Gemini come assistente valutativo: costruire rubriche con l'Intelligenza Artificiale Generativa

Gemini como asistente de evaluación: crear rúbricas con inteligencia artificial generativa

Gemini as an assessment assistant: creating rubrics with Generative Artificial Intelligence

Marta Salvucci, PhD
Collaboratrice di ricerca e Docente a contratto
Università degli Studi di Urbino Carlo Bo
marta.salvucci@uniurb.it

RIASSUNTO
L'intelligenza artificiale (IA) è sempre più integrata nel settore dell'istruzione, offrendo vantaggi quali l'apprendimento personalizzato, la valutazione adattiva, l'analisi predittiva e gli agenti conversazionali. L'IA generativa (Gen-AI) solleva interrogativi su cosa dovrebbero insegnare gli educatori e su come valutare gli studenti, evidenziando la necessità di una "collaborazione con l'IA" e di un coinvolgimento degli studenti di qualità superiore. I sistemi Gen-AI, come Gemini, generano testo, immagini e contenuti multimediali apprendendo modelli dai dati di addestramento. Gemini, un modello linguistico di grandi dimensioni, svolge molteplici compiti, producendo output contestualmente rilevanti, ma la guida umana rimane fondamentale. I modelli linguistici di grandi dimensioni (LLM) non sono in grado di ragionare, creare o valutare criticamente in modo indipendente, rendendo l'intervento dell'insegnante essenziale per un apprendimento significativo.
Nella valutazione, l'IA consente test adattivi, feedback automatizzati e monitoraggio basato sui dati, sostenendo un approccio antropocentrico in cui gli insegnanti mantengono la responsabilità primaria. Le rubriche esemplificano questa sinergia: criteri strutturati, comportamenti osservabili e livelli di padronanza guidano la valutazione. Gemini può aiutare nella stesura delle rubriche, ma il contributo umano affina i descrittori, aggiunge chiarezza e garantisce l'allineamento con gli obiettivi di apprendimento.
Nel complesso, la Gen-AI migliora la precisione e l'efficienza dell'istruzione, ma non può sostituire gli insegnanti. Un'integrazione efficace dipende da educatori competenti che guidino criticamente l'IA, ne convalidino i risultati e ne garantiscano un uso etico e pedagogicamente corretto, rendendo l'IA uno strumento di supporto per favorire l'apprendimento e il successo degli studenti.

RESUMEN
La inteligencia artificial (IA) se integra cada vez más en la educación, ofreciendo ventajas a través del aprendizaje personalizado, la evaluación adaptativa, el análisis predictivo y los agentes conversacionales. La IA generativa (Gen-IA) plantea interrogantes sobre qué deben enseñar los educadores y cómo evaluar a los alumnos, lo que pone de relieve la necesidad de una «colaboración con la IA» y una mayor implicación de los alumnos. Los sistemas de Gen-IA, como Gemini, generan texto, imágenes y contenido multimedia aprendiendo patrones a partir de datos de entrenamiento. Gemini, un modelo de lenguaje a gran escala, realiza múltiples tareas y produce resultados contextualmente relevantes, pero la orientación humana sigue siendo crucial. Los modelos de lenguaje a gran escala (LLM) no pueden razonar, crear ni evaluar críticamente de forma independiente, lo que hace que la intervención del profesor sea esencial para un aprendizaje significativo.
En la evaluación, la IA permite realizar pruebas adaptativas, proporcionar retroalimentación automatizada y llevar a cabo un seguimiento basado en datos, lo que respalda un enfoque antropocéntrico en el que los profesores conservan la responsabilidad principal. Las rúbricas ejemplifican esta sinergia: criterios estructurados, comportamientos observables y niveles de dominio guían la evaluación. Gemini puede ayudar a redactar rúbricas, pero la aportación humana perfecciona los descriptores, aporta claridad y garantiza la alineación con los objetivos de aprendizaje.
En general, la IA de última generación mejora la precisión y la eficiencia educativas, pero no puede sustituir a los docentes. Una integración eficaz depende de educadores cualificados que guíen críticamente a la IA, validen los resultados y garanticen un uso ético y pedagógicamente sólido, convirtiendo a la IA en una herramienta de apoyo para fomentar el aprendizaje y el éxito de los estudiantes.

