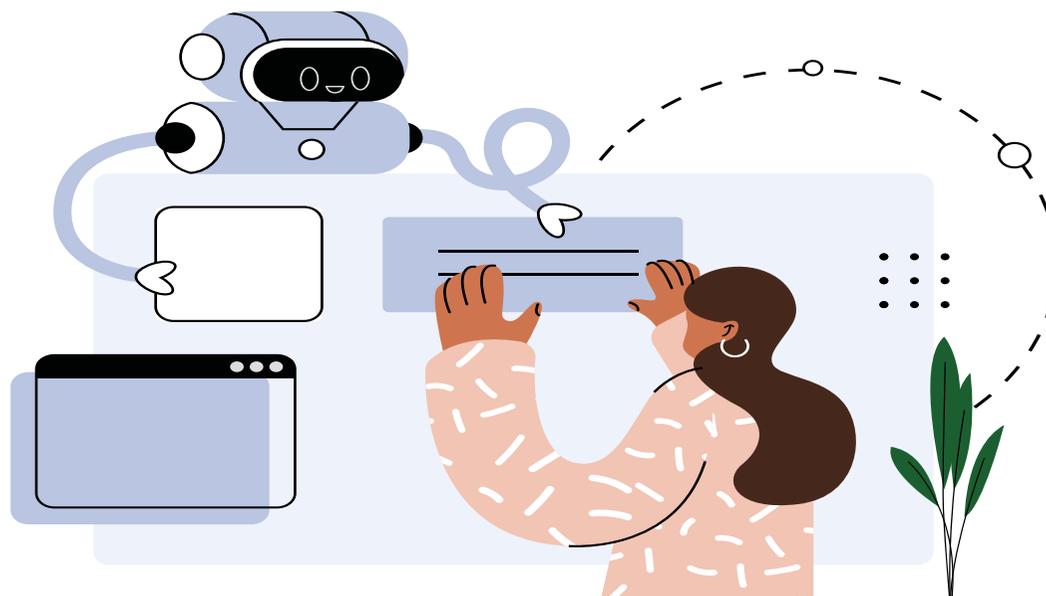


Guida all'utilizzo dell'IA nell'istruzione e nella ricerca



INTRODUZIONE

La pubblicazione di ChatGPT alla fine del 2022, il primo strumento di intelligenza artificiale generativa (GenIA) facile da usare e reso ampiamente disponibile al pubblico, ha provocato scompiglio ovunque e in tutti i settori, e sta alimentando la competizione tra le grandi aziende tecnologiche per posizionarsi nel campo dello sviluppo dei modelli GenIA.

In tutto il mondo, la preoccupazione iniziale nell'ambito dell'istruzione era che ChatGPT, e strumenti simili di GenIA, sarebbero stati utilizzati dagli studenti per imbrogliare nei loro compiti, minando così il valore della valutazione dell'apprendimento, delle certificazioni e delle qualifiche (Anders, 2023). Mentre alcune istituzioni educative hanno vietato l'uso di ChatGPT, altre hanno accolto con cautela l'arrivo di GenIA (Tlili, 2023). Molte scuole e università, ad esempio, hanno adottato un approccio progressista, ritenendo che "anziché cercare di vietarne l'uso, gli studenti e il personale debbano essere sostenuti nell'uso etico, efficace e trasparente degli strumenti GenIA" (Russell Group). Questo approccio riconosce che GenIA è ampiamente diffuso, è probabile che diventi sempre più sofisticato e ha delle potenzialità specifiche per il settore dell'istruzione.

GenIA vanta una miriade di possibili utilizzi: può automatizzare l'elaborazione delle informazioni e la presentazione dei risultati in tutte le rappresentazioni simboliche chiave del pensiero umano; consente la consegna di risultati finali fornendo prodotti di conoscenza semilavorati. Liberando gli esseri umani da alcune categorie di abilità di pensiero di ordine inferiore, questa nuova generazione di strumenti AI potrebbe avere profonde implicazioni su come consideriamo l'intelligenza umana e le modalità di apprendimento.

Ma GenIA solleva anche molte preoccupazioni immediate legate a questioni come la sicurezza, la privacy dei dati, i diritti d'autore e la manipolazione. Alcune di queste sono rischi più ampi legati all'intelligenza artificiale che sono stati ulteriormente esacerbati da GenIA, mentre altri sono emersi con questa ultima generazione di strumenti. È ora necessario che ciascuna di queste questioni e preoccupazioni sia pienamente compresa e affrontata.

La Guida che state leggendo è progettata per rispondere a questa urgenza. Tuttavia, un insieme tematico di orientamenti su GenIA per l'istruzione non dovrebbe essere inteso come una rivendicazione che GenIA sia la soluzione alle sfide fondamentali dell'istruzione. Nonostante l'esagerazione mediatica, è improbabile che GenIA da sola risolva i problemi che affliggono i sistemi educativi di tutto il mondo. Nel rispondere alle questioni educative di lunga data, è fondamentale sostenere l'idea che la capacità umana e l'azione collettiva, e non la tecnologia, siano il fattore determinante nelle soluzioni efficaci alle sfide fondamentali affrontate dalle società.

Pertanto, questa Guida mira a sostenere la pianificazione di regolamenti appropriati, politiche e programmi di sviluppo delle capacità umane, per garantire che GenIA diventi uno strumento che beneficia e responsabilizza davvero insegnanti, studenti e ricercatori. Basandosi sulla Raccomandazione dell'UNESCO sull'Etica dell'Intelligenza Artificiale, la Guida si basa su un approccio centrato sull'essere umano, che promuove la capacità umana di agire e prendere decisioni, l'inclusione, l'equità, la parità di genere, la diversità culturale e linguistica, nonché il pluralismo delle opinioni e delle espressioni.

La Guida esamina innanzitutto cosa sia GenIA e come funziona, presentando le diverse tecnologie e modelli disponibili (Sezione 1), prima di identificare una serie di controversie etiche e questioni politiche legate sia all'IA in generale che a GenIA in modo specifico (Sezione 2). Segue una discussione dei passi e degli elementi chiave da esaminare quando si cerca di regolamentare GenIA basandosi su un approccio centrato sull'essere umano - che garantisce un uso etico, sicuro, equo e significativo (Sezione 3). La Sezione 4 propone misure che possono essere adottate per sviluppare quadri politici coerenti e completi per regolare l'uso di GenIA nell'istruzione e nella ricerca, mentre la Sezione 5 esamina le possibilità di utilizzo creativo di GenIA nella progettazione dei programmi, nell'insegnamento, nell'apprendimento e nelle attività di ricerca. La Sezione 6, infine, offre considerazioni sulle implicazioni a lungo termine di GenIA per l'istruzione e la ricerca.

1. Cos'è l'AI generativa e come funziona?

1.1 Cos'è l'Intelligenza Artificiale Generativa?

L'Intelligenza Artificiale Generativa (GenIA) è una tecnologia di Intelligenza Artificiale (IA) che genera automaticamente contenuti in risposta a prompt scritti in interfacce di conversazione in linguaggio naturale. Aniché offrire semplicemente pagine web, attingendo da contenuti preesistenti, GenIA produce effettivamente nuovi contenuti. I contenuti possono apparire in formati che comprendono tutte le rappresentazioni simboliche del pensiero umano: testi scritti in linguaggio naturale, immagini (da fotografie a dipinti digitali e cartoni animati), video, musica e codice software. GenIA viene addestrato utilizzando dati raccolti da pagine web, conversazioni sui social media e altri media online. Genera i suoi contenuti analizzando statisticamente le distribuzioni di parole, pixel o altri elementi nei dati che ha elaborato, identificando e ripetendo schemi comuni (ad esempio, parole seguono tipicamente altre parole).

Sebbene GenIA possa produrre nuovi contenuti, non può generare nuove idee o soluzioni per le sfide del mondo reale, poiché non comprende oggetti del mondo reale o le relazioni sociali che alimentano il linguaggio. Inoltre, nonostante la sua produzione scorrevole e impressionante, GenIA non può essere considerato sempre affidabile in termini di precisione: persino il produttore di ChatGPT riconosce che *"Sebbene strumenti come ChatGPT possano spesso generare risposte che sembrano ragionevoli, non si può fare affidamento sulla loro accuratezza"* (OpenAI, 2023). Nella maggior parte dei casi, gli errori passeranno inosservati a meno che l'utente non abbia una conoscenza solida dell'argomento in questione.

1.2 Come funziona l'Intelligenza Artificiale Generativa?

Le tecnologie specifiche alla base di GenIA fanno parte della famiglia di tecnologie di IA chiamate Apprendimento Automatico (ML), che utilizzano algoritmi per consentirle di migliorare continuamente e automaticamente le sue prestazioni traendo conoscenze dai dati immagazzinati. Il tipo di ML che ha portato a molti dei progressi in AI che abbiamo visto negli ultimi anni, come l'uso dell'AI per il riconoscimento facciale, è noto come Reti Neurali Artificiali (ANNs), che si ispirano al funzionamento del cervello umano e alle sue connessioni sinaptiche tra neuroni. Esistono molti tipi di ANNs.

Sia le tecnologie di generazione di testo che di immagini di GenIA si basano su un insieme di tecnologie di IA che sono state a disposizione dei ricercatori per diversi anni. Ad esempio, ChatGPT utilizza un Generative Pre-trained Transformer (GPT), mentre il GenIA per le immagini utilizza tipicamente ciò che sono noti come Generative Adversarial Networks (GAN) (vedi Tabella 1).

TABELLA 1

Machine learning (ML)	Un tipo di IA che utilizza i dati per migliorare automaticamente le proprie prestazioni.	
Artificial Neural Network (ANN)	Un tipo di apprendimento automatico ispirato alla struttura e al funzionamento del cervello umano (ad esempio, le connessioni sinaptiche tra i neuroni).	
IA generatrice di testo	General-Purpose Transformers	Un tipo di rete neurale artificiale (ANN) in grado di concentrarsi su diverse parti dei dati per determinare come esse siano correlate tra loro.
	Large Language Models (LLM)	Un tipo di General-purpose Transformer che viene addestrato su vaste quantità di dati testuali

	Generative Pre-trained Transformer (GPT)	Un tipo di LLM (Large Language Model) che è pre-addestrato su quantità ancora più ampie di dati, il che consente al modello di catturare le sfumature del linguaggio e generare testo coerente e consapevole del contesto
IA generatrice di immagini	Generative adversarial networks (GANs)	Tipi di reti neurali utilizzate per la generazione di immagini.
	Variational Autoencoders (VAEs)	

1.2.1. Come funzionano i modelli di testo GenIA

L'Intelligenza Artificiale Generativa basata su testo utilizza un tipo di Rete Neurale Artificiale (ANN) conosciuto come un Transformer ad uso generale e un tipo di Transformer ad uso generale chiamato Large Language Model. È per questo motivo che i sistemi GenIA basati su testo sono spesso indicati come Large Language Models, o LLM. Il tipo di LLM utilizzato dai modelli GenIA basati su testo è noto come Generative Pre-trained Transformer, o GPT (da qui il 'GPT' in 'ChatGPT').

ChatGPT è basato su GPT-3, che è stato sviluppato da OpenAI. Questa è stata la terza generazione del loro GPT, la prima è stata lanciata nel 2018 e la più recente, GPT-4, a marzo 2023 (vedi Tabella 2). Ogni generazione del GPT di OpenAI è migliorata progressivamente rispetto alla precedente attraverso progressi nelle architetture di intelligenza artificiale, nei metodi di addestramento e nelle tecniche di ottimizzazione. Un aspetto ben noto del suo progresso continuo è l'uso di crescenti quantità di dati per addestrare il numero esponenziale di "parametri" metaforicamente intesi come manopole che possono essere regolate per perfezionare le prestazioni del GPT. Questi parametri includono i "pesi" del modello, parametri numerici che determinano come il modello elabora l'input e produce l'output.

Oltre ai progressi nell'ottimizzazione delle architetture di intelligenza artificiale e dei metodi di addestramento, questo rapido sviluppo è stata reso possibile anche grazie alle enormi quantità di dati e ai miglioramenti nelle capacità di calcolo disponibili alle grandi aziende. Dal 2012, le capacità di calcolo utilizzate per addestrare i modelli GenIA si sono raddoppiate ogni 3-4 mesi. A titolo di confronto, la Legge di Moore aveva un periodo di raddoppio di due anni (OpenAI, 2018; Università di Stanford, 2019).

