

Paolo Barabanti

IA per valutare

*Il ruolo dell'Intelligenza Artificiale nella
costruzione di prove e rubriche di valutazione
degli apprendimenti*



Indice

Prefazione

MARCO LAZZARI - Professore Ordinario Università degli Studi di Bergamo 5

Introduzione 7

1. L'Intelligenza Artificiale: origini, tecnologie e ambivalenze 9

Una rivoluzionaria intelligenza non intelligente	10
L'Intelligenza Artificiale: definizioni, origini e sviluppi	13
Elementi di base e modello dell'Intelligenza Artificiale	14
Il machine learning	15
Il deep learning	15
L'Intelligenza Artificiale generativa	17
I Large Language Models	18
I chatbot e il prompt	19
I Big Data	22
L'ambivalenza dell'Intelligenza Artificiale	24

2. L'Intelligenza Artificiale nei contesti educativi e di apprendimento 27

Dalle macchine per l'insegnamento all'IA nell'educazione	28
L'Artificial Intelligence in Education	30
Usi e ambiti applicativi dell'Intelligenza Artificiale in istruzione	30
La personalizzazione delle esperienze di apprendimento	31
L'automazione delle attività didattiche e valutative	32
Il supporto informativo ai processi educativi	33
Tensioni pedagogiche, etiche e organizzative	33
Condizioni, limiti e responsabilità	35
Condizioni per l'integrazione dell'Intelligenza Artificiale a scuola	36
L'adozione dell'IA dal punto di vista dei docenti: i risultati dell'indagine OCSE-TALIS 2024	40
Verso un'alfabetizzazione critica all'Intelligenza Artificiale	43

3. Intelligenza Artificiale e valutazione degli apprendimenti 46

Potenzialità del processo di valutazione degli apprendimenti tramite l'Intelligenza Artificiale	46
---	----

L'Intelligenza Artificiale a supporto del processo di valutazione	49
La definizione e formulazione degli obiettivi	49
Fase istruttoria	49
Fase di giudizio	52
Validità e attendibilità	53
Trasparenza e responsabilità	53
Il contributo alla personalizzazione	54
Modelli per orientare	55
Assistita anziché automatizzata	55
Il modello SAMR	55
L'Artificial Intelligence Assessment Scale	57
Il prompt	59
Gli elementi costitutivi del prompt	60
Le strategie di prompting	64
4. Chatbot, prompting e valutazione degli output: metodologia e disegno della ricerca	67
Finalità, obiettivi e ipotesi di ricerca	67
Metodo e strumenti	69
La costruzione dei prompt	69
La selezione delle strategie di prompting	70
Chatbot utilizzati e contesti didattici	72
Valutazione della qualità degli output	74
5. Analisi e discussione dei risultati	81
Item vero/falso	81
Item a scelta multipla	84
Item di matching univoco	89
Item a domanda aperta univoca	92
Item di cloze guidato	95
Le rubriche	97
Discussione dei risultati	101
Verifica delle ipotesi di ricerca	103
Conclusione: una responsabilità professionale non delegabile	106
Postfazione	
SERAFINA PASTORE - Professoressa Associata Università degli Studi di Bari "Aldo Moro"	110
Bibliografia	113

Prefazione

MARCO LAZZARI

Professore Ordinario Università degli Studi di Bergamo

Sarà che negli anni dell'università vivevo a duecento metri dal Museo Nazionale di San Matteo, un vero scrigno di arte medievale, ma sono un appassionato di pittura pre-giottesca e ho un debole per i politici. Ho pensato allora di organizzare questa prefazione come un trittico, con una tavola per ciascuna delle sezioni principali del bel libro di Paolo Barabanti (ho avuto la fortuna di leggere un impaginato senza indice, così non mi lascio fuorviare dall'*intentio auctoris* che potrebbe essere espressa dal sommario). Nella prima tavola il lettore troverà l'Intelligenza Artificiale e il suo ingresso nei processi di insegnamento e apprendimento; nella seconda, ciò che l'Intelligenza Artificiale può fare nella e per la valutazione; nella terza, una ricerca empirica condotta dall'autore. Chi ha fretta può saltare direttamente al libro, che sta in piedi benissimo per conto suo; chi resta, spero trovi in queste righe un motivo in più per voltare pagina.