ABSTRACT
Artificial Intelligence (AI) is increasingly integrated into education, offering benefits through personalized learning, adaptive assessment, predictive analytics, and conversational agents. Generative AI (Gen-AI) raises questions about what educators should teach and how to assess students, highlighting the need for "collaboration with AI" and higher-quality student engagement. Gen-AI systems, like Gemini, generate text, images, and multimedia by learning patterns from training data. Gemini, a large language model, performs multiple tasks, producing contextually relevant outputs, but human guidance remains crucial. LLMs cannot reason, create, or critically evaluate independently, making teacher intervention essential for meaningful learning.
In assessment, AI enables adaptive tests, automated feedback, and data-driven monitoring, supporting an anthropocentric approach where teachers retain primary responsibility. Rubrics exemplify this synergy: structured criteria, observable behaviors, and mastery levels guide evaluation. Gemini can assist in drafting rubrics, but human input refines descriptors, adds clarity, and ensures alignment with learning goals.
Overall, Gen-AI enhances educational precision and efficiency but cannot replace teachers. Effective integration depends on skilled educators who critically guide AI, validate outputs, and ensure ethical, pedagogically sound use, making AI a supportive tool for fostering learning and student success.

Parole chiave: IA generativa; Istruzione; Valutazione; Modelli linguistici di grandi dimensioni (LLM); Rubriche; Guida da parte degli insegnanti

Palabras llave: IA generativa; Educación; Evaluación; Modelos de lenguaje a gran escala (LLM); Rúbricas; Orientación docente

Keywords: Generative AI; Education; Assessment; Large Language Models (LLM); Rubrics; Teacher guidance

1. GEN-AI NELLA DIDATTICA

L'IA è ormai radicata nell'istruzione dove promette vantaggi tramite piattaforme personalizzate, sistemi di valutazione adattivi, analisi predittive intelligenti e agenti conversazionali (Akgun & Greenhow, 2021). A sostegno di ciò, tra le raccomandazioni per una società della buona IA, Floridi (2022) suggerisce di integrare nei curricoli scolastici l'informatica come disciplina fondamentale, mentre le recenti Linee Guida del Ministero dell'Istruzione e del Merito (2025) inquadrano l'IA come uno strumento per la personalizzazione dell'apprendimento e il supporto alla professionalità docente.

Tuttavia, strumenti generativi come ChatGPT, capaci di risposte di qualità elevata su molteplici compiti educativi, hanno sollevato interrogativi fondamentali su cosa debbano insegnare gli educatori e come debbano essere valutati gli studenti. Occorre attivare «forme di collaborazione con l'IA» (Rossi, 2024) e, per fare ciò, risulta necessario, in primis, modificare il sistema formativo e di valutazione in modo da richiedere agli studenti[1] un impegno maggiore e una qualità di lavoro più elevata.

Quando parliamo di intelligenza artificiale generativa (dall'inglese Gen-AI) intendiamo l'insieme di sistemi capaci di creare testi, immagini e altri tipi di contenuti multimediali sfruttando modelli generativi. Tali modelli apprendono schemi e strutture presenti nei dati di addestramento e li utilizzano per generare nuovi contenuti che ne riprendono le caratteristiche principali e le stesse proprietà strutturali (Sengar et al., 2025).

