelligenza artificiale generativa come pharmakon pedagogico

Se si confrontano le principali analisi del Novecento sulla tecnica, da autori come Günther Anders (1956) e Arnold Gehlen (1957) a figure cardine come Martin Heidegger (1976) e Herbert Marcuse (1964), si nota come l'approccio di Bernard Stiegler (1994) presenta un elemento di indubbia originalità nell'affrontare quella che Heidegger (1976), figura centrale anche per Stiegler, definiva "la questione della tecnica". Sebbene l'idea di *pharmakon* in Stiegler derivi dalla lettura che Derrida (1978) ha proposto del Fedro di Platone¹, la sua elaborazione si discosta in modo significativo da quella del maestro (Vignola, 2024, pp. 186-187). Per Stiegler, il *pharmakon* coincide con la "ritenzione terziaria"², la memoria esteriorizzata nei supporti tecnici, dalla scrittura alle memorie digitali, la quale, a seconda dell'uso, può costituire tanto un rimedio quanto una minaccia. Da questa identificazione, il concetto diventa il perno attorno a cui ripensare la politica, l'economia, la filosofia e, in senso più ampio, le forme di produzione del sapere (Vignola, 2015, pp. 35-56). Ogni tecnologia è un *pharmakon* nel senso che comporta simultaneamente potenziale terapeutico e tossico: ogni innovazione tecnica è simultaneamente rimedio e veleno, cura e intossicazione.

L'intelligenza artificiale generativa rappresenta il *pharmakon* contemporaneo per eccellenza, particolarmente nel contesto del lavoro cognitivo e pedagogico. I *Large Language Models* (LLM), fondati su architetture *Transformer* e meccanismi di *self-attention*, hanno il potenziale di operare come dispositivi tecno-semiottici in grado di ristrutturare l'orizzonte ontologico del lavoro intellettuale (Neuhaus, 2023). Tuttavia, l'evidenza di questa trasformazione cognitiva rimane ancora limitata e largamente teorica; di conseguenza, le loro capacità si configurano spesso come una semplice "illusione di comprensione" (Marchetti *et alii*, 2025), dovuta alla mancanza di meccanismi cognitivi genuini. Essi non si limitano a supportare le pratiche cognitive preesistenti, bensì ne modificano la configurazione, spostando l'attività dall'elaborazione diretta alla mediazione, orchestrazione e valutazione di contenuti generati. In quanto ritenzione terziaria di nuovo tipo, gli LLM manifestano la natura farmacologica in forma paradigmatica: non sono sem-

1 In Derrida, invece, il *pharmakon* è impiegato come strumento critico volto a mettere in crisi la metafisica della presenza e il logocentrismo.

2 Nella filosofia di Bernard Stiegler, la "ritenzione terziaria" si riferisce all'esteriorizzazione della memoria e della conoscenza attraverso artefatti tecnologici. È un concetto cruciale per comprendere come la tecnologia plasma il nostro pensiero e le nostre strutture sociali, in particolare nell'era digitale. Stiegler sostiene che la ritenzione terziaria, a differenza della ritenzione primaria (biologica) e secondaria (sociale), è intrinsecamente legata agli oggetti tecnologici e alla loro capacità di immagazzinare e trasmettere informazioni, plasmando la nostra comprensione del passato, del presente e del futuro.



plici archivi di conoscenza consultabili, ma sistemi che attivamente generano e ricombinano sapere secondo logiche probabilistiche; questa caratteristica rende la loro ambivalenza più radicale rispetto alle tecnologie precedenti.

La loro capacità di processare, sintetizzare e generare testo a velocità e scale sovraumane si manifesta in tensioni costitutive: libera il lavoratore cognitivo dalle routine ripetitive proprio mentre minaccia di renderlo obsoleto; democratizza l'accesso al sapere mentre solleva complesse questioni di governance legate al controllo, alla privacy e all'accountability che richiedono framework politici specifici (Chan, 2023); aumenta la produttività pedagogica attraverso ruoli complementari tra IA e docenti nella gestione della classe e personalizzazione didattica (Jeon, Lee, 2023), mentre rischia di erodere la relazione educativa autentica, richiedendo l'implementazione di salvaguardie per prevenirne l'abuso (Kamalov, Santandreu Calonge, Gurrib, 2023).

Questa ambivalenza farmacologica si manifesta con particolare intensità nel lavoro del docente, che rappresenta un caso paradigmatico di lavoro cognitivo ad alta complessità relazionale. Tale complessità emerge dalle molteplici sfide nell'implementazione dell'IA nella pratica didattica (Celik *et alii*, 2022). L'insegnante si trova infatti nella paradossale posizione di dover sviluppare competenze digitali adeguate per preparare studenti per un mondo permeato dall'IA utilizzando strumenti di IA (Ng *et alii*, 2023) che potrebbero potenzialmente ridefinire le competenze tradizionali che cerca di trasmettere. Di fronte a tale paradosso, la possibilità di sviluppare un approccio critico verso l'intelligenza artificiale è fondamentale per l'integrazione educativa ed è strettamente legata alla necessità di una comprensione approfondita delle tecnologie IA (Walter, 2024); tuttavia, questa stessa comprensione richiede proprio quelle capacità metacognitive e di riflessione profonda che l'automazione rischia progressivamente di atrofizzare.