Prima tavola. Da qualche tempo assistiamo a un curioso fenomeno di trasmutazione professionale: umanisti che, folgorati sulla via dei chatbot, si improvvisano esperti di Intelligenza Artificiale e riempiono pagine ironizzando su quante erre ci siano in *strawberry*, mettendo alla berlina l'allucinazione del giorno oppure sollevando gravi questioni etiche senza peraltro proporre una soluzione. La prima parte di questo libro è l'antidoto a quella letteratura: la sua prima sezione, agile ma ricca e precisa, ci spiega come funzionano davvero i sistemi di Intelligenza Artificiale, da dove vengono, che cosa possono e che cosa non possono fare. Barabanti non concede nulla né agli entusiasmi né agli anatemi e ci ricorda con nettezza che il sistema «non “capisce” realmente il contenuto» che elabora e che «il modello non conosce realmente il significato delle parole che seleziona». Per questo l'autore rifiuta la scorciatoia del rigetto e chiama in causa la «consapevole mediazione umana» come condizione per un uso sensato degli strumenti di Intelligenza Artificiale generativa. E nel momento in cui lo sguardo si sposta sulla scuola, la questione viene posta nei termini che mi paiono più onesti e più fecondi: «non si tratta più di stabilire se utilizzare oppure no l'IA ma di capire come studenti, studentesse e insegnanti la possano impiegare in modo consapevole ed esperto». Barabanti accetta la sfida della complessità e prova a immaginare le condizioni che possono favorire la coevoluzione di ritrovati tecnici e istituzioni scolastiche, per una scuola che accompagni e riconosca le persone che le sono affidate. Di qui l'invito, che sottoscrivo, a «riconoscere l'*AI literacy* come una priorità educativa strategica» per la nostra scuola.

Seconda tavola. Il pannello centrale è dedicato all'incontro tra l'Intelligenza Artificiale e la valutazione degli apprendimenti. Anche qui l'autore evita la tentazione di presentare come inaudito quello che non lo è: l'idea di affidare alle macchine compiti valutativi ha una storia lunga più di mezzo secolo, ma ciò che è cambiato è la portata. Infatti, la valutazione assistita dall'Intelligenza Artificiale generativa, oltre a poter automatizzare le pratiche che già esistono, può incidere sulla natura stessa dell'atto valutativo, sul rapporto tra valutatore e valutato e sui criteri con cui le evidenze di apprendimento diventano giudizi. Per orientarsi in questo territorio Barabanti offre al lettore mappe e *framework*, con una distinzione che possiamo considerare una bussola: valutazione assistita, anziché automatizzata. E dedica pagine preziose al prompt, che ci viene proposto come un autentico dispositivo progettuale, nel quale si riversano le competenze docimologiche e didattiche di chi lo scrive. O non si riversano, e allora sono guai.

Terza tavola. Il terzo pannello è quello che, da lettore di ricerche altrui per mestiere, ho apprezzato di più (senza con questo voler sminuire gli altri): una ricerca empirica pregevole, ricca e articolata, costruita attorno alla «tensione tra potenzialità generativa dei sistemi di IA e correttezza docimologica degli strumenti di valutazione». Barabanti sottopone tre chatbot di ampia diffusione a un confronto sistematico, con un disegno a metodi misti e quattro ipotesi chiare: la strutturazione del prompt migliora la qualità degli output? La presenza di esempi riduce gli errori formali? Modelli diversi producono risultati diversi a parità di richiesta? Lo stesso modello dà risposte diverse se si ripete l'interrogazione? Non anticipo i risultati per non rovinare la sorpresa, ma dico che ci restituiscono un quadro istruttivo di vittorie parziali e di fallimenti momentanei dell'Intelligenza Artificiale generativa, dal quale l'autore ricava indicazioni operative fondate sull'evidenza. Appoggiandosi a un'avvertenza metodologica espressa nella precedente discussione teorica, che vale come programma: «l'integrazione dell'IA nella valutazione degli apprendimenti richiede modelli interpretativi in grado di evitare sia approcci riduzionisti centrati sull'automazione, sia forme di adozione acritica della tecnologia». Né apocalittici, né integrati, verrebbe da dire; piuttosto, professionisti consapevoli.

Nei polittici che amo, sotto le tavole principali corre spesso una predella, con le sue storie minute che completano il racconto; e talvolta una cimasa corona l'insieme. Questo volume, che si apre con un trittico, si chiuderà con una postfazione: al mio pannello iniziale risponderà dunque un'altra voce, e mi pare giusto così, perché un libro che invita alla mediazione consapevole e al confronto merita di essere incorniciato da più sguardi. In mezzo, il lettore troverà un lavoro serio e leggibile, rigoroso senza essere arcigno, aggiornato senza essere modaiolo, che indica ciò che serve per governare gli strumenti dell'Intelligenza Artificiale senza soccombere alla loro seduzione. Lo dico con la libertà che spero si possa concedere a chi scrive le prefazioni dei libri degli amici, ma anche con la convinzione di chi di testi ne ha letti tanti: questo è un libro utile, e la sua lettura è tempo bene speso.