L'introduzione dell'IA generativa nella scuola è una necessità per lo sviluppo della cittadinanza digitale. Secondo i framework europei DigComp 2.1 e 2.2, comprendere i bias e gestire i dati è oggi una competenza di base dei cittadini (Vourikari et al., 2022). L'adozione della Gen-AI nella didattica deve rispettare rigorosi requisiti tecnici e normativi. Il MIM e il DigComp 2.2 sottolineano l'importanza di tutelare i dati personali e di adottare sistemi in grado di rendere comprensibili le decisioni algoritmiche, evitando che la scuola si trasformi in una "scatola nera" imposta dall'esterno.

Qui nello specifico sperimenteremo uno strumento di GenAI noto come Gemini. Gemini è un modello generativo di intelligenza artificiale con caratteristiche distintive rispetto a sistemi come ChatGPT. Addestrato su dataset diversificati, Gemini può svolgere diversi compiti in un unico sistema, generando contenuti contestualizzati su più paragrafi e argomenti. Fa parte della categoria dei *Large Language Models* (LLM), ovvero modelli di linguaggio di grandi dimensioni capaci di comprendere, generare e completare testi basandosi sui dati di addestramento. Gemini analizza i contesti e le probabilità di risposta, utilizzando strategie avanzate come beam search, ensemble, rinormalizzazione e training incrementale, migliorando efficienza e adattabilità ai nuovi dati. Testato in scenari reali, dimostra prestazioni elevate grazie alla capacità

[1] Il termine studente e i sostantivi al maschile sovraesteso sono utilizzati in questo scritto come forme neutre, senza alcuna intenzione discriminatoria o di esclusione.

di integrare più compiti, gestire contesti complessi e produrre risposte di alta qualità (Alsajri et al., 2024).

In questo articolo non vogliamo demonizzare l'uso della GenAI, anzi, lo vogliamo sostenere ma con il fondamentale apporto del docente. Questo perché, come ci ricorda Thorpe (2024), i LLM commettono errori. Inoltre, non essendo umani, non possono ragionare o pensare in modo critico e creativo come facciamo noi. Non possiedono la stessa sfumatura, il contesto o l'interpretazione individuale e potrebbero non essere altrettanto motivati a controllare i propri pregiudizi e limiti. Questo evidenzia, ancora una volta, il valore unico che solo le persone possono apportare alla conoscenza. Inoltre, gli LLM si basano su idee già esistenti, rendendoli fondamentalmente convenzionali. Infatti, i testi generati dalla GenAI sono formulati da umani in quanto vengono ripresi dal web. Per creare qualcosa di nuovo, per generare vera conoscenza, è, quindi, indispensabile l'intervento umano.

2. INTELLIGENZA ARTIFICIALE E VALUTAZIONE

Le nuove tecnologie e le soluzioni alimentate dall'intelligenza artificiale stanno producendo cambiamenti rilevanti in ogni ambito della ricerca scientifica, influenzando in modo significativo anche il settore della valutazione in ambito educativo. L'integrazione dell'IA nei sistemi di valutazione rappresenta uno dei cambiamenti più radicali nel panorama educativo contemporaneo.

Grazie all'IA, le prove possono essere adattate al livello e alle esigenze di ciascuno studente, risultando più motivate e adeguate alle reali capacità individuali. L'automazione consente di creare, somministrare e correggere test in modo rapido ed efficiente, liberando tempo per attività didattiche di maggiore valore. Le tecnologie di elaborazione del linguaggio permettono inoltre di analizzare in modo accurato produzioni scritte e orali, offrendo feedback immediati e personalizzati in diversi formati. L'IA amplia le tipologie di verifica, includendo simulazioni e modalità interattive anche con elementi di gamification, e rende possibile una valutazione più frequente e continua. Attraverso l'analisi dei dati, docenti e istituzioni possono individuare lacune, migliorare i materiali e monitorare con precisione i progressi. In questo modo, la valutazione non è più solo uno strumento di misurazione, ma diventa parte integrante di un percorso di apprendimento dinamico, personalizzato ed efficace (Jana et al., 2025).