TABELLA 2

Modello	Data di lancio	Quantità di dati di addestramento	Numero di parametri	Caratteristiche
GPT-1	2018	40GB	117 milioni	Capace di compiti di elaborazione del linguaggio naturale come il completamento di testi e la risposta a domande
GPT-2	2019	40GB	1,500 milioni	Capace di compiti di elaborazione del linguaggio naturale più complessi, come la traduzione automatica e la sintesi

GPT-3	2020	17,000GB	175,000 milioni	In grado di compiere compiti avanzati di elaborazione del linguaggio naturale, come la scrittura di paragrafi coerenti e la generazione di interi articoli. È inoltre in grado di adattarsi a nuovi compiti con solo alcuni esempi.
GPT-4	2023	1,000,000GB	170,000,000 milioni	Maggiore affidabilità e in grado di elaborare istruzioni più complesse.

Una volta che il GPT è stato addestrato, la generazione di una risposta testuale a un prompt coinvolge i seguenti passaggi:

1. Il prompt viene suddiviso in unità più piccole (chiamate token) che vengono inserite nel GPT.
2. Il GPT utilizza modelli statistici per prevedere parole o frasi probabili che potrebbero formare una risposta coerente al prompt.
 - a. Il GPT identifica modelli di parole e frasi che comunemente co-occorrono nel suo grande modello di dati pre-costruito (che comprende testo estratto da Internet e altrove).
 - b. Utilizzando questi modelli, il GPT stima la probabilità di specifiche parole o frasi che appaiono in un determinato contesto.
 - c. Iniziando con una previsione casuale, il GPT utilizza queste probabilità stimate per prevedere la prossima parola o frase più probabile nella sua risposta.
3. Le parole o frasi previste vengono convertite in testo leggibile.
4. Il testo leggibile viene filtrato attraverso ciò che sono conosciuti come "guardrail" per rimuovere eventuali contenuti offensivi.
5. I passaggi dal 2 al 4 vengono ripetuti fino a quando una risposta è completata. La risposta è considerata completata quando raggiunge un limite massimo di token o soddisfa criteri di arresto predefiniti.
6. La risposta viene post-elaborata per migliorare la leggibilità, applicando formattazione, punteggiatura e altre migliorie (come iniziare la risposta con parole che potrebbe usare un essere umano, come 'Certamente', 'Certo' o 'Mi dispiace').

Mentre i GPT e la loro capacità di generare automaticamente testo sono stati disponibili per i ricercatori dal 2018, ciò che ha reso così innovativo il lancio di ChatGPT è stata la sua accessibilità gratuita tramite un'interfaccia facile da usare: il che significa che chiunque con accesso a Internet poteva esplorare lo strumento. Il lancio di ChatGPT ha scatenato forte interesse e reazioni in tutto il mondo e ha rapidamente portato altre aziende tecnologiche globali a cercare di rimanere al passo, insieme a numerose start-up, sia lanciando i propri sistemi simili sia costruendo nuovi strumenti basati su di essi.

Ecco alcune delle alternative a ChatGPT rilasciate entro luglio 2023:

- Alpaca: Una versione perfezionata di Meta's Llama, dalla Stanford University, che mira a affrontare le informazioni false, gli stereotipi sociali e il linguaggio tossico dei LLM.
- Bard: Un LLM di Google, basato sui suoi sistemi LaMDA e PaLM 2, che ha accesso a Internet in tempo reale, il che significa che può fornire informazioni aggiornate.
- Chatsonic: Realizzato da Writesonic, si basa su ChatGPT e raccoglie dati direttamente.
- Ernie (noto anche come Wenxin Yiyan 文心一言): Un LLM bilingue di Baidu, ancora in sviluppo, che integra una vasta conoscenza con enormi dataset per generare testi e immagini.

- Hugging Chat: Realizzato da HuggingFace, che ha enfatizzato l'etica e la trasparenza durante lo sviluppo, l'addestramento e la distribuzione. Inoltre, tutti i dati utilizzati per addestrare i loro modelli sono open source.
- Jasper: Una serie di strumenti e API che, ad esempio, possono essere addestrati a scrivere nello stile preferito dell'utente. Può anche generare immagini.
- Llama: Un LLM open source di Meta che richiede meno potenza di calcolo e meno risorse per testare nuovi approcci, convalidare il lavoro degli altri ed esplorare nuovi casi d'uso.
- Open Assistant: Un approccio open source progettato per consentire a chiunque abbia sufficiente esperienza di sviluppare il proprio LLM. È stato costruito su dati di addestramento curati da volontari.
- Tongyi Qianwen (通义千问): Un LLM di Alibaba che può rispondere a prompt in inglese o cinese. Sta venendo integrato nella suite di strumenti aziendali di Alibaba.
- YouChat: Un LLM che incorpora capacità di ricerca in tempo reale per fornire contesto e informazioni aggiuntive al fine di generare risultati più precisi e affidabili.

La maggior parte di questi strumenti è gratuita (entro certi limiti), mentre alcuni sono open source. Molti altri prodotti, basati su uno di questi LLM, stanno per essere lanciati. Ecco alcuni esempi:

- ChatPDF: Riassume e risponde alle domande su documenti PDF inviati.
- Elicit: The AI Research Assistant: Mira ad automatizzare parti del flusso di lavoro dei ricercatori, identificando documenti rilevanti e riassumendo informazioni chiave.
- Perplexity: Fornisce un "hub di conoscenza" per persone che cercano risposte rapide e accurate su misura per le loro esigenze.

Allo stesso modo, gli strumenti basati su LLM vengono incorporati in altri prodotti, come i browser web. Ad esempio, le estensioni per il browser Chrome basate su ChatGPT includono:

- WebChatGPT: Concede a ChatGPT l'accesso a Internet per consentire conversazioni più accurate e aggiornate.
- Compose AI: Autocompleta frasi in email e altrove.
- TeamSmart AI: Fornisce un "team di assistenti virtuali".
- Wiseone: Semplifica le informazioni online.

Inoltre, ChatGPT è stato incorporato in alcuni motori di ricerca ed è in fase di implementazione in ampi portafogli di strumenti di produttività (ad esempio Microsoft Word ed Excel), rendendolo ancora più disponibile in uffici e istituti scolastici di tutto il mondo (Murphy Kelly, 2023).

Infine, come esempio interessante del passaggio alla GenIA per le immagini, il più recente GPT di OpenAI, il GPT-4, è in grado di accettare immagini oltre al testo nei suoi prompt. In questo senso, è multimodale. Di conseguenza, alcuni sostengono che il nome "Large Language Model" (LLM) stia diventando meno appropriato, motivo per cui i ricercatori dell'Università di Stanford hanno proposto il termine "foundation model" (modello fondamentale) (Bommasani et al., 2021). Questa alternativa non è però ancora particolarmente diffusa.

1.2.2. Come funzionano i modelli di GenIA per le immagini

I modelli di GenIA per le immagini e la musica utilizzano tipicamente un tipo diverso di Rete Neurale Artificiale (ANN) noto come Generative Adversarial Networks (GANs), che può anche essere combinato con Variational Autoencoders. I GAN hanno due parti (due "avversari"), il "generatore" e il "discriminatore". Nel caso dei GAN per le immagini, il generatore crea un'immagine casuale in risposta a un prompt, e il discriminatore cerca di distinguere tra questa immagine generata e le immagini reali. Il generatore utilizza quindi il risultato del discriminatore per regolare i suoi parametri, al fine di creare un'altra immagine. Il processo viene ripetuto, possibilmente migliaia di volte, con il generatore che crea immagini sempre più realistiche che il discriminatore è sempre meno in grado di distinguere dalle immagini reali. Ad esempio, un GAN di successo addestrato su un dataset di migliaia di fotografie di paesaggi potrebbe generare nuove ma irreali immagini di paesaggi che sono quasi indistinguibili dalle fotografie reali. Nel frattempo, un GAN addestrato su un dataset di musica popolare (o persino di un singolo artista) potrebbe generare nuovi pezzi di musica che seguono la struttura e la complessità della musica originale.

A luglio 2023, i modelli di GenIA per le immagini disponibili sono tutti in grado di generare immagini da prompt di testo. La maggior parte di essi è gratuita, sia pure entro certi limiti:

- Craiyon: In precedenza noto come DALL • E mini.
- DALL•E 2: Lo strumento di GenIA per le immagini di OpenAI.
- DreamStudio: Lo strumento di GenIA per le immagini di Stable Diffusion.
- Fotor: Incorpora GenIA in una gamma di strumenti di modifica delle immagini.
- Midjourney: Uno strumento di GenIA per le immagini indipendente.
- NightCafe: Interfaccia a Stable Diffusion e DALL•E 2.
- Photosonic: Generatore di arte AI di WriteSonic.

Ecco alcuni esempi di **GenIA video** di facile accesso:

- Elai: Può convertire presentazioni, siti web e testo in video.
- GliaCloud: Può generare video da contenuti di notizie, post sui social media, eventi sportivi in diretta e dati statistici.
- Pictory: Può creare automaticamente video brevi da contenuti a lunga durata.
- Runway: Offre una gamma di strumenti di generazione e modifica video (e di immagini).

Infine, questi sono alcuni esempi di GenIA per la **musica** di facile accesso:

- Aiva: Può creare automaticamente colonne sonore personalizzate.
- Boomy, Soundraw e Voicemod: Possono generare canzoni da qualsiasi testo e non richiedono conoscenze di composizione musicale.

1.3 Progettazione di prompt per generare output desiderati

Anche se l'uso di GenIA può essere facile come digitare una domanda, la realtà è che per l'utente non è ancora semplice ottenere esattamente l'output desiderato. Ad esempio, la rivoluzionaria immagine AI "Théâtre D'opéra Spatial", che ha vinto un premio alla Colorado State Fair negli Stati Uniti, ha richiesto settimane di scrittura di prompt e il perfezionamento di centinaia di immagini per generare la presentazione finale (Roose, 2022). La sfida simile di scrivere prompt efficaci per il GenIA testuale ha portato all'aumento di un numero crescente di posizioni di "prompt-engineering" sui siti di

reclutamento (Popli, 2023). Il termine "Prompt-engineering" si riferisce ai processi e alle tecniche per comporre un input al fine di produrre un output GenIA che soddisfi pienamente le aspettative dell'utente.

Il prompt-engineering ha successo quando il prompt articola una catena coerente di ragionamento.

Concentrarsi su un problema particolare o su una catena di pensiero in un ordine logico è essenziale per il prompt-engineering. Ecco alcune raccomandazioni:

- Utilizzare un linguaggio semplice, chiaro e diretto che possa essere facilmente compreso, evitando un linguaggio complesso o ambiguo.
- Includere esempi per illustrare la risposta desiderata.
- Includere il contesto, che è cruciale per generare risposte rilevanti e significative.
- Rifinire e iterare il prompt secondo necessità, sperimentando con diverse variazioni.
- Avere un approccio etico, evitando prompt che potrebbero generare contenuti inappropriati, tendenziosi o dannosi.

È importante anche riconoscere immediatamente che gli output del GenIA non possono essere considerati affidabili senza una valutazione critica. Come OpenAI scrive riguardo al loro GPT-4 più sofisticato: *"Nonostante le sue capacità, il GPT-4 ha limitazioni simili ai modelli GPT precedenti. In particolare, non è ancora completamente affidabile (esso 'allucina' fatti e commette errori di ragionamento). Bisogna prestare grande attenzione nell'utilizzare gli output dei modelli linguistici, specialmente in contesti ad alto rischio, con il protocollo esatto (come la revisione umana, il supporto con contesto aggiuntivo o l'evitare usi ad alto rischio del tutto) che si adatta alle esigenze di un caso d'uso specifico."*

Date le qualità degli output del GenIA, è fondamentale condurre rigorosi test utente e valutazioni delle performance prima di convalidare gli strumenti per l'adozione su larga scala o in contesti ad alto rischio. Tali esercizi dovrebbero essere progettati con una metrica di performance più rilevante per il tipo di compito per il quale gli utenti chiedono al GenIA di fornire output. Ad esempio, per risolvere problemi matematici, la "precisione" potrebbe essere utilizzata come l'elemento principale per quantificare con quanta frequenza uno strumento GenIA produce la risposta corretta; per rispondere a domande sensibili, l'elemento principale per misurare le performance potrebbe essere "tasso di risposta" (la frequenza con cui il GenIA risponde direttamente a una domanda); per la generazione di codice, l'elemento principale potrebbe essere "la frazione dei codici generati che sono direttamente eseguibili" (ovvero se il codice generato potrebbe essere eseguito direttamente in un ambiente di programmazione e superare i test di unità); per il ragionamento visivo, l'elemento principale potrebbe essere "corrispondenza esatta" (ovvero se gli oggetti visivi generati corrispondono esattamente alla verità di riferimento) (Chen, Zaharia e Zou, 2023).

