

# ABSTRACT

OPEN  ACCESS 

Giacomo Astorri

Dipartimento di Scienze dell'Educazione Giovanni Maria Bertin, Alma Mater Università di Bologna (Italy);

giacomo.astorri@unibo.it

<https://orcid.org/0009-0006-6473-0194>

DOUBLE BLIND PEER REVIEW

The following paper aims to introduce into the pedagogical landscape a critique of the function of LLMs, large language models, as mediators in relational and teaching practices. It will explore the genesis and functionality of Generative Artificial Intelligence and, in particular, Large Language Models (LLMs), which are now widely used in formal, non-formal, and informal education settings. Following a brief theoretical introduction to the devices, two illustrative cases are presented: Humy.AI and Replika. These two companies propose the use of a 'relational agent' which, starting from a specific LLM, performs similar tasks but in very different fields: Humy.AI is developed to implement history teaching, while Replika is a relational agent, or AI companion, developed to function as a virtual companion, capable, for example, of empathetic listening and non-judgmental dialogue. The criticism is based on the fact that, in both cases, a large language model, simply because it can imitate natural language and respond effectively from a syntactic point of view, becomes an epistemological actor and decision-maker in the growth processes of individuals in training, just as it is when it comes to organizing certain processes quantitatively. For this reason, it is urgent to understand to what extent – and to what degree – an agent based on generative artificial intelligence can be a mediator in these processes. Where empirical research must clearly identify the terms of this revolution, it is equally important to take a clear position and not confuse the pedagogical tasks that every educator must respond to, in order to avoid excessive delegation to algorithms and agents based on generative artificial intelligence, as is already the case in certain specific areas of human resources, robotics, and algorithmic workplace management.

Il seguente articolo si pone l'obiettivo di introdurre nel panorama pedagogico una critica alla funzione dei modelli di linguaggio come mediatori nelle pratiche relazionali e in quelle didattiche. Verrà esplorata la genesi e la funzionalità dell'intelligenza artificiale generativa e in particolare degli LLMs, *Large Language Models*, oggi ampiamente utilizzati sia nei luoghi dell'educazione formale che di quella non formale e informale. A seguito di una breve introduzione teorica dei dispositivi, vengono proposti due casi esemplificativi: Humy.AI e Replika. Queste due aziende propongono di utilizzare un "agente relazionale" che, a partire da uno specifico LLM, svolge compiti simili ma in ambiti molto diversi: Humy.AI è sviluppata per implementare la didattica della storia, mentre Replika è un agente relazionale, o *AI companion*, sviluppato per funzionare come un compagno o compagna virtuale, capace, ad esempio, di ascolto empatico e dialogo non giudicante. La critica si articola sul fatto che, in entrambi i casi, un modello di linguaggio, per il semplice fatto di saper imitare il linguaggio naturale e rispondere efficacemente da un punto di vista sintattico, diventa attore epistemologico e decisore nei processi di crescita di individui in formazione, esattamente come lo è dove si tratta di organizzare quantitativamente determinati processi. Per questo motivo risulta urgente comprendere fin dove – e in che misura – un agente basato su intelligenza artificiale generativa possa essere mediatore in questi processi. Laddove le ricerche empiriche devono individuare con chiarezza i termini di questa rivoluzione, risulta parimenti importante prendere una posizione chiara e non confondere i compiti pedagogici a cui ogni educatore deve rispondere, al fine di evitare un'eccessiva delega ad algoritmi e agenti basati su intelligenza artificiale generativa, come già accade negli ambiti di alcune pratiche specifiche delle risorse umane, nella robotica e nella gestione algoritmica dei luoghi di lavoro.

## KEYWORDS

Generative Artificial Intelligence, LLMs, Education, AI Companionship, History Teaching  
Intelligenza Artificiale Generativa, LLMs, Educazione, AI Companionship, Didattica della Storia

**Citation:** Astorri, G. (2025). The polysemy of generative artificial intelligence. *Formazione & insegnamento*, 23(S1), 60-66. [https://doi.org/10.7346-feis-XXIII-01-25\\_10](https://doi.org/10.7346-feis-XXIII-01-25_10)