La crescente introduzione dell'intelligenza artificiale generativa in ambito pedagogico pone una questione epistemologica cruciale, che eccede la sfera puramente tecnologica. La cognizione umana, come teorizzato dall'approccio *4E cognition - embodied, embedded, enacted, extended* (Newen, Gallagher, & De Bruin, 2018; Menary, 2010), emerge dall'integrazione inscindibile tra corporeità, azione situata e interazione sociale, configurando il pensiero come processo radicato in dinamiche sensomotorie e relazionali. Al contrario, i sistemi di intelligenza artificiale di tipo connessionista elaborano informazioni attraverso la manipolazione di stati computazionali intrinsecamente inaccessibili, con la loro generazione automatica che risulta opaca in modo strutturalmente unico e non esplicativo (Boge, 2021). Questa opacità, nota come problema della *black box*, esprime la difficoltà strutturale di ricostruire e comprendere i processi decisionali che conducono l'algoritmo a una determinata conclusione (Von Eschenbach, 2021). Emerge così un paradosso costitutivo del pharmakon algoritmico: la complessità computazionale che rende opachi i meccanismi decisionali è precisamente ciò che abilita capacità generative di portata inedita.



È necessario distinguere tra sistemi reattivi e sistemi agentici, benché entrambi operino secondo logiche computazionali che divergono radicalmente dall'esperienza fenomenologica umana. I modelli linguistici di grandi dimensioni basati su architetture transformer, quali ChatGPT, Claude o Gemini, processano input linguistici attraverso meccanismi di attenzione che identificano pattern statistici in vasti corpora testuali, trasformando regolarità probabilistiche in performance comunicative che simulano comprensione senza possederla. Parallelamente, i sistemi agentici rappresentano un'evoluzione verso agenti autonomi orientati agli obiettivi (Murugesan, 2025): già operativi seppur in fase embrionale, introducono capacità di azione autonoma nel mondo digitale e fisico, sollevando questioni ancora più complesse riguardo a sicurezza, efficacia, privacy ed etica. La delega di operazioni a entità algoritmiche autonome richiede framework di governance che ancora non possediamo pienamente (Bengio *et alii*, 2023).

Il rischio centrale che questi sistemi pongono all'educazione è quello che Morozov (2013) identifica come "soluzionismo tecnologico": la convinzione che ogni problema sociale, politico o educativo sia essenzialmente un problema tecnico risolvibile attraverso l'applicazione di algoritmi, dati e automazione. Nel contesto educativo, il soluzionismo tecnologico si manifesta non tanto in promesse mirabolistiche quanto nella tendenza più sottile a presentare l'IA come risposta naturale e necessaria a sfide pedagogiche complesse: dalla personalizzazione dell'apprendimento alla valutazione delle competenze, dal carico di lavoro docente all'inclusione educativa. Queste aspettative sono spesso basate su fraintendimenti delle possibilità tecniche attuali e su visioni eccessivamente ristrette delle funzioni dell'educazione nella società (Holmes, Tuomi, 2022), trasformando questioni pedagogiche multidimensionali in problemi di ottimizzazione computazionale.

Il lavoro pedagogico, inteso come pratica di formazione umana integrale, si rivela simultaneamente attratto dalle possibilità di ottimizzazione offerte dall'IA e irriducibile a esse nella sua essenza trasformativa. L'educazione autentica richiede, come argomenta Nussbaum (2011a), la coltivazione di capacità intrinsecamente relazionali: il pensiero critico, l'immaginazione narrativa e la capacità di riconoscere l'umanità dell'altro. Queste dimensioni, fondamentali per la formazione democratica, nascono dall'incontro pedagogico tra soggetti e resistono a qualsiasi riduzione algoritmica (Chan, Tsi, 2024).

La sfida che si delinea non consiste nel rifiutare aprioristicamente l'innovazione tecnologica né nell'abbracciarla acriticamente, ma nel comprendere quali dimensioni del processo educativo possano beneficiare del supporto algoritmico e quali invece costituiscano il nucleo inalienabile dell'incontro formativo, preservando quella dimensione etica, estetica e affettiva che definisce l'autenticità dell'esperienza educativa.



2. Dialettica del *pharmakon*: liberazione e alienazione nel lavoro educativo

La natura farmacologica dell'IA richiede lo sviluppo di quello che Stiegler (2013) definisce "farmacologia positiva": pratiche e istituzioni capaci di orientare la tecnologia verso effetti trasformativi minimizzando quelli tossici. Nel contesto pedagogico, questa ambivalenza si manifesta attraverso una dialettica complessa tra liberazione e alienazione, potenziamento e proletarizzazione, inclusione e atrofizzazione che investe il cuore stesso del lavoro educativo.

Il "veleno" dell'IA nel lavoro pedagogico opera attraverso un doppio movimento di *datafication* e automazione. Il rischio non risiede tanto nella sostituzione meccanica del docente quanto nella progressiva erosione della complessità dell'insegnamento come pratica sociale, emotiva ed etica irriducibile alla trasmissione di informazioni (Selwyn, 2019). Questa riduzione dell'atto pedagogico a funzioni computabili costituisce la precondizione per quella che Stiegler (2015) chiama "proletarizzazione cognitiva": l'esteriorizzazione di saperi e competenze in sistemi automatici che il docente utilizza senza comprenderne la logica operativa.

La ricerca empirica conferma questa ambivalenza farmacologica dell'IA educativa. Ghamrawi et al. (2023) documentano come l'IA possa simultaneamente espandere e far regredire la leadership docente mediante personalizzazione e automazione. Parallelamente, Jeon et al. (2023) identificano una relazione complementare tra LLM e docenti, sottolineando come l'efficacia educativa dipenda dall'*expertise* pedagogica umana, insostituibile nell'orchestrazione didattica.