In breve, a un livello superficiale, il GenIA è facile da usare; tuttavia, gli output più sofisticati richiedono competenza umana e devono essere valutati in modo critico prima di essere utilizzati.

Implicazioni per l'istruzione e la ricerca

Sebbene GenIA possa aiutare insegnanti e ricercatori a generare testi utili e altri output per sostenere il loro lavoro, non si tratta necessariamente di un processo immediato. Possono essere necessarie diverse interazioni di un prompt prima di ottenere l'output desiderato. Una preoccupazione è che gli studenti, in quanto per definizione meno esperti dei docenti, potrebbero accettare senza rendersene conto e senza un coinvolgimento critico un output GenIA superficiale, inaccurato o addirittura dannoso.

1.4 Emergenza di EdGPT e le sue implicazioni

Dato che i modelli GenIA possono servire da base o punto di partenza per lo sviluppo di modelli più specializzati o specifici per un determinato dominio, alcuni ricercatori hanno suggerito che i GPT dovrebbero essere ribattezzati "modelli fondamentali" (Bommasani et al., 2021). Nell'ambito dell'istruzione, sviluppatori e ricercatori hanno iniziato a perfezionare un modello fondamentale per sviluppare l' "EdGPT". I modelli EdGPT sono addestrati con dati specifici per scopi educativi. In altre parole, l'EdGPT mira a perfezionare il modello che è stato derivato da grandi quantità di dati di addestramento generali con quantità minori di dati di alta qualità specifici per un determinato dominio educativo.

Ciò potrebbe dare all'EdGPT una maggiore capacità di supportare il raggiungimento delle trasformazioni elencate nella Sezione 4.3. Ad esempio, i modelli EdGPT mirati alla progettazione curricolare possono consentire ai docenti e agli studenti di generare materiali didattici appropriati come piani di lezione, quiz e attività interattive che si allineano strettamente con un approccio pedagogico efficace e obiettivi curriculari specifici e livelli di sfida per determinati studenti. Allo stesso modo, nel contesto di un tutor per le competenze linguistiche individuali, un modello fondamentale perfezionato con testi appropriati per una lingua particolare potrebbe essere utilizzato per generare frasi esemplari, paragrafi o conversazioni per l'esercizio. Quando gli studenti interagiscono con il modello, esso può rispondere con testo pertinente e grammaticalmente corretto al livello adeguato a loro. In teoria, gli output dei modelli EdGPT potrebbero anche contenere meno pregiudizi generali o contenuti altrimenti contestabili rispetto a GPT standard, ma potrebbero comunque generare errori. È fondamentale notare che, a meno che i modelli GenIA sottostanti e l'approccio non cambino significativamente, EdGPT potrebbe comunque generare errori e avere limitazioni in altri modi, come suggerimenti per piani di lezione o strategie didattiche. Di conseguenza, è comunque importante che gli utenti principali di EdGPT, in particolare insegnanti e studenti, adottino una prospettiva critica su qualsiasi output.

Attualmente, il perfezionamento dei modelli fondamentali per un uso più mirato di GPT nell'istruzione è in una fase iniziale. Gli esempi esistenti includono EduChat, un modello fondamentale sviluppato dall'Università Normale dell'Est della Cina per fornire servizi per l'insegnamento e l'apprendimento, e i cui codici, dati e parametri sono condivisi come open source. Un altro esempio è MathGPT sviluppato dal TAL Education Group - un LLM che si concentra sulla risoluzione di problemi matematici e sulla lectio magistralis per utenti in tutto il mondo.

Tuttavia, prima che si possa compiere un progresso significativo, è essenziale dedicare sforzi al perfezionamento dei modelli fondamentali non solo attraverso l'aggiunta di conoscenze specifiche del settore e la rimozione di pregiudizi, ma anche attraverso l'aggiunta di conoscenze sui metodi di apprendimento rilevanti e su come questi possono essere riflessi nella progettazione di algoritmi e modelli. La sfida consiste nel determinare fino a che punto i modelli EdGPT possono andare oltre la conoscenza del settore per mirare anche a una pedagogia centrata sullo studente e a interazioni positive tra insegnante e studente. La sfida ulteriore è determinare fino a che punto i dati degli studenti e degli insegnanti possano essere raccolti ed utilizzati eticamente per informare un EdGPT. Infine, è necessaria una robusta ricerca per garantire che EdGPT non comprometta i diritti umani degli studenti né sminuisca il ruolo degli insegnanti.

2. Controversie legate all'intelligenza artificiale generativa (GenIA) e le loro implicazioni per l'istruzione

Dopo aver discusso in precedenza cosa sia la GenIA e come funzioni, questa sezione esamina le controversie e i rischi etici sollevati da tutti i sistemi GenIA e considera alcune delle implicazioni per il settore dell'istruzione.

2.1 Peggioramento della povertà digitale

Come già detto in precedenza, la GenIA si basa su enormi quantità di dati e una potenza di calcolo massiccia, oltre che sulle sue innovative iterazioni nelle architetture dell'IA e nei metodi di addestramento, che sono principalmente disponibili solo alle più grandi aziende tecnologiche internazionali e a poche economie (principalmente gli Stati Uniti, la Repubblica Popolare Cinese e, in misura minore, l'Europa). Ciò significa che la possibilità di creare e controllare la GenIA è fuori dalla portata della maggior parte delle aziende e dei Paesi, specialmente quelli del Sud del mondo.

Poiché l'accesso ai dati diventa sempre più essenziale per lo sviluppo economico dei Paesi e per le opportunità digitali delle persone, quei Paesi e quelle persone che non hanno accesso o non possono permettersi dati sufficienti si trovano in una situazione di "povertà di dati" (Marwala, 2023). La situazione è simile per l'accesso alla potenza di calcolo. La rapida diffusione della GenIA nei Paesi e nelle regioni tecnologicamente avanzate ha accelerato esponenzialmente la generazione e l'elaborazione dei dati e ha contemporaneamente intensificato la concentrazione della ricchezza legata all'IA nel Nord del mondo. Come conseguenza immediata, le regioni povere di dati sono state ulteriormente escluse e messe a rischio a lungo termine di essere colonizzate dagli standard incorporati nei modelli GPT. Gli attuali modelli ChatGPT sono addestrati su dati di utenti online che riflettono i valori e le norme del Nord del mondo, rendendoli inadatti per algoritmi AI localmente rilevanti nelle comunità povere di dati in molte parti del Sud del mondo o nelle comunità più svantaggiate del Nord del mondo.

Implicazioni per l'istruzione e la ricerca

I ricercatori, gli insegnanti e gli studenti dovrebbero adottare una visione critica delle orientazioni valoriali, degli standard culturali e delle consuetudini sociali incorporate nei modelli di addestramento GenIA. I decisori politici dovrebbero essere consapevoli e intraprendere azioni per affrontare il peggioramento delle disuguaglianze causate dall'ampia divisione nell'addestramento e nel controllo dei modelli GenIA.

2.2 Superamento dell'adeguamento alla regolamentazione nazionale

I fornitori principali di GenIA sono stati anche criticati perché non consentono che i loro sistemi siano soggetti a una rigorosa revisione accademica indipendente (Dwivedi et al., 2023). Le tecnologie fondamentali di una società GenIA tendono ad essere protette come proprietà intellettuale aziendale. Nel frattempo, molte delle aziende che stanno iniziando a utilizzare la GenIA stanno trovando sempre più difficile mantenere la sicurezza dei loro sistemi (Lin, 2023). Inoltre, nonostante le richieste di regolamentazione provenienti dall'industria dell'IA stessa, la stesura di legislazioni sulla creazione e l'uso di tutta l'IA, compresa la GenIA, spesso rimane indietro rispetto al rapido sviluppo. Questo spiega in parte le sfide incontrate dagli enti nazionali o locali nel comprendere e governare le questioni legali ed etiche.

Sebbene la GenIA possa potenziare le capacità umane nel completare determinati compiti, vi è un limitato controllo democratico sulle aziende che promuovono la GenIA. Questo solleva la questione della regolamentazione, in particolare per quanto riguarda l'accesso e l'uso dei dati nazionali, compresi i dati sulle istituzioni locali e sugli individui, nonché i dati generati sul territorio dei paesi. È necessaria una legislazione appropriata affinché le agenzie governative locali possano acquisire un certo controllo sulle crescenti onde di GenIA per garantirne la governance come bene pubblico.

Implicazioni per l'istruzione e la ricerca

I ricercatori, gli insegnanti e gli studenti dovrebbero essere consapevoli della mancanza di regolamentazioni adeguate per proteggere la proprietà delle istituzioni e degli individui nazionali e i diritti degli utenti nazionali della GenIA, nonché per affrontare le questioni legislative generate dalla GenIA.

2.3 Utilizzo di contenuti senza consenso

Come già notato in precedenza, i modelli GenIA sono costruiti utilizzando grandi quantità di dati (ad esempio testi, suoni, codice e immagini) spesso raccolti da Internet e di solito senza il permesso dei proprietari. Di conseguenza, molti sistemi di GenIA per le immagini e alcuni sistemi di GenIA per il codice sono stati accusati di violare i diritti di proprietà intellettuale. Attualmente ci sono diversi casi legali internazionali che riguardano questa questione.

Inoltre, alcuni hanno sottolineato che i GPT potrebbero violare leggi come il *Regolamento generale sulla protezione dei dati dell'Unione europea (2016)* o GDPR, soprattutto il diritto delle persone all'oblio, poiché attualmente è impossibile rimuovere i dati di qualcuno (o i risultati di quei dati) da un modello GPT, una volta che è stato addestrato.

Implicazioni per l'istruzione e la ricerca

- I ricercatori, gli insegnanti e gli studenti devono conoscere i diritti dei proprietari dei dati e verificare se gli strumenti GenIA che stanno utilizzando violano eventuali regolamentazioni esistenti.
- I ricercatori, gli insegnanti e gli studenti devono anche essere consapevoli del fatto che le immagini o il codice creati con la GenIA potrebbero violare i diritti di proprietà intellettuale e che le immagini, i suoni o il codice che creano e condividono su Internet potrebbero essere sfruttati da altre GenIA.

2.4 Modelli inesplicabili utilizzati per generare output

È da tempo risaputo che le reti neurali artificiali (ANN) sono di solito "scatole nere", cioè che il loro funzionamento interno non è aperto all'ispezione. Di conseguenza, le ANN non sono "trasparenti" o "spiegabili", e non è possibile stabilire come siano stati determinati i loro output.

Sebbene l'approccio generale, compresi gli algoritmi utilizzati, sia in genere spiegabile, i modelli specifici e i loro parametri, compresi i pesi del modello, non sono ispezionabili, motivo per cui un output specifico che viene generato non può essere spiegato. Ci sono miliardi di parametri/pesi in un modello come GPT-4 (vedi Tabella 2) e sono i pesi nel loro insieme che conservano i pattern appresi che il modello utilizza per generare i suoi output. Poiché i parametri/pesi non sono trasparenti nelle ANN (**Tabella 1**), non è possibile spiegare il modo preciso in cui un output specifico viene creato da questi modelli.

La mancanza di trasparenza della GenIA è sempre più problematica man mano che la GenIA diventa sempre più complessa (**vedi Tabella 2**), producendo spesso risultati inaspettati o indesiderati. Inoltre, i modelli GenIA ereditano e perpetuano i bias presenti nei dati di addestramento, che, date le modalità non trasparenti dei modelli, sono difficili da individuare e affrontare. Come è ovvio, questa opacità è anche una delle principali cause di sfiducia nella GenIA (Nazaretsky et al., 2022a). Se gli utenti non capiscono come un sistema GenIA ha ottenuto un output specifico, sono meno propensi ad adottarlo o utilizzarlo (Nazaretsky et al., 2022b).