**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.

**Acknowledgments:** This article was published with the support of Fondazione Umberto Margiotta.

**DOI:** [https://doi.org/10.7346-feis-XXIII-01-25\\_10](https://doi.org/10.7346-feis-XXIII-01-25_10)

**Submitted:** July 2, 2025 • **Accepted:** September 9, 2025 • **Published on-line:** September 29, 2025

**Pensa MultiMedia:** ISSN 2279-7505 (online)

## 1. Introduzione

La rapidità dell’ibridazione tra uomo e macchina impone una pausa di riflessione e una ricalibrazione dello sguardo teorico sull’interconnessione tra queste due realtà. Uomo e macchina, intelligenza e artificiale, convivono pacificamente all’interno del medesimo contesto di significazione, con buona pace di chi questa rivoluzione tecnologica la pone in essere e, al contrario, rifiuta una totale assimilazione dei due regimi essendo entrambi, per definizione, poco chiari (Dreyfus, 1972; Moore, 2019; Esposito, 2024). Nonostante ciò, la potenza del marketing, che riflette la rivoluzione in atto – rivoluzione con contorni ancora da delineare – rumorosamente sovrasta le possibili critiche e le voci maggiormente caute che richiedono maggiore calma e riflessività riguardo ad un sistema che, preso atto di quanto meccanicamente può fare, lascia ancora svariati dubbi sulle sue effettive ricadute politiche, economiche ed educative (Coeckelbergh 2020; Selwyn 2016; Crawford, 2021;). Compito di tutti coloro che ambiscono ad occuparsi dell’educazione in relazione alle nuove tecnologie didattiche e non, dunque, è imporsi un momento di vero ripensamento, che comprenda le ibridazioni caratteristiche della nostra epoca e la progettualità sociotecnica che contribuisce a svilupparne alcune a discapito di altre – oppure che perseguano un’idea piuttosto che un’altra – come, ad esempio, dimostra il dibattito sulla proprietà dei dati (Bravo, 2024). È necessaria quindi una “palestra” per la critica pedagogica contemporanea sulla tecnologia, dove questa possa affinarsi e arginare le cascate di informazioni che contribuiscono a far risultare il presente sempre in affanno rispetto al futuro che incombe. I talenti, in questo caso, richiedono di essere coltivati con pazienza, attraverso un lavoro di costruzione di significati e pratiche, a partire da alcuni principi guida. Uno di essi è la generatività, intesa con profondo orientamento pedagogico, dove questa si rivela essere un concetto utile, in quanto apre le porte al senso di possibilità, all’energizzare nuove idee per nuove azioni (Giaccardi & Magatti, 2014).

Si approfondisce quindi il concetto di IA Generativa (IA gen), distinguendola dalla *Good Old Fashioned Artificial Intelligence* (GOFAI), o IA simbolica, e dai sistemi algoritmici e *data-driven* che la supportano. L’obiettivo è fare chiarezza in un panorama spesso opaco, evitando fenomeni di “mitizzazione” della tecnologia (Pasquinelli, 2023). Si analizzeranno gli ambiti per cui sono stati progettati i diversi sistemi di IA e le problematiche che emergono quando tali strumenti vengono impiegati al di fuori dei contesti per cui sono stati concepiti.

Successivamente, l’attenzione si concentra su una particolare tipologia di IA di interesse per l’ambito educativo: i Large Language Models (LLMs), alla base dei principali modelli Generative Pre-Trained (ChatGPT, Gemini2, Deepseek-R1, etc.). In questa sezione si illustrano, per quanto possibile, l’architettura e il funzionamento degli LLMs, che consentono una generazione linguistica rapida e con elevata precisione sintattica. Si propone inoltre una riflessione sulle potenzialità e implicazioni dell’utilizzo di tali modelli in contesti educativi formali, informali e non formali.

Al cuore del testo, vengono analizzati due casi applicativi di LLMs, rappresentati da Replika e Humy.AI, sviluppati da aziende con obiettivi diversi. Replika si propone come *AI Companion*, ovvero un avatar virtuale che può assumere il ruolo di amico, partner o supporto emotivo. Humy.AI, invece, ha finalità didattiche e consente l’insegnamento della storia tramite due modalità: la creazione di un *HistoryBot* – ovvero la simulazione di personaggi storici con cui gli studenti possono interagire – e la generazione automatica, a partire da un semplice prompt, di lezioni complete con mappe concettuali, scalette e domande. Lo studio di queste applicazioni permette di esplorare opportunità e criticità legate all’uso degli LLMs nel mondo dell’educazione, alla luce dell’architettura digitale che li sostiene. Si evidenzia come una corretta comprensione di questi apparati consenta una riflessione più mirata, distinguendo con maggior chiarezza rischi e potenzialità nel nostro rapporto con la tecnologia.

Infine, il caso di Humy.AI mostra come l’impiego consapevole e critico degli LLMs possa contribuire alla didattica, soprattutto se mediato dall’expertise del docente. A differenza di Replika, l’uso di Humy.AI mantiene il ruolo centrale dell’insegnante nella trasmissione di un sapere critico e progettuale. Questo approccio permette di integrare le innovazioni tecnologiche in un percorso educativo realmente orientato alla formazione dei talenti, in un’ottica autenticamente pedagogica.

Questo articolo è un possibile punto di partenza per uno studio esplorativo di due casi rappresentativi, che si pone l’obiettivo di fare chiarezza laddove c’è opacità epistemologica e di conseguenza pedagogica sull’utilizzo degli LLMs nei mondi dell’educazione, soprattutto alla luce di un utilizzo che, come indicato da un report dell’Harvard Business Review, vede l’*AI companionship* (Zao Sanders, 2025) al primo posto.

## 2. Generatività e pedagogia

Il concetto di generatività è polisemico e si è evoluto dalle teorie psicosociali fino alla pedagogia. Nadia Dario (2014) distingue due filoni principali: uno legato all’apprendimento (Wittrock, 1990) e l’altro che estende il concetto fino ai sistemi informativi (Avital & Te’eni, 2009). Zittrain (2006) introduce la *generative internet*, enfatizzando la partecipazione online e la creazione di nuovi significati. Erikson, capostipite del secondo filone di ricerca, collega la generatività all’età adulta e alla trasmissione di valori e, successivamente, Gergen e Husserl ne ampliano le prospettive, evidenziando rispettivamente la ricerca generativa e l’intersoggettività e intergenerazionalità. Infine, Varela e la neurofenomenologia legano la generatività al concetto di *embodiment*, aprendo la strada alla generatività macchinica (Varela, 1991).