In sintesi l'IA può supportare i docenti nei processi di valutazione in diversi modi: nella creazione di domande e prove d'esame, nell'analisi automatica dei testi scritti, nell'utilizzo dei dati generati durante il percorso di apprendimento, nella realizzazione di valutazioni adattive e personalizzate, nella correzione automatica di compiti ed esami, nel monitoraggio del coinvolgimento degli studenti e del rispetto dell'integrità

accademica, oltre che nella valutazione dei processi di apprendimento e delle performance collaborative (Bower et al., 2024).

Un ambito particolarmente innovativo è la scuola primaria, dove l'IA generativa viene utilizzata per potenziare il giudizio descrittivo: i docenti possono elaborare feedback più ricchi, oggettivi e personalizzati, aiutando lo studente a comprendere il proprio percorso di crescita (Di Padova et al., 2024).

Alla base di tutto ciò vi deve essere un modello di valutazione assistita dall'Intelligenza Artificiale di tipo antropocentrico, in cui umano e macchina collaborano senza che quest'ultima assuma piena responsabilità. Il ruolo principale nella valutazione deve restare al docente, così da preservare un approccio personale ed equilibrato nel processo educativo (De Martino et al., 2025). Per fare ciò, però, è fondamentale che i docenti possiedano le competenze per utilizzare strumenti di IA come supporto alla valutazione, infatti, il successo dipende dalla capacità del sistema educativo di formare docenti consapevoli, in grado di coniugare innovazione e rigore pedagogico in modo che l'IA possa diventare un alleato efficace (Isidori et al., 2025).

In particolare, il progettista (o il docente, in questo caso) deve conoscere non solo le funzioni basilari degli strumenti fondati su IA e possedere la capacità di applicarle alla progettazione della valutazione, ma deve essere in grado di riflettere criticamente rispetto a tali applicazioni, tenendo conto, in particolare, degli aspetti etici (equità, responsabilità, trasparenza, etica, privacy e sicurezza) relativi al loro utilizzo (Di Padova et al., op cit).

Se regolamentata e utilizzata in modo consapevole, infatti, l'Intelligenza Artificiale può favorire un modello educativo maggiormente centrato sullo studente, contribuendo a migliorare sia le pratiche di valutazione sia la qualità complessiva dell'istruzione (Mousa, 2025).

3. PROGETTARE RUBRICHE CON GEMINI

Uno strumento di valutazione a supporto del docente è la rubrica di valutazione. McTighe e Ferrare (1996) definiscono la rubrica come

uno strumento generale di valutazione impiegato per valutare la qualità dei prodotti e delle prestazioni in un determinato ambito. La rubrica consiste in una scala di punteggi prefissati e in una lista di criteri che descrivono le caratteristiche di ogni punteggio della scala. Le rubriche sono frequentemente accompagnate da esempi di prodotti o di prestazioni che hanno lo scopo di illustrare ciascuno dei punteggi (p.8).

Per valutare le competenze attraverso una rubrica è necessario un approccio strutturato. Occorre chiarire quali competenze si intendono osservare, scomponendole nei loro elementi fondamentali: quali dimensioni le caratterizzano? Quali aspetti risultano essenziali per descriverle in modo completo? Successivamente è importante definire i criteri di valutazione, cioè gli elementi in base ai quali formulare il giudizio.

In altre parole: su quali parametri si fonda la valutazione? Una volta stabiliti i criteri, bisogna individuare le evidenze osservabili, ossia gli indicatori concreti che permettono di riconoscere la presenza e la qualità della competenza. Quali comportamenti, produzioni o atteggiamenti rendono visibile l'apprendimento? Infine, è necessario stabilire il livello di padronanza raggiunto, graduando la performance lungo una scala che consenta di descrivere in modo chiaro il grado di sviluppo della competenza.