Implicazioni per l'istruzione e la ricerca

I ricercatori, gli insegnanti e gli studenti dovrebbero essere consapevoli che i sistemi GenIA operano come scatole nere e che è quindi difficile, se non impossibile, conoscere il processo che ha generato un determinato contenuto. La mancanza di spiegazione su come vengono generati gli output tende a imprigionare gli utenti nella logica definita dai parametri progettati nei sistemi GenIA. Questi parametri possono riflettere valori e norme culturali o commerciali specifici che implicitamente influenzano il contenuto prodotto.

2.5 Contenuti generati da IA che inquinano Internet

Poiché i dati di addestramento di GPT sono generalmente prelevati da Internet, che troppo spesso include linguaggio discriminatorio e altre forme inaccettabili, gli sviluppatori hanno dovuto implementare ciò che chiamano "guardrail" per evitare che gli output di GPT risultino offensivi e/o non etici. Tuttavia, a causa dell'assenza di regolamentazioni rigorose e di meccanismi di monitoraggio efficaci, si stanno diffondendo sempre di più su Internet materiali di parte generati dalla GenIA, inquinando una delle principali fonti di contenuti o conoscenze per la maggior parte degli studenti in tutto il mondo. Questo è particolarmente importante perché il materiale generato dalla GenIA può sembrare molto accurato e convincente, e invece spesso contiene errori e idee tendenziose. Ciò rappresenta un alto rischio per gli alunni che non hanno una solida conoscenza preliminare dell'argomento in questione. Questo a spetto rappresenta anche un rischio ricorsivo per i futuri modelli GPT, che verranno addestrati su testo prelevato da Internet che i modelli GPT stessi hanno creato e che potenzialmente includono anche i loro pregiudizi ed errori.

Implicazioni per l'istruzione e la ricerca

- I ricercatori, gli insegnanti e gli studenti devono essere consapevoli che i sistemi GenIA sono in grado di produrre materiali offensivi e non etici.
- Devono anche essere informati circa le questioni a lungo termine che potrebbero sorgere per l'affidabilità delle conoscenze, quando futuri modelli GPT si baseranno su testo generato da modelli GPT precedenti.

2.6 Mancanza di comprensione del mondo reale

I GPT di testo sono talvolta definiti in modo dispregiativo come "Pappagalli Stocastici" perché, come già evidenziato in precedenza, sebbene siano in grado di produrre testi che sembrano convincenti, talvolta producono errori e generano affermazioni dannose (Bender et al., 2021). Tutto ciò avviene perché i GPT ripetono solo i modelli linguistici presenti nei dati di addestramento (di solito testo prelevato da Internet).

I GPT di testo partono da modelli di modelli casuali (o "stocastici") e senza comprenderne il significato, proprio come un pappagallo può imitare suoni senza realmente comprendere ciò che sta dicendo.

La discrepanza tra l'apparente comprensione del testo da parte dei modelli GenIA e la realtà che essi non comprendono il linguaggio e il mondo reale può indurre insegnanti e studenti a attribuire un ingiustificato grado di fiducia all'output. Ciò comporta seri rischi per l'istruzione. Infatti, GenIA non si basa su osservazioni del mondo reale o su altri aspetti chiave del metodo scientifico, né è allineata con i valori umani o sociali. Per questi motivi, non può generare contenuti veramente originali e affidabili sul mondo reale, sugli oggetti e sulle loro relazioni, sulle persone e sulle relazioni sociali, sulle relazioni tra uomo e oggetto o tra uomo e tecnologia. Se il contenuto apparentemente nuovo generato dai modelli GenIA possa essere riconosciuto come conoscenza scientifica è tuttora oggetto di dibattito.

Come già evidenziato, i GPT possono spesso produrre testi inesatti o non affidabili. In effetti, è ben noto che i GPT inventano alcune cose che non esistono nella vita reale. Alcuni chiamano ciò "allucinazione", anche se altri criticano l'uso di un termine così antropomorfo e quindi fuorviante. È una situazione ben nota alle aziende produttrici di GenIA. Non a caso, in fondo all'interfaccia pubblica di ChatGPT, si legge: "ChatGPT può produrre informazioni inaccurate su persone, luoghi o fatti".

Alcuni sostengono che la GenIA rappresenti un passo significativo nel percorso verso l'Intelligenza Artificiale Generale (AGI), un termine che suggerisce una classe di IA più intelligente degli esseri umani. Tuttavia, ciò è stato a lungo oggetto di critiche poiché l'IA non progredirà mai verso l'AGI almeno fino a quando non unirà in qualche modo, in simbiosi, sia l'IA basata sulla conoscenza (nota anche come IA simbolica o basata su regole) che l'IA basata sui dati (nota anche come apprendimento automatico) (Marcus, 2022). (...)

Implicazioni per l'istruzione e la ricerca

- L'output di un testo GenIA può sembrare straordinariamente simile a quello umano, come se comprendesse il testo che ha generato. Tuttavia, il GenIA non comprende nulla e si limita a costruire frasi di senso compiuto unendo le parole sulla base dei dati immagazzinati. Il testo generato può anche essere scorretto.
- I ricercatori, gli insegnanti e gli studenti devono essere consapevoli del fatto che un GPT non comprende il testo che genera, che può, e spesso lo fa, generare affermazioni errate: è perciò necessario adottare un approccio critico a tutto ciò che un GPT genera.

2.7 Ridurre la diversità di opinioni e marginalizzare ulteriormente voci già emarginate

ChatGPT e strumenti simili tendono a produrre solo risposte standard che assumono i valori dei proprietari/creatori dei dati utilizzati per addestrare i modelli. Pertanto, se una sequenza di parole appare frequentemente nei dati di addestramento, come nel caso di argomenti comuni e non controversi e delle convinzioni dominanti o mainstream, è probabile che sia ripetuta dal GPT nella sua produzione.

Ciò rischia di limitare e minare lo sviluppo di opinioni plurali. Le popolazioni povere di dati, comprese le comunità emarginate del Nord del mondo, hanno una presenza digitale minima o limitata online. Di conseguenza, le loro voci non vengono ascoltate e le loro preoccupazioni non vengono rappresentate nei dati utilizzati per addestrare i GPT e quindi raramente compaiono nelle produzioni. Per queste ragioni, data la metodologia di preaddestramento basata su dati da pagine web di Internet e conversazioni sui social media, i modelli GPT possono ulteriormente emarginare persone già svantaggiate.

Implicazioni per l'istruzione e la ricerca

- Sebbene gli sviluppatori e i fornitori di modelli GenIA abbiano la responsabilità principale di affrontare continuamente i pregiudizi nei dati e nelle produzioni di questi modelli, i ricercatori, gli insegnanti e gli studenti devono sapere che l'output di un testo GenIA rappresenta solo la visione più comune o dominante del mondo al momento in cui sono stati prodotti i dati di addestramento e che parte di esso è problematica o distorta.
- Gli studenti, gli insegnanti e i ricercatori non dovrebbero mai accettare le informazioni fornite dal GenIA alla lettera e dovrebbero sempre valutarle criticamente.
- I ricercatori, gli insegnanti e gli studenti devono anche essere consapevoli di come le voci delle minoranze possano essere escluse, perché per definizione sono meno comuni nei dati di addestramento.

2.8 Generazione di deepfake più avanzati

Oltre alle controversie comuni a tutti i GenIA, quelli basati su GAN possono essere utilizzati per alterare o manipolare immagini o video esistenti al fine di generarne di falsi, difficilmente distinguibili da quelli reali. Il GenIA sta rendendo sempre più semplice creare questi "deepfake" e le cosiddette "fake news". In altre parole, il GenIA sta facilitando a certi attori la possibilità di compiere atti non etici, immorali e criminali, come diffondere disinformazione,

promuovere discorsi di odio e incorporare i volti delle persone, senza il loro consenso o conoscenza, in contesti completamente falsi e talvolta compromettenti.

Implicazioni per l'istruzione e la ricerca

Sebbene sia responsabilità dei fornitori di GenIA proteggere i diritti d'autore e i diritti all'immagine degli utenti, i ricercatori, gli insegnanti e gli studenti devono anche essere consapevoli che qualsiasi immagine che condividono su Internet potrebbe essere inclusa nei dati di addestramento di GenIA e potrebbe essere manipolata e utilizzata in modi non etici.

3. Regolamentazione dell'uso dell'IA generativa nell'istruzione

E' necessario regolamentare per affrontare le controversie legate all'IA generativa e sfruttare i potenziali vantaggi della GenIA nell'istruzione. La regolamentazione della GenIA per scopi educativi richiede una serie di passi e misure politiche basate su un approccio centrato sull'essere umano per garantirne un uso etico, sicuro, equo e significativo.

3.1 Un approccio centrato sull'essere umano all'IA

La *Raccomandazione del 2021 sull'Etica dell'Intelligenza Artificiale* fornisce il quadro normativo necessario per iniziare ad affrontare le numerose controversie legate all'IA generativa, comprese quelle legate all'istruzione e alla ricerca. Si basa su un approccio all'IA centrato sull'essere umano, e sostiene che l'uso dell'IA debba essere al servizio dello sviluppo delle capacità umane per un futuro inclusivo, giusto e sostenibile. Tale approccio deve essere guidato dai principi dei diritti umani e dalla necessità di proteggere la dignità umana e la diversità culturale che definisce il patrimonio comune del sapere. In termini di governance, un approccio centrato sull'essere umano richiede una regolamentazione adeguata in grado di garantire la capacità umana di agire e prendere decisioni, la trasparenza e la responsabilità pubblica.

Il *Consenso di Pechino sull'Intelligenza Artificiale (AI) e l'Educazione* del 2019 approfondisce ulteriormente cosa implica un approccio centrato sull'essere umano per l'uso dell'IA nel contesto dell'istruzione. Il Consenso afferma che l'uso delle tecnologie AI nell'istruzione dovrebbe potenziare le capacità umane per lo sviluppo sostenibile e la collaborazione efficace tra esseri umani e macchine nella vita, nell'apprendimento e nel lavoro. Esso richiama anche a ulteriori azioni per garantire un accesso equo all'IA al fine di sostenere le persone svantaggiate e affrontare le disuguaglianze, promuovendo nel contempo le diversità linguistiche e culturali. Il Consenso suggerisce di adottare approcci a tutto governo, intersettoriali e multistakeholder nella pianificazione delle politiche sull'IA nell'istruzione.

"AI and education: Guidance for policy-makers" (UNESCO, 2022b) affina ulteriormente cosa significhi un approccio centrato sull'essere umano quando si esaminano i benefici e i rischi dell'IA nell'istruzione e il ruolo dell'istruzione come mezzo per sviluppare competenze nell'ambito dell'IA. Esso propone raccomandazioni concrete per la formulazione di politiche per guidare l'uso dell'IA al fine di: (i) consentire l'accesso inclusivo a programmi di apprendimento, soprattutto per gruppi vulnerabili come gli studenti con disabilità; (ii) sostenere opzioni di apprendimento personalizzate e aperte; (iii) migliorare le disposizioni e la gestione basate sui dati per ampliare l'accesso e migliorare la qualità dell'apprendimento; (iv) monitorare i processi di apprendimento e avvertire gli insegnanti dei rischi di fallimento; e (v) sviluppare comprensione e competenze per un uso etico e significativo dell'IA.

3.2 Passi per regolare la GenIA nell'istruzione

Prima della pubblicazione di ChatGPT, i governi stavano sviluppando o adattando quadri normativi per regolare la raccolta e l'uso di dati e l'adozione di sistemi AI in vari settori, compreso l'istruzione: ciò ha fornito un contesto legislativo e politico per la regolamentazione delle nuove applicazioni emergenti di AI. In seguito alla pubblicazione di più modelli concorrenziali di GenIA a partire da novembre 2022, i governi hanno adottato diverse risposte politiche: dal divieto della GenIA alla valutazione delle esigenze di adattamento dei quadri esistenti, fino alla formulazione urgente di nuove normative.

Le strategie governative per regolare e agevolare l'uso creativo della GenIA sono state mappate e valutate nell'aprile 2023 (UNESCO, 2023b). La revisione suggerisce una serie di sei passi che le agenzie governative possono intraprendere per regolamentare l'IA generativa e riaffermare il controllo pubblico al fine di sfruttarne i potenziali in vari settori, compreso l'istruzione.