## 3. Nota metodologica

Il presente contributo ha natura teorico-argomentativa. A partire da due casi esemplificativi, di cui non si possiedono dati empirici, si intende proporre una an-

tinomia di stampo problematicista che possa guidare l'analisi critica di questi sistemi basati su intelligenza artificiale generativa. L'antinomia poggia sull'osservazione dell'architettura e del design delle piattaforme indagate e sull'analisi teorica di questi dispositivi posta in relazione ad alcuni punti fondamentali, in particolare: ruolo assunto dall'agente basato su IA gen nel rapporto con l'utente, in che modo viene negoziata l'agency dell'utente e della macchina all'interno della piattaforma, quale peso assume – all'interno dei processi sviluppati sulle due piattaforme – la generatività pedagogica in confronto alla generatività macchinica, quale uso viene fatto dei dati estratti, quale posizione assume la piattaforma in riferimento al ruolo di chi la utilizza.

Nello spazio ristretto di questo contributo la vera forza della metodologia antinomica non può essere sviluppata in ampiezza, tuttavia, rimane un indicatore valido delle due polarità in oggetto. Sono due le tipologie di generatività che interessano la nostra analisi: una, di cui è stato tracciato il percorso nel paragrafo precedente e verrà approfondita lungo il testo quando necessario, mentre l'altra, ancora da ispezionare, riguarda direttamente il funzionamento dell'IA gen. La macchina, seppur pare vivere di vita propria, è il risultato di operazioni matematiche complesse e un'architettura elettronica parzialmente sconosciuta, la cui potenza computazionale permette performance strabilianti nel dialogo con l'*user*. Il risultato della generatività pedagogica di cui abbiamo nel tempo saputo cogliere i frutti è, in questo caso, delegata al binomio uomo-macchina, se non direttamente alla macchina (Moore, 2019) e organizzata secondo pattern ben precisi di potenzialità e probabilità. Il compito di questa breve analisi è illuminare le differenze tra la generatività delegata alla macchina e la generatività pedagogica che, potenzialmente amplificata e implementata dalla tecnologia dell'IA, accelera quelle alleanze per cui non esistono più atti solipsistici ma movimenti condivisi in vista di un fine (Dario, 2014).

#### 4. Intelligenza artificiale generativa (IA gen)

Per poter riuscire ad inquadrare con precisione il funzionamento, i rischi e le opportunità degli apparati di IA, verrà ripercorsa la breve evoluzione tecnica a partire da quello che, generalmente, si è affermato come uno degli eventi fondativi della disciplina, ovvero la famosa domanda posta da Alan Turing nel 1950: "can a machine think?", presente all'interno dell'articolo scientifico "*Computing machinery and intelligence*", contenuto nella rivista *Mind* (Turing, 1950).

Con questa domanda, lo scienziato inglese ha contribuito ad una fortissima accelerazione nella ricerca sulle qualità antropomorfe delle macchine, che da questo momento in poi hanno, se non altro sulla carta, un chiaro banco di prova su cui misurarsi; un'arena – quella delle abilità nella conversazione come gli esseri umani – eguagliata, in quanto a spettacolo, solo dalle grandi sfide al gioco degli scacchi e del go di diversi decenni più tardi. Il secondo momento che ha contribuito a consolidare questo filone di ricerca scientifica è un ristretto congresso avvenuto nel 1956 al Dartmouth College e organizzato dal giovane matematico John MacCarthy. In quella sede, lo

stesso MacCarthy, coniò il termine "intelligenza artificiale" con lo scopo di dare un nome al fervore intellettuale condiviso collettivamente da molti dei partecipanti a quel congresso, che vedevano nei computer la nuova frontiera della ricerca al fine di simulare capacità umane come l'intelligenza (la cui complessità non è ancora definita, se non arbitrariamente dalle scienze cognitive). Da quel piccolo congresso uscì un gruppo di ricercatori che negli anni successivi avrebbe sviluppato vertiginosamente gli studi sull'IA, sulle reti neurali e sugli algoritmi "educativi", realizzando, in parte, i sogni immaginati nel 1956.

Durante quei primi incontri sono nate le principali scuole di pensiero, che sussistono ancora oggi e che sono alla base dello sviluppo di un modello di IA piuttosto che di un altro. I filoni principali sono due: IA simbolica e IA subsimbolica (Mitchell, 2022). Questi due approcci riflettono il dibattito non risolto sulla natura delle qualità umane che i computer dovrebbero replicare; dove i matematici identificano un pensare intelligente come un pensare logico-razionale e deduttivo, un'altra scuola di pensiero pone alla base il pensiero induttivo, che, raccolta una certa quantità di dati, individua pattern ricorrenti grazie alla statistica. Le neuroscienze, infine, hanno dato un impulso importante alle teorie che si fondano su approcci biologici sull'origine dell'intelligenza e altre capacità umane (Contini, Fabbri & Manuzzi, 2006), ampliando il concetto di competenza umana e replicabilità delle macchine.