Un esempio paradigmatico dell'ambivalenza si manifesta nella questione della personalizzazione didattica. L'IA viene spesso proposta come strumento per adattare l'insegnamento agli "stili di apprendimento" individuali. Tuttavia, nonostante la popolarità di teorie come quella di Kolb (1984) o il modello VARK (Fleming & Mills, 1992), la ricerca empirica non fornisce evidenze solide a supporto di questi costrutti (An & Carr, 2017; Willingham, Hughes, Dobolyi, 2015). Pelletier (2023) critica come l'IA venga erroneamente equiparata alla "personalised learning", costruendo un discorso retorico che presenta soluzioni tecnologiche come automaticamente benefiche mentre trascura questioni educative storicamente centrali che richiederebbero risposte sfumate, ponderate ed empiricamente fondate. Quando gli LLM promettono di adattarsi agli "stili" individuali, perpetuano questa reificazione algoritmica, trasformando miti pedagogici scientificamente infondati in soluzioni computazionali apparentemente neutre.

Diversa e più promettente è la questione della multimodalità, che non presuppone tipologie fisse di apprendenti ma riconosce la necessità di multiple vie di accesso al contenuto secondo i principi dell'*Universal Design for Learning* (CAST, 2024). La multimodalità abilita i sistemi a processare e generare informazioni attraverso canali sensoriali differenti, permettendo traduzione fluida tra modalità secondo necessità contestuali. I modelli multimodali possono soddisfare efficacemente le specifiche esigenze cognitive, comportamentali ed emotive degli



studenti con bisogni speciali (Farhah *et alii*, 2025): questa capacità non serve a confermare presunte tipologie cognitive statiche, bensì a garantire accessibilità universale. Uno studente con dislessia può beneficiare di una fruizione audio non perché sia un “apprendente uditivo” per natura, ma perché in quel momento quella modalità rimuove una barriera all’apprendimento.

La dimensione “curativa” dell’IA si manifesta quando la tecnologia amplifica anziché sostituire le capacità pedagogiche umane. Kim (2023) documenta come la collaborazione insegnante-IA miri a ottimizzare punti di forza complementari preservando i ruoli didattici degli insegnanti, mentre l’automazione di compiti routinari libera tempo prezioso per interazioni formative più profonde. I docenti esperti utilizzano sistemi generativi non come surrogati del pensiero pedagogico ma come interlocutori dialogici per esplorare possibilità didattiche alternative, evolvendo verso quella che Kim definisce una partnership costruttiva. Questa potenzialità terapeutica si realizza solo quando il docente mantiene la capacità di interrogare criticamente gli output algoritmici, comprendendone i limiti e orientandone l’uso secondo principi educativi esplicativi.

L’ambiente educativo contemporaneo è già saturo di mediazioni digitali: *Learning Management Systems*, piattaforme didattiche e strumenti di *analytics* hanno progressivamente trasformato ogni aspetto dell’apprendimento in dato computabile: tempi di risposta, pattern di errore, frequenze di accesso, percorsi di navigazione. Thompson e Prinsloo (2022) utilizzano il concetto di “*data gaze*” per criticare questa “dataficazione intensificata” che rende visibili e quantificabili dimensioni dell’apprendimento, creando simultaneamente opportunità di analisi e rischi di riduzionismo metrico. In questo ecosistema già denso di mediazioni algoritmiche, l’introduzione dell’IA generativa rappresenta un salto qualitativo: mentre i sistemi precedenti si limitavano a tracciare, organizzare e analizzare dati esistenti, gli LLM generano attivamente nuovi contenuti attraverso ricombinazione probabilistica, presentando sfide distintive che divergono dai modelli precedenti (Huang *et alii*, 2023). Questo introduce un livello di complessità e imprevedibilità che può essere tanto fonte di innovazione quanto di allucinazioni, bias e standardizzazione occulta.

La sfida centrale non è tecnica ma epistemologica: richiede una “saggezza farmacologica”, una forma di *phronesis* pedagogica capace di navigare tra le potenzialità trasformative e i rischi alienanti dell’automazione. La *phronesis* aristotelica, come evidenzia Kristjánsson (2024), si configura come integratore e arbitro psicоморale per l’eccellenza nelle decisioni etiche professionali. Questa saggezza non si limita alla competenza operativa ma esige la capacità meta-riflessiva di mantenere una tensione produttiva tra innovazione tecnologica e valori pedagogici fondamentali. La questione non è se utilizzare l’IA, ma come preservare, nella rinegoziazione tecnologica del rapporto educativo, quello spazio di riconoscimento reciproco attraverso cui docente e studente si costituiscono come soggetti autonomi nel processo formativo. Il *pharmakon* tecnologico dell’IA richiede



quindi non un rifiuto luddista né un'accettazione acritica, ma una pratica consapevole che sappia navigare tra cura e veleno, trasformazione e alienazione, mantenendo al centro la dimensione irriducibilmente umana dell'educazione.

3. Oltre apocalittici e integrati: l'inaffidabilità costitutiva degli LLM

Umberto Eco (1964) nel suo celebre “Apocalittici e integrati” identificava due atteggiamenti polarizzati di fronte alla cultura di massa che oggi si ripresentano, *mutatis mutandis*, di fronte all’IA. Gli apocalittici contemporanei vedono nell’automazione cognitiva la fine dell’autenticità umana; gli integrati celebrano ogni innovazione come democratizzazione del sapere e liberazione cognitiva.

Il dibattito contemporaneo sugli LLM, tuttavia, soffre di una patologia discorsiva che va oltre la semplice polarizzazione. Il campo semantico è saturo di distorsioni concettuali: strategie commerciali mascherate da riflessione teorica, analogie cognitive che antropomorfizzano processi statistici, filosofie improvvise che trasformano vincoli computazionali in opportunità pedagogiche. Alzahrani (2024) documenta come voci critiche sollecitino approcci più misurati, evidenziando preoccupazioni sistematiche riguardo a trasparenza, bias e questioni etiche occultate da narrazioni ottimistiche. Questo inquinamento epistemico cristallizza le comunità di pratica in *echo chambers* autoreferenziali, impermeabili all’evidenza empirica che contraddice le loro narrazioni preferite.