Occorre, quindi, effettuare scelte fondamentali: esplicitare i risultati di apprendimento attesi in termini concreti e osservabili, progettare compiti autentici o prove di prestazione per verificarli e, come già esplicitato, definire l'impianto valutativo con scale, livelli e descrittori qualitativi, così da garantire coerenza, trasparenza e affidabilità al processo di valutazione (CSTOLDI, 2009).

Quando parliamo di valutazione nella scuola, spesso pensiamo a un momento finale, ma la valutazione può – e dovrebbe – essere molto di più. Può diventare uno strumento di ricerca capace di raccogliere informazioni affidabili sul percorso di apprendimento e di guidare azioni concrete per migliorarlo. Nella didattica, assume davvero valore solo se sostiene l'apprendimento degli studenti, accompagnandoli e orientandoli lungo tutto il percorso, e non limitandosi a misurarlo a posteriori.

L'insegnante ha il compito di mettere i propri studenti in condizione di essere pienamente consapevoli di ciò che accadrà durante le lezioni: cosa impareranno, come verranno guidati nel processo di apprendimento e come saranno valutati. Le rubriche, infatti, stabiliscono cosa è rilevante e quali standard di qualità ci si aspetta. Ancora più importante, non si limitano a fornire un semplice punteggio: rappresentano una guida dettagliata per valutare le prestazioni degli studenti, organizzandole, come abbiamo detto, secondo una gerarchia di livelli (D'Ugo, 2012).

La rubrica si compone di alcuni elementi essenziali. Prima di tutto, le dimensioni o tratti: gli aspetti specifici del lavoro dello studente che vogliamo osservare. Poi, le descrizioni, che chiariscono cosa intendiamo per ciascuna dimensione, evitando ambiguità. Serve anche una scala di valore, come abbiamo più volte ribadito, che permetta di graduare le performance, e, infine, degli esempi o standard di eccellenza, modelli concreti che rendono chiari i diversi livelli di padronanza (Herman et al., 1992).

L'uso di rubriche standardizzate è raccomandato dagli studiosi in quanto riduce i pregiudizi e garantisce valutazioni coerenti tra diversi valutatori. L'uso di criteri uniformi assicura coerenza nella valutazione, rendendo i giudizi trasparenti (Nguyen et al., 2025).

Come abbiamo compreso sino ad ora, le rubriche costituiscono uno strumento fondamentale per stabilire aspettative chiare e fornire criteri di valutazione strutturati. Spesso, però, risulta complesso allineare tali criteri agli obiettivi di apprendimento e alle finalità di un compito. È qui che entra in gioco l'intelligenza artificiale generativa (GenAI) in quanto supporto ai docenti per creare rubriche chiare, dettagliate e di qualità. L'introduzione dell'intelligenza artificiale generativa ha aperto nuove strade nella progettazione delle rubriche di valutazione, consentendo ai docenti di arricchi-

re e rendere più efficaci le proprie strategie didattiche. Con il continuo sviluppo di questi strumenti, diventa sempre più utile per gli insegnanti sperimentare e integrare l'IA nelle loro pratiche, sfruttandone le potenzialità per sostenere e potenziare l'apprendimento degli studenti. È importante sottolineare che la GenAI non sostituisce il docente: il ruolo dell'insegnante resta centrale, poiché è lui a definire il compito, gli obiettivi di apprendimento e i criteri di valutazione, fornendo i prompt necessari affinché la GenAI produca una rubrica adeguata alle proprie esigenze. Lee Bonstra (2024) fornisce un metodo per generare i prompt che può essere riassunto nelle seguenti fasi:

- Design (Progettazione): definire chiaramente l'obiettivo, il contesto e i vincoli. Usare verbi d'azione forti (Analizza, Genera, Estrai).
- Implementation (Strutturazione): organizzare il prompt usando delimitatori per separare le istruzioni dai dati.
- Refinement (Affinamento): fornire esempi (Few-shot) se l'output iniziale è ambiguo.
- Evaluation & Iteration (Valutazione): testare il prompt, analizzare gli errori e modificare i parametri o il testo finché l'output non è costante.