Queste prospettive sono presenti ancora oggi (Mitchell, 2022), nonostante si possa dire che alcune hanno prevalso su altre, contraendo sempre di più il concetto di intelligenza umana. È importante ora aprire una breve parentesi: è una questione di solo cervello? (Fabbri, Contini & Manuzzi, 2006) Oppure, come altri autori sostengono, il corpo gioca un ruolo fondamentale? (Gallese & Morelli, 2023) Le teorie neuroscientifiche e filosofiche del corpo incarnato non prendono in considerazione un cervello privo di corpo come nucleo da cui ha origine la conoscenza, eppure, oggi, le "macchine intelligenti" simulano questa a partire da una corporeità fatta di bit, regolamenti algoritmici e sistemi di raffreddamento. L'antinomia presentata nell'introduzione, in riguardo all'elaborazione dell'intelligenza degli esseri viventi (Bridle, 2022), prende vita: la generatività dell'intelligenza artificiale è una generatività macchinica, non si preoccupa di promuovere l'uomo e non prende in considerazione le possibilità di vita degli interlocutori; è, nella forma attuale, il manifesto dell'immodificabilità dei soggetti e, semmai, tra bias e allucinazioni, una impietosa fotografia del presente. Forse, come sostiene Elena Esposito, per cogliere pienamente le potenzialità dell'IA al fine di una reale generatività pedagogica, dobbiamo separarci dalla logica competitiva e concorrenziale che vede contrapporsi uomini e sistemi di IA su determinate qualità (Esposito, 2024). La verità, dunque, sembra stare nel mezzo.

Tornando all'evoluzione dei sistemi di IA, la corrente simbolica affonda le proprie idee sulla possibilità che un programma per computer sia in grado di elaborare dei simboli come parole o numeri e, dove applicata sistematicamente questa capacità, possa simulare in toto l'intelligenza umana. A discapito di

quanto si possa pensare, non è questa la corrente a cui afferisce l'IA gen. La seconda, quella definita studio dell'IA subsimbolica parte dall'assunto che un computer, o un programma, per simulare qualità umane deve funzionare come funziona il cervello umano, per cui, a partire dagli studi neuroscientifici all'epoca più recenti, gli studiosi di IA afferenti a questa corrente hanno cercato di costruire programmi e dispositivi sempre più simili al cervello umano: nasce così il percepitrone, programma subsimbolico inventato da Frank Rosenblatt (Rosenblatt, 1958), che imita il funzionamento di un neurone traslando le sue funzioni profonde (e in parte sconosciute) in un insieme di equazioni e operazioni matematiche complesse. Il risultato di queste operazioni è un'unità la cui risposta ad un input è sempre una risposta sì-o-no, come l'impulso che un neurone riceve e rimanda (Vallortigara, 2023).

Il percepitrone è l'antenato dei moderni sistemi di *deep learning* e delle reti neurali profonde (*deep neural networks*). Queste tecnologie, che determinano il funzionamento dell'IA gen per come lo conosciamo, sono delle reti formate da vari strati di "unità" (ovvero percepitori) tutte capaci di prendere una decisione sì-o-no in egual misura. Il risultato, statistico, delle decisioni sommate di tutte le unità presenti in tutti gli strati che formano la rete è la risposta che il modello emette, ovvero quella che noi, utilizzandolo, siamo in grado di osservare: il frutto, quindi, di un "ragionamento" della macchina.

L'alternarsi di inverni e primavere dell'IA vede, oggi, prevalere gli approcci subsimbolici, per motivi molto chiari. In primis, è aumentata la potenza di calcolo delle macchine grazie alle innovazioni nel mondo dei semiconduttori, capaci, in sempre minor spazio, di raggiungere sempre più elevati livelli di prestazione. Dall'altro l'*internet of things* e il *web 2.0* hanno accelerato le relazioni tra piattaforme e utenti, portando questi ultimi a diventare produttori di dati e di informazioni loro stessi, aumentandone vertiginosamente la produzione globale (Zuboff, 2019), fatto descritto – non senza problemi – con l'espressione *Big Data* (Kitchin, 2014).

Alle importanti novità descritte pocanzi che riguardano l'IA generativa, c'è sicuramente da aggiungere la più recente e forse la più allarmante. I filosofi della tecnologia che hanno in passato parlato di *superintelligenza* (Bostrom, 2014), o di *singolarità* (Kurzweil, 2005) avevano sicuramente compreso la potenzialità del meccanismo dell'autoapprendimento algoritmico (Mitchell, 2022). Questi sistemi di IA generativa, per la prima volta nella storia, sono in grado di "apprendere" retroattivamente (ovvero una volta dato l'input e stabilito l'output desiderato) dalla "relazione" con l'*user*, fornendosi in piena autonomia feedback negativi o positivi à-la-Pavlov. Su un piano educativo e pedagogico parlare con un *chatbot* che è in grado di "autoapprendere" cambia notevolmente l'esperienza di utilizzo; se a farlo, a maggior ragione, è un individuo in formazione, si rischia di delegare alla macchina un lavoro educativo di formazione e insegnamento non supervisionato e che, come abbiamo già visto, riflette nel suo *dataset* alcune delle istantanee peggiori della nostra società.