Introduzione

L'Intelligenza Artificiale generativa ha infranto un muro che credevamo invalicabile: quello della produzione di contenuti complessi in tempi istantanei. Ciò che un tempo richiedeva ore di lavoro e di riflessione, oggi arriva in pochi secondi grazie a un prompt. Non siamo di fronte a una semplice innovazione tecnica ma a un cambiamento antropologico epocale: la macchina non si limita più a potenziare (o sostituire?) i nostri muscoli ma entra (o prende il posto?) nel cuore pulsante delle nostre facoltà superiori.

Di fronte a questi cambiamenti, la scuola non può restare una spettatrice muta. L'Intelligenza Artificiale generativa sta di certo offrendo nuovi attrezzi del mestiere ma sta soprattutto riscrivendo le regole del gioco: incide su come si progetta, su come si insegna e anche su come si valutano gli apprendimenti. Sempre più spesso, infatti, i chatbot diventano *ghostwriter* di verifiche, item e rubriche, assistendo così in modo invisibile, in guisa di “eminenze grigie”, il lavoro docente.

Ma la velocità e la riduzione di fatica possono essere sinonimo di valore? Un output dell'Intelligenza Artificiale può essere formalmente impeccabile, elegante nella sintassi e seducente nella forma, eppure rivelarsi un pacco regalo vuoto o, peggio ancora, un contenitore luccicante di una sorpresa molto rischiosa: quesiti ambigui, distrattori deboli e incoerenze docimologiche possono annidarsi sotto una superficie di apparente perfezione o essere spesso sacrificati sull'altare della tempestività esecutiva.

È proprio in questa faglia tra la potenza di calcolo della macchina e il rigore del docente che si colloca questo volume. Qui non si troveranno né entusiasmi acritici né arroccamenti nostalgici né derive apocalittiche. L'obiettivo è un altro: smontare il giocattolo, analizzarne gli ingranaggi e osservare cosa accade realmente quando la logica dell'algoritmo incontra la complessità della valutazione.

Il percorso di indagine si snoda lungo due binari complementari, progettati per offrire sia una solida base concettuale sia un riscontro empirico. Il primo binario, di natura teorico-ermeneutica, si addentra nell'identità profonda di questi sistemi. Non ci si limita a descrivere la tecnologia ma si analizza l'Intelligenza Artificiale come il punto più avanzato di una secolare “catena di deleghe” che oggi trasferisce alla macchina anche l'apparire intelligente delle funzioni cognitive superiori. Il cuore di questa riflessione è la ridefinizione del prompt: esso viene spogliato della veste di semplice comando tecnico per essere elevato a dispositivo progettuale. Il prompt emerge come una vera e propria bussola metodologica indispensabile per addomesticare il caos generativo della macchina e orientare il calcolo probabilistico verso esiti dotati di senso pedagogico.

Il secondo binario è di carattere empirico. Questa ricerca adotta un approccio comparativo a metodi misti per sottoporre tre tra i modelli più avanzati – ChatGPT, Microsoft Copilot e Gemini – a uno stress test metodologico. Sono state messe alla prova quattro ipotesi fondamentali: l'effetto della strutturazione del prompt sull'innalzamento della qualità degli output; il ruolo dell'esemplificazione nel *prompting* per ridurre gli errori formali e migliorare quanto prodotto; l'esistenza di differenze di qualità tra modelli diversi a parità di prompt; la possibilità che uno stesso chatbot generi output qualitativamente differenti tra tentativi successivi. Attraverso l'incrocio di diverse strategie di interrogazione con compiti valutativi specifici, il volume prova a mappare (parziali) vittorie e (momentanei) fallimenti dell'Intelligenza Artificiale generativa. I risultati mettono a nudo la tensione tra l'istantaneità delle risposte e l'economia di sforzo offerta dalle macchine rispetto al rigore dei criteri docimologici. Il lavoro offre così non solo un'analisi critica ma anche indicazioni operative fondate sull'evidenza, affinché i docenti e le docenti possano governare l'algoritmo senza soccombere alla sua fatale seduzione.

I nostri libri

Vai al catalogo >>



Di seguito trovi una selezione di volumi della collana “Insegnare nel XXI secolo” dedicati ai temi del digitale e dell’Intelligenza Artificiale

Tra carta e schermo

Riflessioni e strategie didattiche per educare alla lettura

Andrea Nardi



Inquadra il QrCode per saperne di più e procedere all'acquisto



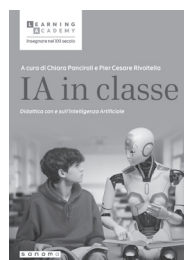
IA in classe

Didattica con e sull’Intelligenza Artificiale

A cura di Chiara Panciroli, Pier Cesare Rivoltella



Inquadra il QrCode per saperne di più e procedere all'acquisto



Educare all’IA

La sfida didattica dell’Intelligenza Artificiale: ChatGPT e Gemini

Massimiliano Badino, Fabio Aurelio D’Asaro, Francesco Pedrazzoli



Inquadra il QrCode per saperne di più e procedere all'acquisto