Inoltre, tra le best practices fondamentali ci ricorda di essere specifici, di utilizzare istruzioni positive e di documentare tutto.

Al fine di dimostrare l'importanza del ruolo del docente nella collaborazione con la GenAI per strutturare rubriche oggettive, è stato messo alla prova Gemini che non nasce per creare rubriche di valutazione ma che, solo se guidato da prompt ben strutturato, può produrre contenuti utili. Riprendendo il lavoro fatto da D'Ugo et al. (2024) insieme a delle insegnanti dell'IC Fermi di Mondolfo, sono state seguite le stesse tappe adottate delle insegnanti per costruire le rubriche di valutazione per la loro scuola. Insieme alle docenti, infatti, sono state costruite delle rubriche di valutazione progettate per rilevare e documentare alcune tra le competenze più significative nello sviluppo degli alunni della scuola primaria, con particolare riferimento alle discipline di italiano e matematica. Le docenti hanno, in primis, riflettuto su quali risultati ritenessero davvero importanti: quali abilità dovrebbe aver maturato l'alunno? In quali comportamenti osservabili si traduce tale apprendimento? Quali segnali indicano che il percorso sta dando frutti o che la competenza è stata effettivamente raggiunta? In seguito, hanno individuato criteri chiari, cioè punti di riferimento che aiutano a distinguere i diversi livelli di padronanza. I criteri, infatti, rendono visibile la qualità della prestazione e permettono di collocarla in modo coerente lungo la scala dei livelli previsti. Infine, hanno tenuto conto del contesto in cui l'alunno può dimostrare ciò che sa fare. Una competenza emerge solo se viene messa in gioco in una situazione adeguata e significativa. È quindi necessario proporre attività e compiti che offrano l'opportunità concreta di manifestare le abilità acquisite. Solo in questo modo la valutazione diventa autentica e realmente utile a comprendere il livello di apprendimento raggiunto (Ibidem).

A questo punto, riprendendo una delle competenze oggetto di osservazione delle docenti dell'IC Fermi e riprendendo le varie fasi esposte sopra, è stato elaborato il seguente prompt per Gemini:

Elabora una rubrica di valutazione analitica per la scuola primaria.

Disciplina: Italiano
Classe: Terza primaria

Competenza di riferimento:
Ascoltare e comprendere messaggi e brevi testi.

Dimensione specifica da valutare:
"Fornire risposte corrette alle domande di comprensione del testo e saper esporre oralmente il contenuto ascoltato."

Consegna di lavoro (modalità e materiali):
- L'alunno ascolta un testo narrativo o informativo adeguato alla classe terza.
- Risponde in autonomia a un questionario a scelta multipla.
- Espone oralmente il contenuto del testo ascoltato.
- Il compito si svolge in una situazione nota.

Segui questi passaggi nella costruzione della rubrica:
- Definisci in modo sintetico la competenza osservata.
- Individua criteri osservabili coerenti con la consegna (ad esempio: correttezza delle risposte, autonomia, comprensione delle informazioni esplicite, organizzazione dell'esposizione orale, uso di lessico adeguato, completezza).

Costruisci una rubrica in tabella con 5 livelli di padronanza:
- Avanzato
- Intermedio
- Base
- In via di prima acquisizione
- Non ancora acquisito

Per ciascun livello scrivi un descrittore:
- chiaro e concreto
- osservabile
- progressivo (differenze nette tra i livelli)
- riferito a comportamenti reali dell'alunno
- adatto a bambini di classe terza
- coerente con una situazione nota

Nella prima rubrica elaborata da Gemini in seguito al prompt sopra citato alcune formulazioni erano troppo generiche e mancavano indicatori concreti osservabili.

Per rendere la valutazione più oggettiva, è stato necessario un ulteriore intervento "umano", fornendo istruzioni precise: usare verbi chiari, quantificatori e brevi esempi tra parentesi per indicare cosa osservare.