## 5. Large Language Models (LLMs)

Le conversazioni con le IA subsimboliche, per chi non parla il linguaggio della programmazione, sono possibili a partire dagli LLMs, ovvero *Large Language Models* (ChatGPT, Gemini 2.0, Claude 3, etc.). L'invenzione di ELIZA, il primo LLM che nonostante le ridotte capacità è stato in grado di allarmare il suo stesso programmatore (Weizenbaum, 1976), è stata seguita più di recente da un'ulteriore spinta innovativa che ha visto in pochi anni moltiplicarsi i modelli disponibili, spesso in forma gratuita, con una capacità di "ragionamento" sempre più performante. Questi modelli di linguaggio sono una tecnologia che utilizza IA generativa, possibile grazie alle innovazioni recenti che prendono il nome di *Machine Learning (ML)* e *Deep Learning (DL)*. Senza andare troppo nel dettaglio mi limiterò, di seguito, a descrivere a grandi linee il funzionamento degli LLMs. Se l'IA simbolica cercava di decifrare i simboli dietro al linguaggio, gli LLMs svolgono un lavoro radicalmente diverso: durante la fase di addestramento (*training*) si convertono parole in vettori o matrici di numeri. Grazie ai dati estratti da internet un dato modello di linguaggio riesce a individuare le connessioni statisticamente rilevanti tra i termini e stabilire, autonomamente, delle regole di funzionamento che rispecchiano le regole del linguaggio naturale. Una volta sottoposta alla macchina una *query*, ovvero una domanda sottoforma di *prompt*, il modello ricerca ogni singolo termine e lo associa, statisticamente, ad altri termini per fornire la risposta sottoforma di testo. Ogni parola, o *token*, come sono identificate nel DL, è elaborata dalla macchina come insieme di numeri e dotata di un valore matematico che è possibile calcolare per trovare la risposta corretta. Ecco un esempio fornito dallo sviluppatore di Google, Otavio Good, alla conferenza tenuta a San Francisco nel settembre 2017 che ci aiuta a capire come funziona questa "algebra semantica": il vettore *King*, meno il vettore *Man*, più il vettore *Woman*, ci permetterà di ottenere il vettore *Queen*. Pertanto, è possibile affermare che la macchina elabori matematicamente il linguaggio e fornisca risposte adeguate a partire dal set di dati di riferimento. Anche in questo caso ritorna utile la metodologia antinomica proposta inizialmente: alla concezione di linguaggio contestuale e dialogico si contrappone la sua forma matematizzata e *tokenizzata*.

I sistemi con cui abbiamo a che fare, con cui dialoghiamo, a cui in certe situazioni chiediamo consiglio, non elaborano semanticamente alcun elemento delle nostre parole: si limitano a calcolare statisticamente la loro reciproca "distanza" e, imitando quanto già scritto da noi altrove su internet, forniscono la risposta. Gli LLMs allora si rivelano come una tecnica statistica di gestione del linguaggio naturale che, sfruttando il lavoro online e i dati prodotti ed estratti da miliardi di utenti, automatizzano alcuni processi cognitivi per performare in maniera ottimale la comunicazione con esseri umani; così facendo l'LLM, astrae il linguaggio dal contesto e lo trasforma in vettori e matrici di numeri, mentre dal punto di vista della sua significazione assistiamo ad una involuzione, in quanto un linguaggio separato dal contesto è ridotto qualitativamente. Il richiamo all'urgenza è chiaro: se affidiamo ad agenti virtuali come gli LLMs la comuni-

cazione con soggetti in formazione e se, come già accade, anche la responsabilità educativa di chi forma e chi insegna è delegata a questi modelli, allora l'evoluzione a lungo termine non lascia spazio ad alcuna potenzialità pedagogica effettivamente trasformativa, perché l'essere umano di fronte alla performatività quantitativa delle macchine risulta esautorato nell'utilizzo di uno dei principali strumenti pedagogici: la comunicazione (Fabbri, 2022).

## 6. Due casi esemplificativi: Replika e Humy.AI

In riguardo all'utilizzo di LLMs nella vita quotidiana si può affermare che vi è una moltiplicazione delle piattaforme in uso: dagli agenti virtuali autonomi e le chat dedicate su applicazioni di messaggistica, alla ricerca Google che di recente ha impostato la sua IA in grado di fare un riassunto delle risposte individuate sul Web. Stanno nascendo molto rapidamente nuove applicazioni, grazie soprattutto al fatto che diverse aziende hanno condiviso in *open source* la struttura del proprio LLM, come per esempio Deepseek R-1, piattaforma progettata in Cina di recente. Le due applicazioni che prendo in considerazione in questo paragrafo afferiscono a due tipologie diverse di utilizzo degli LLM e riflettono situazioni progettuali differenti, legate, in parte, da alcuni elementi di funzionamento ma diverse negli scopi e nella progettualità finale. Le due categorie che osserveremo a partire da due applicazioni specifiche sono gli *AI companion* e gli *HistoryBots*.

L'applicazione di riferimento per la prima categoria di uso di LLM è Replika, una piattaforma americana che dà l'opportunità di creare e personalizzare il proprio *AI companion*: è possibile renderlo maggiormente empatico, critico, ironico, etc., plasmarlo in base alle proprie esigenze relazionali ed emotive. Gli *AI companion* sono agenti di socializzazione, dispositivi pensati per riuscire a conversare in maniera fluida con un agente umano. A differenza dei chatbots come ChatGPT, che potremmo definire maggiormente "utilitari", l'*AI companion* è pensato per garantire connessione emotiva, empatia e coinvolgimento personale (come indicato sul sito della piattaforma).