Fase 1: Approvazione di Regolamenti Generali sulla Protezione dei Dati (GDPR) internazionali o regionali o sviluppo di GDPR nazionali

L'addestramento dei modelli GenIA ha comportato la raccolta e l'elaborazione di dati online da parte dei cittadini di molti paesi. L'uso, da parte dei modelli GenIA, di dati e contenuti senza il consenso rappresenta una sfida ulteriore per la protezione dei dati. I Regolamenti Generali sulla Protezione dei Dati, con il GDPR dell'UE entrato in vigore nel 2018 come uno degli esempi precursori, forniscono il quadro legale necessario per regolare la raccolta e il trattamento dei dati personali da parte dei fornitori di GenIA. Secondo il portale Worldline sulla legislazione sulla protezione dei dati e sulla privacy delle Nazioni Unite per la Conferenza sul Commercio e lo Sviluppo (UNCTAD), 137 dei 194 paesi hanno istituito legislazioni per garantire la protezione dei dati e della privacy. Tuttavia, rimane poco chiaro in che misura questi quadri normativi vengano effettivamente attuati in tali paesi. Pertanto, è sempre più difficile assicurarsi che vengano adeguatamente implementati, compresa la regolare monitoraggio delle operazioni dei sistemi GenIA. È anche urgente che i paesi che ancora non dispongono di leggi generali sulla protezione dei dati le sviluppino.

Fase 2: Adozione/revisione e finanziamento di strategie governative complete sull'IA

La regolamentazione dell'IA generativa deve far parte integrante delle più ampie strategie nazionali sull'IA, affinché possano garantire un uso sicuro ed equo dell'IA in tutti i settori dello sviluppo, compreso l'istruzione. La formulazione, l'approvazione, il finanziamento e l'attuazione delle strategie nazionali sull'IA richiedono un approccio a tutto tondo per poter garantire la coordinazione delle azioni intersettoriali necessarie a rispondere in modo integrato alle sfide emergenti.

Entro l'inizio del 2023, circa 67 paesi avevano sviluppato o pianificato strategie nazionali sull'IA, di cui 61 avevano la forma di una strategia autonoma sull'IA e 7 facevano parte di strategie nazionali più ampie sull'ICT o sulla digitalizzazione. Comprensibilmente, dato il suo carattere innovativo, nessuna di queste strategie nazionali aveva ancora affrontato specificamente l'IA generativa al momento della stesura.

È fondamentale che i paesi rivedano le strategie nazionali sull'IA esistenti o le sviluppino, garantendo disposizioni per regolamentare l'uso etico dell'IA in tutti i settori, compreso l'istruzione.

Fase 3: Consolidare e attuare regolamenti specifici sull'etica dell'IA

Per affrontare le dimensioni etiche poste dall'uso dell'IA, sono necessarie regolamentazioni specifiche. La revisione del 2023 dell'UNESCO sulle strategie nazionali sull'IA esistenti indica che l'identificazione di tali questioni etiche e la formulazione di principi guida sono comuni solo a circa quaranta strategie nazionali sull'IA. E anche in questi casi, i principi etici dovranno essere tradotti in leggi o regolamenti applicabili. Questo avviene raramente: solo una ventina di paesi, infatti, aveva definito regolamentazioni chiare sull'etica dell'IA, compresi i riferimenti all'istruzione, sia come parte delle strategie nazionali sull'IA che in altre forme. È interessante notare che, sebbene l'istruzione sia evidenziata come un ambito politico in circa quarantacinque strategie nazionali sull'IA, i riferimenti all'istruzione sono formulati più in termini di competenze e sviluppo del talento nell'IA necessari per sostenere la competitività nazionale, e meno in termini di questioni etiche.

I paesi che non hanno ancora regolamentazioni sull'etica dell'IA devono urgentemente definirle e attuarle.

Fase 4: Adeguare o far rispettare le leggi sul copyright esistenti per regolamentare i contenuti generati dall'IA

L'uso sempre più diffuso dell'IA generativa ha introdotto nuove sfide per il copyright, sia per quanto riguarda i contenuti o le opere coperte da copyright su cui i modelli vengono addestrati, sia per lo status delle produzioni di conoscenza 'non umane' che essi generano.

Attualmente, solo Cina, paesi dell'Unione Europea (UE) e Stati Uniti hanno adeguato le leggi sul copyright per tener conto delle implicazioni dell'IA generativa. Ad esempio, l'Ufficio del Copyright degli Stati Uniti ha stabilito che l'output dei sistemi di IA generativa, come ChatGPT, non è protetto dalla legge sul copyright degli Stati Uniti, sostenendo che "il copyright può proteggere solo materiale che è prodotto dalla creatività umana" (Ufficio del Copyright degli Stati Uniti, 2023). Mentre nell'UE, la proposta di legge sull'IA dell'UE richiede agli sviluppatori di strumenti AI di divulgare i materiali coperti da copyright utilizzati per costruire i loro sistemi (Commissione europea, 2021). La Cina, attraverso la sua regolamentazione sull'IA generativa

pubblicata nel luglio 2023, richiede l'etichettatura delle produzioni dell'IA generativa come contenuto generato dall'IA e le riconosce solo come produzioni di sintesi digitale.

La regolamentazione dell'uso di materiali coperti da copyright nell'addestramento dei modelli di IA generativa e la definizione dello status di copyright delle produzioni dell'IA generativa stanno emergendo come nuove responsabilità delle leggi sul copyright esistenti. È urgente adeguare le leggi esistenti a queste esigenze.

Fase 5: Sviluppare quadri regolatori sull'IA generativa

Il rapido sviluppo delle tecnologie dell'IA sta costringendo le agenzie di governo nazionali/locali ad accelerare il rinnovo delle regolamentazioni. Al luglio 2023, solo un paese, la Cina, aveva emanato una regolamentazione ufficiale specifica sull'IA generativa. Il Regolamento Provvisorio sulla Gestione del Servizio di IA Generativa, pubblicato il 13 luglio 2023 (Amministrazione Cinese per lo Spazio Cibernetico, 2023a), richiede ai fornitori di sistemi di IA generativa di etichettare correttamente e legalmente il contenuto, le immagini e i video generati dall'IA, in conformità con il suo Regolamento esistente sulla Sintesi Profonda nel Contesto dei Servizi di Informazione Online. È necessario sviluppare ulteriori quadri normativi nazionali specifici sull'IA generativa basati su una valutazione delle lacune nelle regolamentazioni e leggi locali esistenti.

Fase 6: Costruire capacità per l'uso adeguato dell'IA generativa nell'istruzione e nella ricerca

Le Scuole devono sviluppare capacità per comprendere i potenziali benefici e rischi dell'IA, compresa l'IA generativa, per l'istruzione. È solo basandosi su questa comprensione che possono convalidare l'adozione di strumenti basati sull'IA. Inoltre, i docenti e i ricercatori devono essere sostenuti per rafforzare le loro capacità per l'uso adeguato dell'IA generativa, anche attraverso la formazione e il coaching continuo. Alcuni paesi hanno avviato programmi di sviluppo di capacità di questo tipo, tra cui Singapore, che offre una piattaforma dedicata allo sviluppo delle capacità di IA delle istituzioni scolastiche attraverso il suo AI Government Cloud Cluster, che include un repository dedicato di modelli GPT (Ocampo, 2023).

Fase 7: Riflettere sulle implicazioni a lungo termine dell'IA generativa per l'istruzione e la ricerca

L'impatto delle versioni attuali dell'IA generativa sta solo cominciando a dispiegarsi, e i loro effetti sull'istruzione devono ancora essere completamente esplorati e compresi. Nel frattempo, continuano a essere sviluppate e implementate versioni più potenti dell'IA generativa e di altre classi di IA. Tuttavia, rimangono cruciali domande sulle implicazioni dell'IA generativa per la creazione, la trasmissione e la convalida della conoscenza, per l'insegnamento e l'apprendimento, per la progettazione dei curricula e la valutazione, nonché per la ricerca e il copyright. La maggior parte dei paesi si trova nella fase iniziale dell'adozione dell'IA generativa nell'istruzione, anche se gli impatti a lungo termine devono ancora essere compresi. Per garantire un uso centrato sull'essere umano dell'IA, è urgente condurre un dibattito pubblico aperto e un dialogo politico sulle implicazioni a lungo termine. Un dibattito inclusivo che coinvolge il governo, il settore privato e altri partner dovrebbe servire a fornire spunti e contributi per il rinnovo iterativo delle regolamentazioni e delle politiche.

3.3 Regolamenti sull'AI generativa: Elementi chiave

Tutti i paesi devono regolamentare adeguatamente l'AI generativa al fine di garantire che essa contribuisca allo sviluppo nell'ambito dell'istruzione e in altri contesti. Questa sezione propone azioni legate agli elementi chiave che possono essere intraprese da: (1) agenzie di regolamentazione governative, (2) fornitori di strumenti abilitati dall'AI, (3) utenti istituzionali e (4) utenti individuali. Sebbene molti degli elementi nel quadro siano di natura transnazionale, tutti dovrebbero essere considerati anche alla luce del contesto locale; cioè, dei sistemi educativi specifici di ciascun paese e dei quadri regolamentari generali già in atto.

3.3.1. Agenzie di regolamentazione governative

È necessario un approccio a tutto il governo per la coordinazione della progettazione, l'allineamento e l'attuazione delle regolamentazioni sull'IA generativa. Vengono raccomandati i seguenti sette elementi chiave e le relative azioni:

- **Coordinamento intersettoriale:** istituire un organismo nazionale per guidare l'approccio a tutto il governo all'IA generativa e coordinare la cooperazione tra settori.
- **Allineamento della legislazione:** allineare il quadro normativo con i contesti legislativi e regolamentari rilevanti di ciascun paese, ad esempio le leggi sulla protezione dei dati, le normative sulla sicurezza su internet, le leggi sulla sicurezza dei dati prodotti o utilizzati per servire i cittadini e altre leggi rilevanti e pratiche consuete. Valutare l'adeguatezza delle normative esistenti e le eventuali necessarie adattamenti in risposta alle nuove problematiche sollevate dall'IA generativa.
- **Equilibrio tra la regolamentazione dell'IA generativa e la promozione dell'innovazione nell'IA:** promuovere la cooperazione intersettoriale; favorire la collaborazione tra aziende e organizzazioni di governance settoriale, istituzioni educative e di ricerca, nonché pertinenti agenzie pubbliche per sviluppare congiuntamente modelli affidabili; incoraggiare la creazione di ecosistemi open source per favorire la condivisione di risorse di supercalcolo e dataset di pre-formazione di alta qualità; promuovere l'applicazione pratica dell'IA generativa in tutti i settori e la creazione di contenuti di alta qualità per il bene pubblico.
- **Valutazione e classificazione dei potenziali rischi dell'IA:** stabilire principi e un processo per la valutazione e la categorizzazione dell'efficacia, della sicurezza e della protezione dei servizi di IA generativa, prima della loro implementazione e durante il ciclo di vita del sistema. Considerare meccanismi di categorizzazione basati sui livelli di rischio che l'IA generativa può comportare per i cittadini. Classificarli in regolamentazioni rigorose (cioè il divieto di applicazioni o sistemi basati su IA con rischi inaccettabili), regolamentazioni speciali per applicazioni ad alto rischio e regolamentazioni generali per applicazioni che non presentano un alto rischio. Si veda il progetto di legge UE sull'IA per un esempio di questo approccio.
- **Protezione della privacy dei dati:** considerare il fatto che l'uso dell'IA generativa coinvolge quasi sempre gli utenti nella condivisione dei propri dati con il fornitore di IA generativa. Imporre la stesura e l'attuazione di leggi per la protezione delle informazioni personali degli utenti e individuare e contrastare la conservazione, la profilazione e la condivisione illecite dei dati.
- **Definizione e applicazione del limite di età per l'uso dell'IA generativa:** la maggior parte delle applicazioni di IA generativa è principalmente progettata per utenti adulti. Queste applicazioni comportano spesso rischi significativi per i bambini, tra cui l'esposizione a contenuti inappropriati e la possibilità di manipolazione. Alla luce di questi rischi e date le notevoli incertezze che continuano a circondare le applicazioni iterative dell'IA generativa, si raccomanda vivamente di introdurre restrizioni di età per le tecnologie di IA a uso generale al fine di proteggere i diritti e il benessere dei bambini.