È stato, quindi, fornito a Gemini un ulteriore prompt:

La rubrica è valida, ma evita espressioni generiche come 'in modo adeguato'. Utilizza verbi osservabili e concreti, ad esempio individua, seleziona, espone, organizza, utilizza, ricostruisce. Inserisci quantificatori chiari (ad esempio, 'la maggior parte' corrisponde a quante risposte corrette?) e, tra parentesi tonde, fornisci brevi esempi che aiutino a capire esattamente cosa osservare. Di seguito (Tab. 1) la versione finale.

TAB 1. Rubrica generata da Gemini 3 a gennaio 2026

Livello	**Descrittore di Padronanza (Osservazioni e Comportamenti)**
Avanzato	**Seleziona** correttamente la totalità delle risposte (9-10 su 10). **Espone** il contenuto ricostruendo l'intera vicenda con ordine logico e cronologico. **Utilizza** un lessico specifico (es: nomi propri dei personaggi, termini tecnici del testo informativo) e connettivi vari (es: *inizialmente, in seguito, infine*). Procede in totale autonomia senza interruzioni.
Intermedio	**Individua** la maggior parte delle risposte (7-8 su 10). **Organizza** l'esposizione orale seguendo i nodi principali della trama (es: inizio, vicenda, conclusione) senza saltare passaggi chiave. **Utilizza** frasi corrette ma con un lessico comune. **Ricostruisce** i fatti con minimi aiuti (es: un solo invito a proseguire).
Base	**Riconosce** solo le informazioni esplicite e letterali (5-6 su 10 risposte). **Espone** il contenuto limitandosi ai fatti essenziali, talvolta perdendo il filo logico. **Utilizza** una sintassi semplice e ripetitiva (es: uso continuo di *"...e poi... e poi..."*). **Necessita** di domande-guida per completare l'esposizione (es: *"Cosa succede dopo l'incontro con...?"*).
In via di prima acquisizione	**Individua** solo alcune informazioni isolate (3-4 su 10 risposte). **Riferisce** solo brevi frammenti del testo ascoltato, spesso non collegati tra loro (es: ricorda un dettaglio curioso ma non il senso della storia). **Si esprime** con frasi incomplete. **Agisce** solo se sollecitato costantemente e guidato passo-passo dall'insegnante.
Non ancora acquisito	**Non seleziona** risposte corrette o risponde a caso (0-2 su 10). **Non espone** il contenuto o riferisce elementi estranei al testo ascoltato. **Mostra** estrema difficoltà nel mantenere l'attenzione durante l'ascolto. Anche se guidato, non riesce a ricostruire alcun segmento della prova.

CONCLUSIONI

La creazione di questa rubrica è frutto di un processo di ingegneria pedagogica guidato dal docente, seguendo le fasi di progettazione, implementazione e affinamento

delineate da Lee Boonstra (2024). La Gen-AI ha supportato la definizione di descrittori concreti e misurabili, ma senza l'intervento del docente, capace di selezionare obiettivi chiari, verbi e livelli adeguati alla maturità cognitiva degli studenti, i risultati sarebbero rimasti generici e privi di valore educativo. In questo contesto, la Gen-AI agisce come assistente intelligente, potenziando precisione e rigore, mentre il ruolo dell'insegnante resta centrale: la formazione dei docenti diventa quindi essenziale non solo per utilizzare la tecnologia, ma per guidarla criticamente, validare i risultati e garantire che ogni strumento digitale contribuisca effettivamente al successo formativo degli studenti. Va sottolineato, tuttavia, che la rubrica richiede ulteriori fasi di validazione: occorre testarne l'affidabilità e la validità di contenuto. Solo attraverso questi passaggi la rubrica può diventare uno strumento realmente efficace per sostenere e monitorare l'apprendimento.