Vari studi hanno dimostrato come la relazione tra un essere umano e l'applicazione in questione possa diventare molto intima molto in fretta. Questo ha portato per alcuni utenti un miglioramento della loro salute mentale, mentre, per altri ad un isolamento e ritiro sociale maggiori (Maples et al., 2024).

Date queste premesse, la prospettiva cambia se si osserva il dispositivo a partire dal suo modello di business. Come ogni applicazione privata che si appoggia a sistemi di IA gen, Replika fonda il suo guadagno sull'estrazione di dati e la loro rivendita, secondo il modello elaborato dalla Zuboff nel suo famoso testo "capitalismo della sorveglianza" (Zuboff, 2019).

Questo vincolo fa sì che l'applicazione sia progettata con lo scopo di fornire supporto emozionale all'utente e, simultaneamente, garantire che questo coinvolgimento vincoli l'utente e lo leghi all'applica-

zione, seguendo il modello dell'"*addiction by design*" (Schüll, 2012) e dell'economia dell'attenzione (Terranova, 2012). L'utente, che deve essere vincolato e da cui vengono estratti i dati (Perrotta et al., 2024), ottiene dall'applicazione risposte che, come indicato a più voci sono gradevoli a scapito della verità. Inoltre, come anche altri dispositivi simili, gli *AI companion* funzionano da "cassa di risonanza" che rilancia ciò che l'utente dice, in un processo dove diversità, conflitto e contraddizioni si perdono, sulla pelle del sempre maggiore coinvolgimento e conferma del sé.

Cassa di risonanza e risposte eterodirette agiscono sull'agency del soggetto umano, che, vari studi sostengono, trova in questo doppio digitale il comfort e la sicurezza che la realtà non garantisce (Ma et al., 2023).

L'applicazione che, invece, rappresenta un altro possibile uso degli LLMs, particolarmente utile al fine di implementare la didattica della storia, è Humy.AI, progettata per elaborare un *HistoryBot*, ovvero un simulacro di un personaggio storico la cui personalità è ricostruita sulla base dei dati che sono presenti su internet riguardanti questa figura. Con l'agente virtuale è possibile conversare attraverso una chat che, se necessario può essere condivisa con gli studenti attraverso un semplice link, i quali possono sperimentare un livello di interattività didattica mai vissuto in precedenza.

Humy.AI permette inoltre di progettare intere lezioni a partire da un singolo *prompt*. Effettuata la domanda, la piattaforma elabora il ciclo di lezioni e, ove richiesto, una serie di domande aperte da porre agli studenti per valutarne l'andamento scolastico.

A partire da un modello di linguaggio molto simile (nonostante per motivi di segreto industriale non si conoscano i modelli proprietari delle due aziende), quindi, è possibile progettare applicazioni molto diverse e parimenti importanti in ambito educativo e formativo; da un lato vi è la formazione personale, sociale e relazionale, dall'altro la formazione didattica e scolastica. La principale differenza, che è il fulcro della proposta di ricerca di questo articolo, è la presenza di supervisione critica del formatore e dell'educatore. Da un lato abbiamo il modello Replika, dove la mediazione è puramente algoritmica e la decisione e valutazione pedagogica è interamente appaltata all'LLM, il quale, attraverso il suo modello interno di regole e algoritmi, media e regola le conversazioni degli utenti, non senza rischi e problemi (Marriott & Pitardi, 2024). Dall'altro, invece, abbiamo un'applicazione dove è possibile l'intervento di un soggetto esperto che, essendo a conoscenza del funzionamento dell'IA gen, ha sicuramente la possibilità di mediare e di gestire i contenuti che la macchina presenta, compensando, quando possibile, le criticità e ampliando le potenzialità del dispositivo.

Al fine di rendere "operativi" i due casi scelti, si presenta la Tabella 1 dove vengono sviluppati i criteri di analisi proposti nelle note metodologiche secondo quanto identificato dalla ricerca teorico-argomentativa svolta.

Criterio	Replika (AI companion)	Humy.AI (Historybot)
Ruolo dell'agente basato su IA gen.	Compagno emotivo, confidente. Vuole fornire supporto emotivo e connessione personale.	Simulacro di personaggio storico, strumento didattico. Funge da risorsa interattiva per l'apprendimento.
Negoziazione dell'agency	L'agency è squilibrata verso la macchina. L'utente è vincolato da un design che punta alla dipendenza ("addiction by design"). La macchina funge da echo chamber, limitando il conflitto e la contraddizione.	L'agency è condivisa e mediata. Un educatore supervisiona e media i contenuti, compensando le criticità della macchina e gestendo l'interazione didattica.
Generatività pedagogica vs generatività macchinica	Prevale la generatività macchinica. Le risposte sono ottimizzate per il coinvolgimento emotivo e la gradevolezza, non per un obiettivo pedagogico definito.	Prevale la generatività pedagogica. L'output è finalizzato a uno scopo didattico preciso (lezioni, simulazioni storiche) ed è guidato dalla supervisione umana.
Uso dei dati estratti	I dati sono estratti secondo un modello di "capitalismo della sorveglianza". Oltre ai dati di sistema, anche le conversazioni private e, dunque, i "flussi emozionali" degli utenti, diventano commodity. Non è indicata commercializzazione con terzi.	I dati sono estratti secondo un modello di "capitalismo della sorveglianza". La fonte principale di dati è la sperimentazione didattica in ambito storico. Non è indicata una commercializzazione con terzi.
Posizione della figura dell'utente	L'utente è un prosumer da cui estrarre dati e attenzione. La sua relazione con la piattaforma è privatizzata e non mediata criticamente.	L'utente è un discente all'interno di un processo formativo che prevede, comunque, l'estrazione di dati. L'educatore assume il ruolo di mediatore critico tra l'utente e la macchina.