Attualmente, i termini d'uso di ChatGPT richiedono che gli utenti abbiano almeno 13 anni e che gli utenti sotto i 18 anni abbiano il permesso dei loro genitori o tutori legali per utilizzare i Servizi.

Questi limiti di età o soglie fanno riferimento al Children's Online Privacy Protection Act degli Stati Uniti d'America (Federal Trade Commission, 1998). Approvata nel 1998 prima dell'ampio utilizzo dei social media e ben prima della creazione di applicazioni IA generative facili da usare e potenti come ChatGPT, la legge statunitense specifica che le organizzazioni o i singoli fornitori di social media non possono fornire servizi a bambini di età inferiore ai 13 anni senza il permesso dei genitori. Molti analisti considerano questa soglia troppo bassa e hanno sostenuto legislazioni per innalzare l'età a 16 anni. Il GDPR dell'Unione Europea (2016) specifica che gli utenti devono avere almeno 16 anni per utilizzare i servizi dei social media senza il permesso dei genitori.

L'emergere di vari chatbot GenIA richiede che i paesi considerino attentamente - e deliberino pubblicamente - la soglia di età appropriata per conversazioni indipendenti con piattaforme GenIA. La soglia minima dovrebbe essere di tredici anni. I paesi dovranno anche decidere se l'autodichiarazione dell'età rimane un mezzo appropriato di verifica dell'età. I paesi dovranno imporre le responsabilità dei fornitori di GenIA per la verifica dell'età e le responsabilità dei genitori o tutori per monitorare le conversazioni indipendenti dei minori.

- Proprietà nazionale dei dati e rischio di povertà dei dati: adottare misure legislative per proteggere la proprietà nazionale dei dati e regolamentare i fornitori di GenIA che operano all'interno dei propri confini. Per i set di dati generati dai cittadini utilizzati a scopi commerciali, stabilire regolamentazioni per promuovere la cooperazione reciproca vantaggiosa in modo che questa categoria di dati non venga sottratta dal paese per essere sfruttata esclusivamente dalle grandi aziende tecnologiche.

3.3.2. Fornitori di strumenti GenIA

I fornitori di GenIA includono organizzazioni e individui responsabili dello sviluppo e della messa a disposizione di strumenti GenIA, e/o che utilizzano tecnologie GenIA per fornire servizi, anche attraverso API programmabili. La maggior parte dei fornitori influenti di strumenti GenIA sono aziende estremamente ricche.

Deve essere chiaro ai fornitori di GenIA che sono responsabili per l'etica fin dalla progettazione, compresa l'implementazione dei principi etici stabiliti nei regolamenti. Le seguenti dieci categorie di responsabilità dovrebbero essere coperte:

- Responsabilità umane: i fornitori di GenIA dovrebbero essere ritenuti responsabili di garantire l'adesione ai valori fondamentali e agli scopi leciti, il rispetto della proprietà intellettuale e il mantenimento delle pratiche etiche, prevenendo nel contempo la diffusione di disinformazione e discorsi d'odio.
- Dati e modelli affidabili: ai fornitori di GenIA dovrebbe essere richiesto di dimostrare l'affidabilità e l'eticità delle fonti di dati e dei metodi utilizzati dai loro modelli e dalle relative uscite. Devono essere obbligati ad adottare dati e modelli di base provenienti da fonti legali comprovate e a rispettare le leggi sulla proprietà intellettuale pertinenti (ad esempio, se i dati sono protetti da diritti di proprietà intellettuale). Inoltre, quando i modelli necessitano di utilizzare informazioni personali, la raccolta di tali informazioni dovrebbe avvenire solo con il consenso informato ed esplicito dei proprietari.
- Generazione di contenuti non discriminatori: i fornitori di GenIA devono vietare la progettazione e la messa in funzione di sistemi GenIA che generano contenuti tendenziosi o discriminatori basati su razza, nazionalità, genere o altre caratteristiche protette. Dovrebbero assicurarsi che siano in atto robusti "guardrail" per evitare che GenIA produca contenuti offensivi, tendenziosi o falsi, garantendo nel contempo che gli esseri umani coinvolti nel definire tali "guardrail" siano protetti e non sfruttati.
- Spiegabilità e trasparenza dei modelli GenIA: i fornitori dovrebbero fornire alle agenzie di regolamentazione pubblica le spiegazioni delle fonti, delle dimensioni e dei tipi di dati utilizzati dai modelli, delle regole per l'etichettatura dei dati nella pre-formazione, dei metodi o algoritmi che i modelli utilizzano per generare contenuti o risposte e dei servizi forniti dai loro strumenti GenIA. Quando necessario, dovrebbero offrire supporto per aiutare le agenzie di regolamentazione a comprendere la tecnologia e i dati. La tendenza di GenIA a generare contenuti con errori e risposte contestabili dovrebbe essere resa trasparente per gli utenti.
- Etichettatura dei contenuti GenIA: In conformità alle leggi o regolamenti pertinenti sull'assistenza AI alla sintesi di informazioni online, i fornitori devono etichettare adeguatamente e legalmente documenti, relazioni, immagini e video generati da GenIA. Ad esempio, l'output GenIA dovrebbe essere chiaramente etichettato come prodotto da una macchina.
- Principi di sicurezza: I fornitori di GenIA dovrebbero garantire un servizio sicuro, robusto e sostenibile per l'intero ciclo di vita di un sistema GenIA.
- Specifiche sull'appropriatezza per l'accesso e l'uso: I fornitori di GenIA dovrebbero fornire specifiche chiare sull'audience appropriata per i loro servizi, nonché scenari e scopi di utilizzo, e aiutare gli utenti degli strumenti GenIA a prendere decisioni razionali e responsabili.

- Riconoscimento dei limiti e prevenzione dei rischi prevedibili: I fornitori di GenIA dovrebbero pubblicizzare chiaramente i limiti dei metodi utilizzati dai sistemi. Devono sviluppare tecnologie per garantire che i dati di input, i metodi e le risposte non arrechino danni prevedibili agli utenti, insieme a protocolli per mitigare danni imprevedibili quando si verificano. Devono anche fornire orientamenti per aiutare gli utenti a comprendere i contenuti generati da GenIA basati su principi etici e prevenire una loro eccessiva dipendenza e assuefazione ai contenuti generati.
- Meccanismi di reclamo e rimedio: I fornitori di GenIA devono istituire meccanismi e canali per la raccolta di reclami da parte degli utenti e del pubblico in generale, e intraprendere azioni tempestive per accettare e processare questi reclami.
- Monitoraggio e segnalazione dell'uso illecito: I fornitori devono cooperare con le agenzie di regolamentazione pubblica per facilitare il monitoraggio e la segnalazione dell'uso illecito. Ciò include quando le persone utilizzano i prodotti GenIA in modi illegali o in violazione di principi etici o valori sociali, come la promozione della disinformazione o discorsi di odio, la generazione di spam o la creazione di malware

3.3.3 Utenti istituzionali

Utenti istituzionali includono autorità educative e istituzioni come università e scuole che hanno responsabilità nella determinazione se adottare GenIA e quali tipi di strumenti GenIA dovrebbero essere acquisiti e implementati all'interno dell'istituzione.

- Auditing istituzionale di algoritmi GenIA, dati e risultati: Implementare meccanismi per monitorare il meglio possibile gli algoritmi e i dati utilizzati dagli strumenti GenIA e i risultati che generano. Questo dovrebbe includere audit e valutazioni regolari, la protezione dei dati degli utenti e il filtro automatico dei contenuti inappropriati.
- Convalida della proporzionalità e protezione del benessere degli utenti: Implementare meccanismi di classificazione nazionali o sviluppare una politica istituzionale per categorizzare e convalidare sistemi e applicazioni GenIA. Assicurarsi che i sistemi GenIA adottati dall'istituzione siano in linea con quadri etici localmente convalidati e non arrechino danni prevedibili agli utenti dell'istituzione, specialmente ai bambini e ai gruppi vulnerabili.
- Esaminare e affrontare gli impatti a lungo termine: Nel tempo, fare affidamento su strumenti o contenuti GenIA nell'istruzione potrebbe avere effetti profondi sullo sviluppo delle capacità umane come le abilità di pensiero critico e creatività. Questi potenziali effetti dovrebbero essere valutati e affrontati.
- Appropriatezza dell'età: Valutare l'implementazione di restrizioni sull'età minima per l'uso indipendente di GenIA nell'istituzione.

3.3.4 Utenti individuali

Gli utenti individuali potrebbero includere tutte le persone nel mondo che hanno accesso a Internet e almeno un tipo di strumento GenIA. Il termine "utenti individuali", come utilizzato qui, si riferisce principalmente a insegnanti, ricercatori e studenti individuali nelle istituzioni educative formali o a coloro che partecipano a programmi di studio non formali.

- Consapevolezza dei termini di riferimento sull'uso di GenIA: All'atto di firma o espressione di consenso agli accordi di servizio, gli utenti dovrebbero essere consapevoli degli obblighi di conformarsi ai Termini di Riferimento specificati nell'accordo e alle leggi o regolamenti che lo sostengono.
- Uso etico delle applicazioni GenIA: Gli utenti dovrebbero utilizzare GenIA in modo responsabile ed evitare di sfruttarlo in modi che potrebbero danneggiare la reputazione e i diritti legali di altre persone.
- Monitoraggio e segnalazione di applicazioni GenIA illegali: Quando scoprono applicazioni GenIA che violano una o più regolamentazioni, gli utenti dovrebbero notificare le agenzie di regolamentazione governative.

4. Verso un quadro politico per l'uso di Intelligenza Artificiale generativa nell'istruzione e nella ricerca

La regolamentazione di GenIA per sfruttare i potenziali benefici per l'istruzione e la ricerca richiede lo sviluppo di politiche adeguate. I dati del sondaggio del 2023 citati in precedenza indicano che solo pochi paesi hanno adottato politiche o piani specifici per l'uso dell'IA nell'istruzione. La sezione precedente ha delineato una visione, i passaggi necessari e gli elementi chiave e le azioni che possono essere intraprese da vari attori. Questa sezione fornisce misure che possono essere adottate per sviluppare quadri politici coerenti e completi per regolare l'uso di GenIA nell'istruzione e nella ricerca.

Un punto di partenza per questo è il documento del 2022 *"IA e istruzione: linee guida per i responsabili delle politiche"* (UNESCO, 2022b), che propone un insieme completo di raccomandazioni per guidare i governi nello sviluppo e nell'attuazione di politiche settoriali sull'IA e sull'istruzione, con un focus sulla promozione dell'istruzione di qualità, dell'equità sociale e dell'inclusione. La maggior parte delle raccomandazioni rimane applicabile e può essere ulteriormente adattata per guidare la formulazione di politiche specifiche su GenIA nell'istruzione. Vengono proposte di seguito otto misure specifiche per la pianificazione di politiche su GenIA nell'istruzione e nella ricerca per integrare l'orientamento esistente.

4.1 Promuovere l'inclusione, l'equità, la diversità linguistica e culturale

È necessario riconoscere e affrontare l'importanza critica dell'inclusione lungo l'intero ciclo di vita di GenIA. In particolare, gli strumenti GenIA non contribuiranno a affrontare le sfide fondamentali nell'istruzione o al raggiungimento degli impegni dell'Obiettivo di Sviluppo Sostenibile 4, a meno che tali strumenti non siano resi accessibili in modo inclusivo (indipendentemente dal genere, dall'etnia, dai bisogni educativi speciali, dallo status socio-economico, dalla posizione geografica...) e se non promuovono per design l'equità, la diversità linguistica e il pluralismo culturale. Per raggiungere questo obiettivo, sono raccomandate tre misure:

- Identificare coloro che non dispongono o non possono permettersi la connessione a Internet o i dati, e prendere misure per promuovere la connettività universale e le competenze digitali al fine di ridurre le barriere per un accesso equo e inclusivo alle applicazioni IA. Stabilire meccanismi di finanziamento sostenibili per lo sviluppo e la fornitura di strumenti abilitati dall'IA per gli studenti con disabilità o esigenze speciali. Promuovere l'uso di GenIA per sostenere i discenti di tutte le età, provenienze e sfondi nell'apprendimento lungo tutto l'arco della vita.
- Sviluppare criteri per la convalida dei sistemi GenIA per garantire che non vi sia alcun pregiudizio di genere, discriminazione contro gruppi emarginati o discorsi di odio incorporati nei dati o negli algoritmi.
- Sviluppare e attuare specifiche inclusive per i sistemi GenIA e mettere in atto misure istituzionali per proteggere le diversità linguistiche e culturali durante la diffusione di GenIA nell'istruzione e nella ricerca su larga scala. Le specifiche rilevanti dovrebbero richiedere ai fornitori di GenIA di includere dati in molte lingue, in particolare lingue locali o indigene, nella formazione dei modelli GPT per migliorare la capacità di GenIA di rispondere e generare testi multilingue. Le specifiche e le misure istituzionali dovrebbero impedire rigorosamente ai fornitori di IA la rimozione intenzionale o non intenzionale delle lingue minoritarie o la discriminazione contro i parlanti di lingue indigene e dovrebbero richiedere ai fornitori di interrompere i sistemi che promuovono lingue dominanti o norme culturali.