BIBLIOGRAFIA

Akgun S., Greenhow C. (2021), "Artificial intelligence in education: Addressing ethical challenges in K-12 settings", *AI and Ethics*, 2, pp. 431–440

Alsajri A., Salaman H.A., Steiti A.M. (2024), "Generative Models in Natural Language Processing: A Comparative Study of ChatGPT and Gemini", *Babylonian Journal of Artificial Intelligence*, pp. 134–145

Boonstra L. (2024), *Prompt Engineering*, Google Cloud / Kaggle

Bower M., Torrington J., Lai J.W.M., Petocz P. (2024), Alfano M., "How should we change teaching and assessment in response to increasingly powerful generative Artificial Intelligence? Outcomes of the ChatGPT teacher survey", "Education and Information Technologies", 29(12), pp. 15403–15439

Castoldi M. (2009), *Valutare le competenze. Percorsi e strumenti*, Roma, Carocci

D'Ugo R., Montanari C., Salvucci M. (2024), "Imparare a valutare per insegnare a valutarsi", *Pedagogia più Didattica*, 10(1)

D'Ugo R. (2012), *La valutazione degli apprendimenti*, in M. Baldacci, *L'insegnamento nella scuola secondaria*, Napoli, Tecnodid, pp. 99-120

De Martino D., Dicataldo M.C., Tinterri A., Dipace A. (2025), "Assessment, ethics and Artificial Intelligence", *Journal of Inclusive Methodology and Technology in Learning and Teaching*, 5(1)

Di Padova M., Tinterri A., Basta A., Amatori G., Dipace A. (2024), "Potenziare il giudizio descrittivo nella scuola primaria con l'uso dell'IA generativa", *IUL Research*, 5(9)

Floridi L. (2022), *L'etica dell'intelligenza artificiale. Sviluppi, opportunità e sfide*, Milano, Raffaello Cortina Editore

Herman J., Aschbacher P., Winters L. (1992), *A Practical Guide to Alternative Assessment*, ASCD

Isidori M.V., Muccini H., Evangelista C. (2025), "Generative intelligence: a possible redefinition of teaching and formative assessment. A survey among teacher trainees", *Giornale italiano di educazione alla salute, sport e didattica inclusiva*, 9(1)

Jana B., Sarkar P., Lepcha P., Topno P. (2025), *Pedagogy to Cybergogy - Education 5.0*, LAP Lambert Academic Publishing

McTighe J., Ferrare S. (1996), *Performance-based assessment in the classroom: A planning framework*, in R.E. Blum, J.A. Arter (eds), *A handbook for student performance assessment in an era of restructuring*, Association for Supervision and Curriculum Development, Alexandria

Ministero dell'Istruzione e del Merito (2025), *Linee guida per l'introduzione dell'intelligenza artificiale nelle istituzioni scolastiche* (Decreto Ministeriale n. 166, 9 agosto 2025)

Mousa M. (2025), "AI-Supported Formative Assessments: Enhancing Student-Centered Learning and Teacher Perceptions", *Journal of Pedagogy and Education Science*, 4(2), pp. 127-141

Nguyen D.L., Le P.T.T., Le T.T. (2025), "Using Gemini for formative assessment in English academic writing: Critical insights into the AI tool's efficacy", *AsiaCALL Online Journal*, 16(1), pp. 328-343

Rossi F. (2024), *L'intelligenza artificiale. Come funziona e dove ci porta la tecnologia che sta trasformando il mondo*, Bari, Laterza

Sengar S.S., Bin Hassan A., Kumar S., Carrol F. (2025), "Generative artificial intelligence: a systematic review and applications", *Multimed Tools Appl*, 84, pp. 23661–23700

Thorpe J. (2024), *AI for Students. Creative hacks for academic success an illustrated guide*, Promptly Books

Vuorikari R., Kluzer S., Punie Y. (2022), *DigComp 2.2: The Digital Competence Framework for Citizens*, Publications Office of the European Union, Luxembourg, pp. 1-128