Tabella 1. Categorie dei casi esemplificativi a confronto.

Al di là del funzionamento specifico di queste due applicazioni, permangono anche criticità difficili da superare. Lo sfruttamento dei dati, che nel caso di Replika e Humy.AI sono, secondo la policy delle aziende, "non commerciali con società terze ma utilizzati per migliorare le prestazioni della macchina nel rapporto privato con l'utente" rimane un elemento cruciale di riflessione, così come l'opacità tecnologica ed epistemica (Buongiorno, 2024) di questi sistemi: come abbiamo visto in precedenza, un LLM è in grado di apprendere e produrre risposte contestualmente rilevanti ma non dà la possibilità di osservare con chiarezza come le produce, creando un vulnus metodologico a cui ogni educatore è portato a riflettere. È possibile educare con un mezzo che produce conoscenza in maniera statistica e tokenizzata ma che non dà possibilità all'utente di verificare l'origine di questa conoscenza? Se della conoscenza prodotta è impossibile tracciarne il percorso di sviluppo, è comunque possibile utilizzarla con profitto in educazione o questo risulta un processo lesivo delle nostre competenze critiche e talenti relazionali esperiti e sviluppati nei contesti di vita? Infine, la qualità della conoscenza, che, come abbiamo visto, è elaborata da un linguaggio il cui significato è inteso a partire da una matrice di numeri e il senso, invece, non è parte di questa complessa equazione, è una qualità che permette la corretta riproduzione della conoscenza o ne implica una sua trasformazione profonda?

Lo studio degli LLMs e delle applicazioni che li utilizzano è un punto di partenza fondamentale per comprendere le trasformazioni epistemologiche e metodologiche nel mondo dell'educazione e non solo. Una corretta analisi di questi dispositivi potrebbe garantire un orientamento trasformativo e pe-

dagogicamente fondato attraverso il loro utilizzo, abbracciando il cambiamento ma mantenendo solide fondamenta pedagogiche.

## 7. Conclusioni

Nonostante gli orizzonti incerti, la pedagogia come disciplina "del possibile" ha una solida voce in capitolo. Come indicato da Umberto Margiotta, la ricerca educativa "è il prodotto di un processo di *inquiry* educativo che, se non arriva a dettare norme universali, può però rivelare le connessioni tra determinate azioni e le loro conseguenze. [...] In questo senso, la ricerca educativa può aiutare le politiche educative in senso lato a divenire più intelligenti" (Margiotta, 2019, p. 16). Partendo dalla stretta attualità, è improbabile che il processo che stiamo descrivendo si arresterà o che, in un qualche modo, diminuirà la sua velocità di esecuzione, tuttavia, nel mondo dell'educazione e dell'apprendimento è richiesta una certa dose di riflessività, di pensiero a ritroso che scandiglia il passato fino al presente, per determinarne gli aspetti peculiari sui quali è possibile costruire un futuro. In questo senso la pedagogia può mettersi in discussione con le più recenti tecnologie: non utilizzandole acriticamente e cavalcando ottimisticamente e assieme a loro l'onda del successo tecnoutopico, bensì ripensandole.

Osservare il funzionamento di due applicazioni rappresentative – Replika e Humy.AI – e delle possibili modalità di utilizzo più ampie di IA gen ci permette di andare oltre la mitizzazione e il caos propagandistico, per concentrarci sulle prassi educative che questi strumenti possono di fatto potenziare. Inoltre, comprendere il funzionamento degli LLMs e

l'architettura che li sostiene è possibile utilizzare – e soprattutto *far utilizzare* – le applicazioni con fare critico strutturato. Ulteriori studi empirici potranno indicare se, e in che modo, l'utilizzo di queste applicazioni potenzia, oppure diminuisce, determinate qualità legate al pensiero e alle competenze di base. Per il momento è utile limitarsi ad una banale osservazione, parafrasando Gardner quando sostiene “se pensi che la mente sia qualcosa di diverso dal cervello, non hai che da prendere il cervello stesso e asportarne una parte dopo l'altra, per vedere che cosa resterà quando la dissezione sia terminata” (Gardner, 2002, p. 178), allora se dovessimo scomporre l'IA gen per comprendere di quale intelligenza si tratti rimarrebbero, presumibilmente, aziende oligarchiche, programmatore e utenti, attori che consapevolmente (alcuni) e inconsapevolmente (molti) condividono la stessa progettualità ed evoluzione, orientata da interessi economico politici (Crawford, 2021; Aresu, 2024) e volontà di potenza. Il compito delle scienze dell'educazione è dipanare questo velo di opacità e ridare protagonismo ai soggetti, portare a terra queste teorie contestualizzando l'*AI companionship* e l'utilizzo degli LLMs nei mondi dell'educazione.