4.2 Tutelare l'agire umano

Con l'evolversi sempre più sofisticato di GenIA, un pericolo chiave è la sua potenziale capacità di compromettere l'agire umano. Man mano che sempre più utenti individuali utilizzano GenIA per sostenere la loro scrittura o altre attività creative, potrebbero involontariamente finire per fare affidamento esclusivamente su di esso, con la conseguenza potenziale di compromettere lo sviluppo delle capacità intellettuali. Sebbene GenIA possa essere

utilizzato per sfidare ed estendere il pensiero umano, non dovrebbe essere permesso di usurpare il pensiero umano. La protezione e l'incremento dell'agire umano dovrebbero sempre essere considerazioni fondamentali quando si progettano e si adottano sistemi di intelligenza artificiale generativa (GenIA) secondo le sette prospettive seguenti:

- Informare gli studenti sui tipi di dati che la GenIA potrebbe raccogliere da loro, su come vengono utilizzati questi dati e sull'effetto che potrebbe avere sulla loro istruzione e sulla loro vita più ampia.
- Proteggere la motivazione intrinseca degli studenti a crescere e imparare come individui. Rinforzare l'autonomia umana sulle proprie metodologie di ricerca, insegnamento e apprendimento nel contesto dell'uso di sistemi GenIA sempre più sofisticati.
- Prevenire l'uso della GenIA quando priverebbe gli studenti delle opportunità di sviluppare abilità cognitive e sociali attraverso l'osservazione del mondo reale, pratiche empiriche come esperimenti, discussioni con altri esseri umani e ragionamento logico indipendente.
- Garantire un'adeguata interazione sociale e un'esposizione appropriata alle produzioni creative realizzate dagli esseri umani e impedire che gli studenti diventino dipendenti dalla GenIA.
- Utilizzare gli strumenti GenIA per ridurre la pressione dei compiti e degli esami, anziché per aggravarla.
- Consultare ricercatori, insegnanti e studenti sulle loro opinioni sulla GenIA e utilizzare il feedback per decidere se e come specifici strumenti GenIA dovrebbero essere implementati su scala istituzionale. Incoraggiare studenti, insegnanti e ricercatori a criticare e mettere in discussione le metodologie dietro i sistemi di AI, l'accuratezza dei contenuti generati e le norme o pedagogie che potrebbero imporre.
- Evitare di trasferire la responsabilità umana ai sistemi GenIA nella presa di decisioni ad alto rischio.

4.3 Monitorare e convalidare i sistemi GenIA per l'istruzione

Come già indicato, lo sviluppo e l'implementazione della GenIA dovrebbero avere fondamenti etici fin dalla fase di progettazione. Successivamente, una volta che la GenIA è in uso e durante tutto il suo ciclo di vita, è necessario monitorarla attentamente e convalidarla - per i suoi rischi etici, la sua idoneità pedagogica e la sua rigore, e il suo impatto sugli studenti, gli insegnanti e le relazioni in classe o nella scuola. A tal proposito, si raccomandano le seguenti cinque azioni:

- Creare meccanismi di convalida per verificare se i sistemi GenIA utilizzati nell'istruzione e nella ricerca sono privi di pregiudizi, in particolare pregiudizi di genere, e se sono formati su dati rappresentativi della diversità (in termini di genere, disabilità, status sociale ed economico, background etnico e culturale e posizione geografica).
- Affrontare il complesso problema del consenso informato, in particolare nei contesti in cui i bambini o altri studenti vulnerabili non sono in grado di dare un consenso genuinamente informato.
- Verificare se gli output della GenIA includono immagini deepfake, notizie false o discorsi d'odio. Se si scopre che la GenIA sta generando contenuti inappropriati, le istituzioni e gli insegnanti dovrebbero essere disposti e in grado di adottare misure rapide e efficaci per mitigare o eliminare il problema.
- Esercitare una rigorosa convalida etica delle applicazioni GenIA prima che vengano ufficialmente adottate nelle istituzioni educative o di ricerca (ossia adottare un approccio etico nel design).

- Prima di prendere decisioni sull'adozione istituzionale, assicurarsi che le applicazioni GenIA in questione non causino danni prevedibili agli studenti, siano educativamente efficaci e valide per le età e le capacità degli studenti target, e siano in linea con principi pedagogici validi (ossia basati sui pertinenti domini del sapere e sugli obiettivi di apprendimento attesi e sullo sviluppo di valori).

4.4 Sviluppare competenze in intelligenza artificiale (AI), comprese le competenze relative alla GenIA, per gli studenti

Gli ultimi sviluppi della GenIA hanno ulteriormente enfatizzato l'urgente necessità per tutti di raggiungere un adeguato livello di alfabetizzazione nelle dimensioni umane e tecnologiche dell'Intelligenza Artificiale, comprendendo in modo generale come funziona e l'impatto specifico della GenIA. Per farlo, ora sono urgentemente necessarie le seguenti cinque azioni:

- Impegnarsi a fornire curricula di AI approvati dal governo per l'istruzione scolastica, per l'istruzione tecnica e professionale, nonché per l'apprendimento permanente. I curricula di AI dovrebbero coprire l'impatto dell'AI sulle nostre vite, comprese le questioni etiche che solleva, nonché una comprensione appropriata all'età degli algoritmi e dei dati, e competenze per l'uso corretto e creativo di strumenti di AI, comprese le applicazioni GenIA;
- Sostenere istituzioni di istruzione superiore e di ricerca per potenziare programmi volti a sviluppare talenti locali in AI;
- Promuovere l'uguaglianza di genere nello sviluppo di competenze avanzate in AI e creare un pool di professionisti equilibrato dal punto di vista di genere;
- Sviluppare previsioni intersettore sugli spostamenti di lavoro nazionali e globali causati dall'automazione dell'ultima GenIA e migliorare le competenze resistenti al futuro a tutti i livelli dei sistemi educativi e di apprendimento permanente in base agli spostamenti prospettici della domanda;
- Fornire programmi speciali per i lavoratori anziani e i cittadini che potrebbero dover imparare nuove competenze e adattarsi a nuovi ambienti.

4.5 Sviluppare capacità per insegnanti e ricercatori affinché possano fare un uso adeguato della GenIA.

Secondo i dati di un sondaggio del 2023 sull'uso governativo dell'IA nell'istruzione (UNESCO, 2023c), solo sette paesi (Cina, Finlandia, Georgia, Qatar, Spagna, Thailandia e Turchia) hanno riportato di aver sviluppato o stanno sviluppando quadri o programmi di formazione sull'IA per gli insegnanti. Solo il Ministero dell'Istruzione di Singapore ha segnalato la creazione di un repository online centrato sull'uso di ChatGPT nell'insegnamento e nell'apprendimento. Ciò dimostra chiaramente che nella maggior parte dei paesi gli insegnanti non hanno accesso a formazione ben strutturata sull'uso dell'IA nell'istruzione, e tanto meno sull'uso della GenIA.

Per preparare gli insegnanti per l'uso responsabile ed efficace della GenIA, i paesi devono intraprendere le seguenti quattro azioni:

- **Formulare o adeguare orientamenti basati su test locali:** Questo passo comporta la creazione o l'adattamento di linee guida basate su test locali per aiutare ricercatori e insegnanti a utilizzare in modo efficace le ampie disponibilità di strumenti GenIA e a orientare la progettazione di nuove applicazioni di IA specifiche per determinati settori. Questo favorirà l'adozione responsabile e mirata della GenIA.
- **Proteggere i diritti degli insegnanti e dei ricercatori:** È importante analizzare e proteggere i diritti dei docenti e dei ricercatori, nonché il valore delle loro pratiche quando utilizzano la GenIA. Questo può includere la definizione delle particolari funzioni degli insegnanti nell'agevolare il pensiero di ordine superiore, nell'organizzare l'interazione umana e nell'incoraggiare i valori umani all'interno dell'ambiente educativo.
- **Definire l'orientamento, la conoscenza e le competenze necessarie:** Questa azione implica la definizione chiara dell'orientamento, della conoscenza e delle competenze di cui gli insegnanti hanno bisogno per comprendere ed utilizzare sistemi GenIA in modo efficace ed etico. Gli

insegnanti dovrebbero essere abilitati a creare strumenti specifici basati su GenIA per agevolare l'apprendimento in classe e nel loro sviluppo professionale.

- **Rivedere dinamicamente le competenze necessarie:** Infine, è importante condurre una revisione dinamica delle competenze necessarie per gli insegnanti per comprendere e utilizzare l'IA nell'insegnamento, nell'apprendimento e per il loro sviluppo professionale. Questo dovrebbe includere l'integrazione di nuovi insiemi di valori, comprensione e competenze sull'IA nei quadri di competenze e nei programmi per la formazione di insegnanti in servizio e in formazione iniziale. In questo modo, gli insegnanti saranno preparati per affrontare le sfide e sfruttare le opportunità offerte dalla GenIA in modo responsabile ed efficace.

4.6 Promuovere opinioni plurali e espressioni plurali di idee

Come già indicato in precedenza, la GenIA non è in grado di comprendere né il prompt né la risposta: le sue risposte si basano su probabilità di modelli linguistici presenti nei dati (Internet) che ha assimilato durante la sua fase di addestramento del modello. Per affrontare alcuni dei problemi fondamentali delle sue uscite, attualmente si stanno studiando nuovi metodi, come la connessione della GenIA con basi di conoscenza e motori di ragionamento. Tuttavia, a causa del suo funzionamento, dei materiali di origine e delle prospettive implicite dei suoi sviluppatori, la GenIA, per definizione, riproduce visioni del mondo dominanti nelle sue uscite e mina le opinioni minoritarie e plurali. Di conseguenza, se le civiltà umane devono prosperare, è essenziale riconoscere che la GenIA non potrà mai essere una fonte autorevole di conoscenza su qualsiasi argomento con cui interagisce.

Di conseguenza, gli utenti devono valutare criticamente le uscite della GenIA. In particolare:

- Comprendere il ruolo della GenIA come una fonte di informazioni veloce, ma talvolta non affidabile. Alcuni plugin e strumenti basati su LLM menzionati in precedenza sono progettati per supportare la necessità di accedere a informazioni validate e aggiornate, ma finora vi sono poche prove robuste che dimostrino la loro efficacia.
- Incoraggiare gli studenti e i ricercatori a mettere in discussione le risposte fornite dalla GenIA. Riconoscere che la GenIA tende tipicamente a ripetere opinioni consolidate o standard, minando così le opinioni plurali e minoritarie e le diverse espressioni di idee.
- Fornire agli studenti sufficienti opportunità per apprendere attraverso il metodo di tentativi ed errori, esperimenti empirici e osservazioni del mondo reale.

4.7 Effettuare test di modelli di applicazioni rilevanti a livello locale e costruire una base di evidenze cumulative.

I modelli di GenIA sono finora dominati da informazioni provenienti dal Nord del mondo e sotto-rappresentano le voci del Sud del mondo e delle comunità indigene. Solo attraverso sforzi determinati, ad esempio sfruttando dati sintetici (Marwala, T. 2023), gli strumenti GenIA potranno diventare sensibili al contesto e alle esigenze delle comunità locali, in particolare di quelle del Sud del mondo. Per esplorare approcci pertinenti alle esigenze locali, collaborando in modo più ampio, si raccomandano le seguenti otto azioni:

- Assicurare che la progettazione e l'adozione della GenIA siano pianificate in modo strategico anziché facilitare un processo di approvvigionamento passivo e non critico.
- Incentivare i progettisti di GenIA a puntare su opzioni di apprendimento aperte, esplorative e diverse.
- Testare e ampliare casi d'uso basati su prove di applicazione dell'IA nell'istruzione e nella ricerca in conformità con le priorità educative, anziché basarsi su novità, miti o hype.