Un esempio paradigmatico di questa confusione è la tendenza a interpretare l’inaffidabilità strutturale degli LLM come opportunità formativa. Quando si sostiene che gli errori dell’IA stimolino il pensiero critico, si compie un’operazione retorica che occulta una realtà tecnica fondamentale: questi sistemi operano attraverso la predizione statistica del token successivo più probabile, senza alcuna rappresentazione semantica del significato. A differenza dei sistemi basati su *knowledge base*³, che organizzano esplicitamente fatti e relazioni semantiche in modo ispezionabile, gli LLM non dispongono di strutture interne di conoscenza: non producono sapere, bensì generano simulacri testuali di coerenza statistica. Tali narrazioni promozionali riflettono la fallace assunzione che le tecnologie digitali costituiscano sempre la soluzione ai problemi educativi (Williamson, 2023). La propensione all’“errore” degli LLM, tecnicamente definita “allucinazione”, non è un difetto correggibile ma una conseguenza diretta dell’architettura probabilistica su cui si fondano. I processi interni rimangono non trasparenti e non ispezionabili per ragioni intrinseche al design architettonicale (Carabantes, 2019).

3 Le knowledge base tradizionali possono assumere la forma di basi di regole (nell’intelligenza artificiale simbolica classica), di ontologie (in logica descrittiva).



Questa opacità strutturale si traduce in una trasformazione qualitativa del lavoro pedagogico: il docente che integra gli LLM nella propria pratica non può ispezionare i processi inferenziali che generano un determinato output, ma deve esercitare una vigilanza epistemica costante attraverso la validazione esterna dei contenuti. A differenza delle *knowledge base*, dove le catene inferenziali sono tracciabili e ispezionabili, l'architettura probabilistica degli LLM trasforma ogni interazione in un atto di fiducia non verificabile dall'interno del sistema. L'integrazione dell'IA generativa nella didattica, perciò, non può esaurirsi in una competenza tecnico-operativa: richiede capacità metacognitive di verifica, discernimento e distanza riflessiva, in grado di spingersi oltre la fascinazione della plausibilità linguistica verso un giudizio sulla verità e sul senso del discorso generato. Solo in questa prospettiva la saggezza farmacologica può diventare strumento di orientamento nell'ecosistema algoritmico. Zhai, Wibowo, Li (2024) documentano empiricamente come l'eccessiva dipendenza dai sistemi di IA comprometta le capacità cognitive degli studenti quando accettano raccomandazioni generate dall'IA senza questionarle.

Questa incomprensione genera anche un paradosso educativo: mentre si moltiplicano le proposte di utilizzare gli LLM come “partner cognitivi” o strumenti di collaborazione docente–IA verso partnership costruttive (Kim, 2023), si ignora che stiamo delegando funzioni epistemiche essenziali a sistemi che operano in assenza di comprensione. Il risultato è un regime di conoscenza in cui la verosimiglianza linguistico-statistica sostituisce progressivamente la verifica fattuale, la fluidità testuale maschera l'assenza di referenzialità e la competenza apparente occulta l'incompetenza strutturale.

Nel contesto pedagogico, questa confusione si traduce in una falsa dicotomia tra tecnofobia e tecnofilia acritica. I primi vedono nell'IA solo una minaccia alla relazione educativa autentica; i secondi la celebrano come soluzione universale ai problemi dell'istruzione. Entrambe le posizioni, come osservava Eco in *Apocalittici e integrati* (1964), rappresentano forme di “pigrizia intellettuale” che eludono un'analisi rigorosa delle specificità tecniche e delle loro implicazioni pedagogiche.

Martha Nussbaum (2011b), attraverso il *capabilities approach*, offre un criterio valutativo che supera la sterile opposizione tra tecnofilia e tecnofobia. Le capacità fondamentali che identifica: dalla ragion pratica all'affiliazione, dall'immaginazione al controllo del proprio ambiente, diventano parametri per stabilire quando l'integrazione tecnologica espande l'*agency* e la dignità umana e quando invece le compromette. Applicato agli LLM in educazione, questo *framework* rimette in luce un'ambivalenza irriducibile: lo stesso sistema può, al tempo stesso, ampliare le possibilità esplorative del docente e minarne l'autorità epistemica, facilitare la personalizzazione didattica e insieme omogeneizzare i contenuti secondo i pattern dominanti nei dati di addestramento.

La questione cruciale diventa quindi: come preservare e potenziare le capacità



umane fondamentali in un ecosistema educativo sempre più mediato da sistemi che operano secondo logiche opache e non deterministiche? La risposta non risiede né nel rifiuto categorico né nell'adozione entusiastica, ma richiede quella che, nello spirito della *phronesis* aristotelica, potremmo definire una “saggezza pratica situata”: la capacità di discernere, contesto per contesto, quando la mediazione algoritmica arricchisce l'esperienza formativa e quando invece la impoverisce.

Richard Sennett (2008) approfondisce questa prospettiva attraverso il concetto di *craftsmanship*, l'eccellenza artigianale che deriva dalla padronanza riflessiva dei propri strumenti. L'artigiano autentico non è solo utilizzatore competente ma comprende intimamente materiali, tecniche e limiti del proprio mestiere. Traslato nel contesto dell'IA educativa, questo significa che il docente deve sviluppare non soltanto competenze operative ma una comprensione critica dell'architettura computazionale che governa questi sistemi. Il problema è che l'opacità algoritmica degli LLM commerciali impedisce quella relazione dialogica con lo strumento che Sennett identifica come essenziale all'etica artigianale. Il docente si trova a operare con *black box* in cui sono visibili input e output, ma restano invisibili i processi intermedi: sistemi che promettono “intelligenza” ma operano attraverso correlazioni statistiche cieche al significato. Questa asimmetria epistemica mina alla base la possibilità di quella maestria consapevole che caratterizza il lavoro artigianale.