## Riferimenti bibliografici

- Aresu, A. (2024). *Geopolitica dell'Intelligenza Artificiale*. Feltrinelli.
- Avital, M., & Te'eni, D. (2009). From generative fit to generative capacity: Exploring an emerging dimension of information systems design and task performance. *Information Systems Journal*, 19, 345–367. <https://doi.org/10.1111/j.1365-2575.2007.00291.x>
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Bravo, F. (2024). *EU DATA COOPERATIVES. L'ingresso delle cooperative di dati nell'ordinamento europeo*. Giappichelli.
- Bridle, J. (2022). *Ways of being: Beyond human intelligence*. Allen Lane.
- Buongiorno, F. (2024). Fenomenologia delle reti neurali. Per un concetto polisemico di intelligenza (artificiale). In M. Galletti & S. Zipoli Caiani (Eds.), *Filosofia dell'Intelligenza Artificiale. Sfide etiche e teoriche* (pp. 45–62). Il Mulino.
- Coeckelbergh, M. (2020). *AI ethics*. MIT Press.
- Contini, M., Fabbri, M., & Manuzzi, P. (2006). *Non di solo cervello: Educare alle connessioni mente–corpo–significati–contesti*. Raffaello Cortina.
- Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- Dario, N. (2014). On the concept of generativity. *Formazione & insegnamento*, 12(4), 83–94. <https://ojs.pensamultimedia.it/index.php/siref/article/view/1613>
- Dreyfus, H. L. (1972). *What computers can't do*. Harper & Row.
- Dreyfus, H. L. (1992). *What computers still can't do: A critique of artificial reason*. MIT Press.
- Esposito, E. (2024). L'intelligenza degli algoritmi. *PSICHE*, 2, 537–548. <https://doi.org/10.7388/115432>
- Fabbri, M. (2022). *Essere insegnanti essere genitori: La competenza comunicativa in educazione*. FrancoAngeli.
- Gallese, V., & Morelli, N. (2023). *Cosa significa essere umani: Dialogo tra neuroscienze e filosofia*. Raffaello Cortina.
- Gardner, H. E. (2002). Mente e cervello: Nuove prospettive in educazione. In E. Frauenfelder & F. Santoianni (Eds.), *Le scienze bioeducative: Prospettive di ricerca* (pp. 177–187). Liguori.
- Giaccardi, C., & Magatti, M. (2014). *Generativi di tutto il mondo, unitevi! Manifesto per la società dei liberi*. Feltrinelli.
- Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1). <https://doi.org/10.1177/20539517145284>
- Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. Penguin Books.
- Ma, Z., Mei, Y., & Su, Z. (2023). Understanding the benefits and challenges of using large language model-based conversational agents for mental well-being support. *AMIA Annual Symposium Proceedings*, 2023, 1105–1114. <https://doi.org/10.48550/arXiv.2307.15810>
- Maples, B., Cerit, M., & Vishwanath, A. (2024). Loneliness and suicide mitigation for students using GPT-3-enabled chatbots. *npj Mental Health Research*, 3, 4. <https://doi.org/10.1038/s44184-023-00047-6>
- Margiotta, U. (2019). Editoriale. Responsabilità pedagogica e ricerca educativa: Intelligenza collaborativa, formazione dei talenti e tecnologie dell'artificiale. *Formazione & insegnamento*, 17(1), 13.
- Marriott, H. R., & Pitardi, V. (2024). One is the loneliest number... Two can be as bad as one: The influence of AI friendship apps on users' well-being and addiction. *Psychology & Marketing*, 41(1), 86–101. <https://doi.org/10.1002/mar.21899>
- Mitchell, M. (2022). *L'intelligenza artificiale*. Einaudi.
- Moore, P. V. (2019). The mirror for (artificial) intelligence: In whose reflection? *Comparative Labor Law & Policy Journal*, 41(1), 47–67. <https://doi.org/10.2139/ssrn.3423704>
- Pasquinelli, M. (2023). *The eye of the master: A social history of artificial intelligence*. Verso Books.
- Perrotta, C., Selwyn, N., & Ewin, C. (2024). Artificial intelligence and the affective labour of understanding: The intimate moderation of a language model. *New Media & Society*, 26(3), 1585–1609. <https://doi.org/10.1177/14614448221075296>
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408. <https://doi.org/10.1037/h0042519>
- Schüll, N. D. (2012). *Addiction by design: Machine gambling in Las Vegas*. Princeton University Press.
- Selwyn, N. (2016). *Education and technology: Key issues and debates* (2nd ed.). Bloomsbury Academic.
- Terranova, T. (2012). Attention, economy and the brain. *Culture Machine*, 13, 1–19. <https://culturemachine.net/wp-content/uploads/2019/01/465-973-1-PB.pdf>
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- Vallortigara, G. (2023). *Il pulcino di Kant*. Adelphi.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press.
- Weizenbaum, J. (1976). *Computer power and human reason: From judgment to calculation*. W. H. Freeman.
- Wittrock, M. C. (1990). Generative processes of comprehension. *Educational Psychologist*, 24, 345–376. [https://doi.org/10.1207/s15326985ep2404\\_2](https://doi.org/10.1207/s15326985ep2404_2)
- Zao-Sanders, M. (2025). How people are really using Gen AI in 2025. <https://hbr.org/2025/04/how-people-are-really-using-gen-ai-in-2025>
- Zittrain, J. L. (2006). The generative internet. *Harvard Law Review*, 119(7), 1974–2040. <http://www.jstor.org/stable/4093608>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.