La via critica che si propone riconosce quindi la necessità di superare tanto la postura apocalittica quanto l'integrazione acritica, attraverso una comprensione rigorosa dei vincoli architettonici e computazionali di questi sistemi. Non si tratta di trovare un compromesso tra estremi, ma di sviluppare criteri valutativi fondati sulla chiarezza tecnica e pedagogica. Solo riconoscendo gli LLM come potenti manipolatori di pattern linguistici privi di comprensione semantica possiamo utilizzarli produttivamente come amplificatori di possibilità testuali da vagliare criticamente, mai come sostituti del giudizio pedagogico; come generatori di variazioni da selezionare, mai come autorità epistemiche; come strumenti per esplorare lo spazio delle formulazioni possibili, mai come garanti di verità.

Questa consapevolezza critica non implica rifiuto ma uso sapiente, non tecnofobia ma lucidità tecnica, non apocalisse ma vigilanza epistemica. Il lavoro pedagogico nell'era degli LLM richiede quindi una nuova forma di professionalità: quella del docente-epistemologo, capace di navigare tra simulacri testuali e conoscenza autentica, tra plausibilità statistica e verità pedagogica, mantenendo salda la bussola dei valori educativi in un mare di rumore computazionale. Tale visione del docente come figura critica trova riscontro empirico nello studio di Mouta, Torrecilla-Sánchez, Pinto-Llorente (2024), i cui partecipanti identificano la necessità per gli insegnanti di diventare “*vital agents of thought and action in the face of the polysemic landscape of AIED technologies*” (p. 3369), confermando che l'educatore non può limitarsi al ruolo di *gatekeeper of high-quality content*, ma deve sviluppare quella dimensione epistemologico-critica qui teorizzata.



4. Saggezza farmacologica: principi operativi e dimensioni irriducibili

Le diagnosi precedenti hanno evidenziato l'inquinamento epistemico del dibattito e l'inaffidabilità costitutiva degli LLM. È quindi indispensabile elaborare una farmacologia positiva per il lavoro pedagogico che, senza ignorare i vincoli strutturali di questi sistemi, sviluppi pratiche consapevoli per mitigarne gli effetti tossici. I principi che seguono sono euristiche operative per navigare un ecosistema in cui la plausibilità statistica viene sistematicamente confusa con la verità epistemica.

Il primo principio, la “trasparenza epistemica”, risponde alla tendenza a mascherare processi probabilistici come comprensione semantica. L'insegnante deve mantenere consapevolezza costante che gli LLM sono ricombinatori statistici privi di referenzialità. Comprendere che questi sistemi generano output attraverso predizione del *token* successivo più probabile, non attraverso comprensione, è precondizione per evitare quella delega epistemica acritica che trasforma il docente in operatore passivo di *black box*. Tale consapevolezza si traduce in competenze operative concrete: distinguere tra un sistema di *retrieval-augmented generation* (RAG) con accesso a fonti web aggiornate e uno perimetrato a documenti controllati⁴; riconoscere il significato di *fine-tuning*⁵ e di *prompting*; individuare i segnali di un'allucinazione rispetto a un'informazione verificata. Come confermano Kalai *et alii* (2025), le allucinazioni non sono anomalie misteriose ma errori di classificazione binaria emergenti da pressioni statistiche intrinseche al training. Il docente deve padroneggiare il *prompting* non come stratagemma procedurale, ma come strumento di dialogo epistemico con sistemi che operano secondo logiche computazionali non isomorfe al ragionamento umano.

Il secondo principio, la “contestabilità”, costituisce l'antidoto alla sostituzione dell'*agency* pedagogica. L'allucinazione non è un errore marginale ma la cifra stessa di sistemi che operano per approssimazione statistica: per questo ogni output deve rimanere aperto, rivedibile, mai assunto come definitivo. La contestabilità non è solo un meccanismo tecnico, ma una forma di vigilanza epistemica che custodisce lo spazio del giudizio umano. È in questo spazio che l'insegnante, con la sua presenza incarnata, può esercitare la libertà di decostruire ciò che l'algoritmo propone, riconoscere le assunzioni implicite: culturali, epistemologiche, valoriali e correggere quelle inevitabili distorsioni che nessuna macchina può redimere. Contestare significa quindi difendersi dalla mera plausibilità linguistica e custodire lo spazio autentico del dialogo educativo.

4 Un esempio di sistema perimetrato è NotebookLM di Google, che opera esclusivamente sui documenti caricati dall'utente senza accesso a fonti esterne, garantendo maggiore controllo sulla fonte delle informazioni ma limitando la capacità generativa al corpus fornito.

5 Il fine-tuning è l'adattamento di un modello pre-addestrato a compiti specifici attraverso training aggiuntivo su dataset mirati.



Il terzo principio, la “complementarità”, risponde alla tentazione di trasformare l’inaffidabilità strutturale in virtù pedagogica attraverso narrazioni salvifiche. Gli LLM sono potenti generatori di variazioni testuali, non produttori di conoscenza: vanno utilizzati come amplificatori di possibilità da vagliare criticamente, mai come sostituti del discernimento umano. La distinzione è cruciale: generazione esplorativa *versus* decisione valutativa, moltiplicazione delle opzioni *versus* selezione pedagogica. Un sistema è complementare quando espande lo spazio delle formulazioni possibili senza sostituire il giudizio pedagogico; diventa invece tossico quando i suoi output statisticamente plausibili ma epistemicamente inaffidabili vengono acriticamente accettati.

Uno scenario operativo può mostrare come questi principi traducano la saggezza farmacologica in pratica didattica concreta. Un docente che si serve di un LLM per progettare un’unità didattica su un tema complesso, storico, letterario o scientifico, non lo interroga come un’encyclopedia affidabile, ma ne delimita consapevolmente lo spazio d’azione fornendo un corpus selezionato di fonti verificate. Attraverso un sistema di *retrieval-augmented generation* perimetrato, il docente circoscrive il materiale su cui l’LLM può operare, trasformandolo da oracolo opaco a strumento di ricombinazione supervisionata. Il modello può così generare molteplici formulazioni didattiche, proporre connessioni inedite tra i documenti, suggerire percorsi narrativi alternativi, ma sempre entro il perimetro epistemico tracciato dall’intenzionalità educativa. L’output non rappresenta mai un punto di arrivo, bensì un materiale grezzo da vagliare e rifinire secondo criteri pedagogici trasparenti.

In questa pratica si attuano simultaneamente i tre principi: la trasparenza epistemica è assicurata dal controllo delle fonti; la contestabilità è preservata dalla tracciabilità degli input; la complementarità riafferma il ruolo insostituibile del giudizio pedagogico nella selezione, nell’interpretazione e nella valutazione degli output generati.

L’articolazione concreta di questi principi suggerisce una farmacologia positiva possibile, non una soluzione definitiva: una tensione produttiva tra consapevolezza critica e uso strumentale che deve ancora misurarsi con la complessità del reale educativo. Tuttavia, tali principi devono radicarsi nel riconoscimento di ciò che resta irriducibile all’automazione: presenza incarnata, creatività pedagogica situata e riconoscimento etico.

La presenza incarnata acquista significato di fronte ai simulacri algoritmici: l’insegnante costituisce una presenza corporea che abita lo spazio pedagogico, testimoniando l’irriducibilità dell’esperienza umana alla manipolazione simbolica, quella dimensione che Merleau-Ponty (2003) pone a fondamento dell’interazione con il mondo.

La creatività pedagogica situata rappresenta l’antitesi della generazione probabilistica: mentre gli LLM operano attraverso *pattern matching* statistico, l’inse-



gnante risponde creativamente all'unicità di ogni situazione educativa, realizzando quella *reflection-in-action* teorizzata da Schön (1983).

Il riconoscimento etico, infine, è radicalmente incomputabile: si rivolge all'alterità irriducibile dell'altro, quella dimensione che per Levinas (1961) precede ogni categorizzazione.

L'integrazione tra rigore operativo (trasparenza epistemica, contestabilità, complementarità) e consapevolezza della dimensione antropologica dell'educare (presenza incarnata, creatività situata, riconoscimento etico) dovrebbe orientare la formazione degli insegnanti, sviluppando quella capacità critica che sappia distinguere tra potenziamento e alienazione epistemica, guidandoli a progettare pratiche didattiche che mantengano senso, empatia e consapevolezza riflessiva. Non basta aggiungere moduli su "IA e didattica" ai curricula esistenti: occorre ripensare radicalmente cosa significhi preparare docenti per un ecosistema saturo di generatori di plausibilità. La formazione deve demistificare il funzionamento degli LLM senza cadere nel tecnicismo estremo, fornendo quella comprensione concettuale che immunizza dalle retoriche salvifiche. Parallelamente, deve confrontarsi con le implicazioni etiche della delega epistemica a sistemi strutturalmente inaffidabili, sviluppando la capacità di riconoscere quando l'efficienza del simulacro compromette l'integrità del processo educativo. Infine, sul piano relazionale, occorre preservare l'autenticità dell'incontro pedagogico quando la mediazione algoritmica minaccia di sostituire la reciprocità costitutiva del rapporto educativo.

Questa formazione multidimensionale non è lusso intellettuale ma necessità epistemica. Nell'era dei simulacri algoritmici, la capacità di distinguere tra amplificazione strumentale e sostituzione ontologica diventa condizione di possibilità per un'educazione che rimanga fedele alla sua vocazione emancipatrice.

In questa prospettiva la saggezza farmacologica si radica in un'unità inscindibile che integra la padronanza tecnica dello strumento secondo la lezione di Sennett sul *craftsmanship*, che si declina nei principi operativi di trasparenza epistemica, contestabilità e complementarità; il discernimento etico-critico tra rimedio e veleno secondo Stiegler, quella *phronesis* pedagogica capace di navigare la natura farmacologica della tecnologia; e la consapevolezza antropologica delle dimensioni irriducibili all'automazione: la presenza incarnata che Merleau-Ponty pone a fondamento dell'esperienza, l'atto creativo che Schön identifica con la riflessione-in-azione e il riconoscimento etico dell'alterità nella prospettiva di Levinas e Nussbaum.

Solo una formazione che intrecci queste dimensioni può aiutare il docente ad abitare consapevolmente il *pharmakon* tecnologico, cercando di mantenere salda quella dimensione di incontro e riconoscimento che nessun algoritmo può davvero replicare.



5. Conclusioni: navigare l'ecosistema algoritmico

Dalle riflessioni emerse si delinea un punto di approdo: il *pharmakon* algoritmico non esige un compromesso tra posizioni opposte, quanto una lucidità radicale capace di riconoscere i vincoli strutturali e smascherare le derive discorsive che comprimono la complessità in slogan. L'IA non “risolverà” i problemi dell'educazione, né necessariamente li aggraverà: introduce piuttosto una complessità epistemica che impone nuove forme di vigilanza critica.

Il lavoro pedagogico nell'era degli LLM richiede un rigore operativo particolare: la consapevolezza costante di interagire con generatori di plausibilità testuale, unita alla determinazione di trasformare questa consapevolezza in pratica educativa riflessiva. Ciò implica sviluppare una “doppia vista”: utilizzare questi strumenti riconoscendone simultaneamente potenzialità e illusioni, e imparando a navigare tra possibilità e vincoli. Significa rifiutare le narrazioni consolatorie che tendono a trasformare limiti strutturali in presunte opportunità pedagogiche o a confondere correlazione statistica con comprensione autentica.

In questo senso, la saggezza farmacologica si configura come una forma di resistenza epistemica: né rifiuto né accettazione acritica della tecnologia, ma capacità di mantenere chiarezza analitica in un ecosistema saturo di rumore computazionale. L'insegnante è chiamato a sviluppare una nuova professionalità, quella del docente-epistemologo, capace di distinguere tra conoscenza e verosimiglianza, tra comprensione e associazioni probabilistiche; e a coltivare una competenza critica che sappia riconoscere quando il simulacro può operare come strumento esplo- rativo e quando, invece, rischia di compromettere l'integrità epistemica del pro-cesso educativo.

Ma il docente-epistemologo deve operare in un ecosistema tecnologico segnato da profonde asimmetrie di potere. Mentre oligopoli tecnologici controllano ar- chitetture e dataset, le istituzioni educative navigano tra dipendenza tecnologica e illusione di autonomia. La volatilità dell'ecosistema, con soluzioni che nascono e scompaiono rapidamente, impone scelte strategiche difficili.

L'imperativo non è adattare gli esseri umani alle macchine né viceversa, ma mantenere consapevolezza della loro radicale alterità. Gli LLM restano strumenti che ampliano possibilità testuali senza produrre conoscenza, simulacri ottimizzati per metriche commerciali che riflettono inevitabilmente i bias dei loro creatori.

Come ricordava Eco, il problema non è stabilire se la tecnologia sia "buona" o "cattiva", ma capire come operare criticamente al suo interno. Nel caso degli LLM, questo significa navigare consapevolmente tra plausibilità e verità, tra pro- messe di democratizzazione e concentrazione oligopolistica. Solo una vigilanza critica può trasformare il *pharmakon* algoritmico da fonte di inquinamento co- gnitivo in occasione di apprendimento riflessivo. La saggezza farmacologica non è una scelta, ma una postura epistemica necessaria a custodire lo spazio del pen- siero critico nell'era dei simulacri algoritmici



Bibliografia

- Alzahrani A. (2024). Unveiling the shadows: beyond the hype of AI in education. *Helion*, 10(9): e30696. <https://doi.org/10.1016/j.heliyon.2024.e30696>
- Anders G. (1956). *Die Antiquiertheit des Menschen 1: Über die Seele im Zeitalter der zweiten industriellen Revolution*. C. H. Beck.
- An D., Carr M. (2017). Learning styles theory fails to explain learning and achievement: recommendations for alternative approaches. *Personality and Individual Differences*, 116: 410-416. <https://doi.org/10.1016/j.paid.2017.04.050>
- Bengio Y. et alii (2024). Managing extreme AI risks amid rapid progress: preparation requires technical research and development, as well as adaptive, proactive governance. *Science*, 384(6698): 842-845. <https://doi.org/10.1126/science.adn0117>
- Boge F. (2021). Two dimensions of opacity and the deep learning predicament. *Minds and Machines*, 32(1): 43-75. <https://doi.org/10.1007/s11023-021-09569-4>
- Carabantes M. (2019). Black-box artificial intelligence: an epistemological and critical analysis. *AI & Society*, 35(2): 309-317. <https://doi.org/10.1007/s00146-019-00888-w>
- CAST (2024). *Universal Design for Learning Guidelines (Version 3.0)*. Center for Applied Special Technology. <http://udlguidelines.cast.org>
- Celik I. et alii (2022). The promises and challenges of artificial intelligence for teachers: a systematic review of research. *TechTrends*, 66(4): 616-630. <https://doi.org/10.1007/s11528-022-00715-y>
- Chan C. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(1): 1-17. <https://doi.org/10.1186/s41239-023-00408-3>
- Chan C., Tsi L. (2024). Will generative AI replace teachers in higher education? A study of teacher and student perceptions. *Studies in Educational Evaluation*, 81: 101395. <https://doi.org/10.1016/j.stueduc.2024.101395>
- Derrida J. (1972). La pharmacie de Platon. In *La dissémination* (pp. 69-197). Éditions du Seuil (trad. it. J. Derrida, *La farmacia di Platone*, Jaca Book, 1978).
- Eco U. (1964). *Apocalittici e integrati: Comunicazioni di massa e teorie della cultura di massa*. Bompiani.
- Farhah N. et alii (2025). AI-driven innovation using multimodal and personalized adaptive education for students with special needs. *IEEE Access*, 13: 69790-69820. <https://doi.org/10.1109/ACCESS.2025.3560863>
- Fleming N.D., Mills C. (1992). Not another inventory, rather a catalyst for reflection. *To Improve the Academy*, 11(1): 137-155. <https://doi.org/10.1002/j.2334-4822.1992.tb00213.x>
- Gehlen A. (1957). *Die Seele im technischen Zeitalter: Sozialpsychologische Probleme in der industriellen Gesellschaft*. Rowohlt.
- Ghamrawi N., Shal T., Ghamrawi N. (2023). Exploring the impact of AI on teacher leadership: regressing or expanding? *Education and Information Technologies*, 29: 8415-8433. <https://doi.org/10.1007/s10639-023-12174-w>
- Heidegger M. (1953). *Vorträge und Aufsätze*. Neske (trad. it. M. Heidegger, *Saggi e discorsi*, Mursia, 1976).



- Holmes W., Tuomi I. (2022). State of the art and practice in AI in education. *European Journal of Education*, 57(4): 542-570. <https://doi.org/10.1111/ejed.12533>
- Huang L. et alii (2023). A survey on hallucination in large language models: principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems*, 43(4): 1-55. <https://doi.org/10.1145/3703155>
- Jeon J., Lee S. (2023). Large language models in education: the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 28: 15873-15892. <https://doi.org/10.1007/s10639-023-11834-1>
- Kalai A.T. et alii (2025). Why language models hallucinate. *arXiv preprint arXiv:2509.04664*. <https://doi.org/10.48550/arXiv.2509.04664>
- Kamalov F., Calonge D., Gurrib I. (2023). New era of artificial intelligence in education: towards a sustainable multifaceted revolution. *Sustainability*, 15(16): 12451. <https://doi.org/10.3390/su151612451>
- Kim J. (2023). Leading teachers' perspective on teacher-AI collaboration in education. *Education and Information Technologies*, 29: 8693-8724. <https://doi.org/10.1007-s10639-023-12109-5>
- Kolb D.A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice-Hall.
- Kristjánsson K. (2024). Aristotelian practical wisdom (phronesis) as the key to professional ethics in teaching. *Topoi*, 43(4): 1031-1042. <https://doi.org/10.1007/s11245-023-09974-7>
- Levinas E. (1961). *Totalité et infini*. Martinus Nijhoff.
- Marchetti A. et alii (2025). Artificial intelligence and the illusion of understanding: a systematic review of theory of mind and large language models. *Cyberpsychology, Behavior, and Social Networking*. <https://doi.org/10.1089/cyber.2024.0536>
- Marcuse H. (1964). *One-dimensional man: Studies in the ideology of advanced industrial society*. Beacon Press.
- Menary R. (2010). Introduction to the special issue on 4E cognition. *Phenomenology and the Cognitive Sciences*, 9(4): 459-463. <https://doi.org/10.1007/S11097-010-9187-6>
- Merleau-Ponty M. (1945). *Phénoménologie de la perception*. Gallimard (trad. it. M. Merleau-Ponty, *Fenomenologia della percezione*, Il Saggiatore, 2003).
- Morozov E. (2013). *To save everything, click here: The folly of technological solutionism*. PublicAffairs.
- Mouta A., Torrecilla-Sánchez E., Pinto-Llorente A. (2024). Comprehensive professional learning for teacher agency in addressing ethical challenges of AIED: insights from educational design research. *Education and Information Technologies*, 30(7): 3343-3387. <https://doi.org/10.1007/s10639-024-12946-y>
- Murugesan S. (2025). The rise of agentic AI: implications, concerns, and the path forward. *IEEE Intelligent Systems*, 40(2): 8-14. <https://doi.org/10.1109/MIS.2025.-3544940>
- Neuhaus F. (2023). Ontologies in the era of large language models – a perspective. *Applied Ontology*, 18(4): 399-407. <https://doi.org/10.3233/AO-230072>
- Newen A., Gallagher S., De Bruin L. (2018). 4E cognition. In *The Oxford handbook of 4E cognition*. Oxford University Press. <https://doi.org/10.1093/oxfordhb-9780198735410.013.1>



- Ng D.T.K. *et alii* (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. *Educational Technology Research and Development*, 71(1): 137-161. <https://doi.org/10.1007/s11423-023-10203-6>
- Nussbaum M.C. (2011a). *Non per profitto: Perché le democrazie hanno bisogno della cultura umanistica*. Il Mulino.
- Nussbaum M.C. (2011b). *Creating capabilities: The human development approach*. Harvard University Press.
- Pelletier C. (2023). Against personalised learning. *International Journal of Artificial Intelligence in Education*, 33(1): 1-5. <https://doi.org/10.1007/s40593-023-00348-z>
- Schön D.A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.
- Selwyn N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.
- Selwyn N. (2022). The future of AI and education: some cautionary notes. *European Journal of Education*, 57(4): 620-631. <https://doi.org/10.1111/ejed.12532>
- Sennett R. (2008). *The craftsman*. Yale University Press.
- Stiegler B. (1994). *La technique et le temps, 1: La faute d'Épiméthée*. Galilée.
- Stiegler B. (2013). *What makes life worth living: On pharmacology*. Polity Press.
- Stiegler B. (2015). *La società automatique: Vol. 1. L'avenir du travail*. Fayard.
- Thompson T., Prinsloo P. (2022). Returning the data gaze in higher education. *Learning, Media and Technology*, 48(2): 153-165. <https://doi.org/10.1080/17439884.2022.2092130>
- Vaswani A. *et alii* (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (vol. 30).
- Vignola P. (2015). Il pharmakon di Stiegler. Dall'archi-cinema alla società automatica. In V. Cuomo (ed.), *Medium: Dispositivi, ambienti e psicotecnologie* (pp. 35-56). Kaiak.
- Vignola P. (2024). Dall'Epimeteo dimenticato alla politica di Ermes. Stiegler e la postproduzione filosofica del mito. *Trópos. Rivista di ermeneutica e critica filosofica*, 16(1): 185-204. <https://doi.org/10.13135/2036-542X/11060>
- Von Eschenbach W.J. (2021). Transparency and the black box problem: why we do not trust AI. *Philosophy & Technology*, 34(4): 1607-1622. <https://doi.org/10.1007/s13347-021-00477-0>
- Walter Y. (2024). Embracing the future of artificial intelligence in the classroom: the relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education*, 21(1): 1-29. <https://doi.org/10.1186/s41239-024-00448-3>
- Willingham D.T., Hughes E.M., Dobolyi D.G. (2015). The scientific status of learning styles theories. *Teaching of Psychology*, 42(3): 266-271. <https://doi.org/10.1177/0098628315589505>
- Williamson B. (2023). The social life of AI in education. *International Journal of Artificial Intelligence in Education*, 34(1): 97-104. <https://doi.org/10.1007/s40593-023-00342-5>
- Zhai C., Wibowo S., Li L. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments*, 11(1): 28. <https://doi.org/10.1186/s40561-024-00316-7>