- Guidare l'uso della GenIA per stimolare l'innovazione nella ricerca, compreso lo sfruttamento delle capacità di calcolo, dei dati su larga scala e delle risposte della GenIA per informare e ispirare il miglioramento delle metodologie di ricerca.
- Esaminare le implicazioni sociali ed etiche dell'integrazione della GenIA nei processi di ricerca.
- Stabilire criteri specifici basati su ricerche e metodologie pedagogiche basate su prove e costruire una base di evidenze sull'efficacia della GenIA nel supportare l'offerta di opportunità di apprendimento inclusive, nel raggiungimento degli obiettivi di apprendimento e di ricerca e nella promozione delle diversità linguistiche e culturali.
- Intraprendere passi iterativi per rafforzare le prove sull'impatto sociale ed etico della GenIA.
- Analizzare i costi ambientali derivanti dallo sfruttamento su larga scala delle tecnologie AI (ad esempio, l'energia e le risorse necessarie per addestrare i modelli GPT) e sviluppare obiettivi sostenibili da raggiungere da parte dei fornitori di AI al fine di evitare di contribuire ai cambiamenti climatici.

4.8 Rivedere le implicazioni a lungo termine in un modo intersettoriale e interdisciplinare.

ChatGPT

Gli approcci intersettoriali e interdisciplinari sono essenziali per l'uso efficace ed etico della GenIA nell'istruzione e nella ricerca. Solo attraverso l'impiego di una serie di competenze e riunendo molteplici attori interessati, sarà possibile identificare prontamente le sfide chiave e affrontarle in modo efficace al fine di minimizzare le implicazioni negative a lungo termine, sfruttando nel contempo i benefici in corso e cumulativi. Pertanto, si raccomandano le seguenti tre azioni:

- Collaborare con fornitori di AI, insegnanti, ricercatori e rappresentanti dei genitori e degli studenti per pianificare adeguamenti a livello di sistema nei quadri curriculari e nelle metodologie di valutazione, al fine di sfruttare appieno il potenziale e mitigare i rischi della GenIA per l'istruzione e la ricerca.
- Riunire competenze intersettoriali e interdisciplinari, tra cui insegnanti, ricercatori, scienziati dell'apprendimento, ingegneri AI e rappresentanti di altri stakeholder, per esaminare le implicazioni a lungo termine della GenIA per l'apprendimento e la produzione di conoscenza, la ricerca e il copyright, il curriculum e la valutazione, e la collaborazione umana e le dinamiche sociali.
- Fornire consigli tempestivi per informare gli aggiornamenti iterativi di regolamenti e politiche.

5. Promuovere l'uso creativo della GenIA nell'istruzione e nella ricerca

Quando ChatGPT è stato lanciato per la prima volta, in tutto il mondo molti docenti hanno espresso preoccupazioni riguardo al suo potenziale nel generare saggi e su come potesse aiutare gli studenti a imbrogliare. Più recentemente, molte persone e organizzazioni, comprese alcune delle università più prestigiose del mondo, hanno sostenuto che "il genio è fuori dalla bottiglia" e che strumenti come ChatGPT sono qui per restare e possono essere utilizzati in modo produttivo nei contesti educativi. Nel frattempo, Internet è ora piena di suggerimenti sull'uso della GenIA nell'istruzione e nella ricerca: si parla dell'uso della GenIA per ispirare nuove idee, generare esempi da molteplici prospettive, sviluppare piani di lezione e presentazioni, riassumere materiali esistenti e stimolare la creazione di immagini. Anche se nuove idee al riguardo compaiono su Internet praticamente ogni giorno, i ricercatori e gli insegnanti stanno ancora cercando di capire esattamente cosa significhi la GenIA per l'insegnamento, l'apprendimento e la ricerca. In particolare, molte persone dietro le proposte di utilizzo potrebbero non aver considerato adeguatamente i principi etici, mentre altre sono guidate dalle potenzialità tecnologiche della GenIA piuttosto che dalle esigenze dei ricercatori, degli insegnanti o degli studenti. Questa sezione illustra alcune modalità in cui è possibile favorire l'uso creativo della GenIA nell'istruzione.

5.1 ChatGPT Strategie istituzionali per agevolare l'uso responsabile e creativo della GenIA.

Come già indicato in precedenza, le istituzioni scolastiche e di ricerca dovrebbero sviluppare, implementare e convalidare strategie adeguate e quadri etici per guidare l'uso responsabile ed etico dei sistemi e delle applicazioni GenIA al fine di soddisfare le esigenze dell'insegnamento, dell'apprendimento e della ricerca. Ciò può essere ottenuto attraverso le seguenti quattro strategie:

- Implementazione istituzionale di principi etici: Assicurare che i ricercatori, gli insegnanti e gli studenti utilizzino gli strumenti GenIA in modo responsabile ed etico, e si avvicinino criticamente all'accuratezza e alla validità delle risposte.
- Orientamento e formazione: Fornire orientamento e formazione a ricercatori, insegnanti e studenti sugli strumenti GenIA per assicurarsi che comprendano le questioni etiche, come i pregiudizi nella marcatura dei dati e negli algoritmi, e che rispettino le normative appropriate sulla privacy dei dati e sulla proprietà intellettuale.
- Sviluppo delle capacità di creare prompt per la GenIA: Oltre alla conoscenza specifica della materia, i ricercatori e gli insegnanti avranno bisogno di competenze nella creazione e nella valutazione critica dei prompt generati dalla GenIA. Dato che le sfide sollevate dalla GenIA sono complesse, i ricercatori e gli insegnanti devono ricevere formazione di alta qualità e supporto per farlo.
- Rilevamento del plagio basato su GenIA nei compiti scritti: La GenIA potrebbe consentire agli studenti di presentare come proprio un testo che non hanno scritto, creando un nuovo tipo di "plagio". I fornitori di GenIA sono tenuti a contrassegnare le loro risposte con filigrane del tipo "generato da AI", mentre sono in fase di sviluppo strumenti per identificare materiale prodotto da AI. Tuttavia, vi sono poche prove che queste misure o strumenti siano efficaci. La strategia istituzionale immediata è di preservare l'integrità accademica e rafforzare la responsabilità attraverso una rigorosa individuazione da parte degli esseri umani. La strategia a lungo termine prevede che istituzioni e insegnanti ripensino la progettazione dei compiti scritti in modo che non vengano utilizzati per valutare compiti che gli strumenti GenIA possono svolgere meglio degli studenti. Invece, dovrebbero affrontare ciò che gli esseri umani possono fare e che la GenIA e altri strumenti AI non possono fare, compresa l'applicazione di valori umani come la compassione e la creatività a sfide complesse del mondo reale.

5.2 Un approccio centrato sull'essere umano e pedagogicamente appropriato all'interazione

I ricercatori e gli insegnanti dovrebbero dare priorità alla capacità umana di agire e prendere decisioni, e a un'interazione responsabile e pedagogicamente appropriata tra esseri umani e strumenti AI quando decidono se e come utilizzare la GenIA. Questo obiettivo include le seguenti cinque considerazioni:

- L'uso dello strumento (o degli strumenti) dovrebbe contribuire alle esigenze umane e rendere l'apprendimento o la ricerca più efficaci rispetto a un approccio senza tecnologia o a un'alternativa diversa;
- L'uso dello strumento (o degli strumenti) da parte degli insegnanti e degli studenti dovrebbe essere basato sulla loro motivazione intrinseca;
- Il processo di utilizzo dello strumento (o degli strumenti) dovrebbe essere controllato dagli insegnanti, dagli studenti o dai ricercatori;
- La scelta e l'organizzazione dello strumento (o degli strumenti) e dei contenuti che generano dovrebbero essere proporzionate all'età degli studenti, ai risultati attesi e al tipo di conoscenza obiettivo (ad esempio, conoscenza fattuale, concettuale, procedurale o metacognitiva) o problema obiettivo (ad esempio, ben strutturato o mal strutturato);
- I processi di utilizzo dovrebbero garantire il coinvolgimento interattivo degli esseri umani con la GenIA e il pensiero di ordine superiore, nonché la responsabilità umana per le decisioni legate all'accuratezza dei contenuti generati dall'IA, alle strategie di insegnamento o di ricerca e al loro impatto sui comportamenti umani.

5.3 Progettare congiuntamente l'uso della GenIA nell'istruzione e nella ricerca

L'uso della GenIA nell'istruzione e nella ricerca non dovrebbe essere imposto con un approccio dall'alto verso il basso né guidato da logiche commerciali. Invece, il suo uso sicuro ed efficace dovrebbe essere progettato congiuntamente da insegnanti, studenti e ricercatori. È anche necessario un processo strutturato di sperimentazione e valutazione per esaminare l'efficacia e l'impatto a lungo termine dei diversi utilizzi.

Per agevolare la co-progettazione raccomandata, questa Guida propone un quadro composto dalle seguenti sei prospettive per consolidare interazioni pedagogicamente appropriate e la priorità dell'agire umano:

- Aree appropriate di conoscenza o problemi;
- Risultati attesi;
- Strumenti GenIA appropriati e vantaggi comparativi;
- Requisiti per gli utenti;
- Metodi pedagogici richiesti e prompt di esempio;
- Rischi etici.

Questa sezione fornisce esempi di come un processo di co-progettazione nell'uso di GenIA possa informare le pratiche di ricerca, assistere nell'insegnamento, fornire orientamenti per l'acquisizione autonoma di competenze fondamentali, agevolare il pensiero di ordine superiore e sostenere gli studenti con esigenze speciali. Questi esempi rappresentano solo la punta dell'iceberg del crescente numero di ambiti in cui la GenIA può avere un potenziale.

5.3.1 Generative AI per la ricerca

I modelli GenIA hanno dimostrato il loro potenziale nell'ampliare le prospettive sugli schemi di ricerca e nell'arricchire l'esplorazione dei dati e le revisioni della letteratura (vedi Tabella 3). Sebbene possa emergere una vasta gamma di casi d'uso, è necessaria una ricerca innovativa per definire il

potenziale dominio di problemi di ricerca e risultati attesi, per dimostrare l'efficacia e l'accuratezza e per garantire che la capacità umana di agire e prendere decisioni e nel comprendere il mondo reale attraverso la ricerca non venga compromessa dall'uso di strumenti AI.

TABELLA 3: Co-progettare l'uso di GenAI per la ricerca

Utilizzi potenziali ma non comprovati	Ambiti appropriati di conoscenza o problemi.	Risultati attesi	Strumenti GenAI appropriati e vantaggi comparativi	Requisiti per gli utenti	Metodi pedagogici umani richiesti e esempi di prompt	Rischi possibili
Consulente AI per gli schemi di ricerca	Potrebbe essere utile in ambiti di ricerca ben strutturati	Sviluppo e risposta a domande di ricerca, suggerendo metodologie adeguate. Trasformazione potenziale: Coach 1:1 per la pianificazione della ricerca.	A partire dalla lista nella Sezione 1.2, valutare se gli strumenti GenAI sono localmente accessibili, open source, rigorosamente testati o convalidati dalle autorità. Inoltre, considerare i vantaggi e le sfide di qualsiasi particolare strumento GenAI e assicurarsi che affronti adeguatamente specifiche esigenze umane	L'utente deve avere una comprensione di base dell'argomento/i; dovrebbe inoltre sviluppare la capacità di verificare le informazioni ed essere in grado di individuare citazioni di articoli di ricerca o testi inesistenti e di rispondere a domande	Idee di base per la definizione dei problemi di ricerca (ad esempio pubblico, target, questioni, contesto), nonché metodologie, risultati attesi e formati. Scrivere 10 potenziali domande di ricerca per [argomento x] e classificarle in base all'importanza per [il campo di ricerca y].	Bisogna stare attenti all'alto rischio che l'IA generativa inventi informazioni (come pubblicazioni di ricerca inesistenti) e al rischio che gli utenti siano tentati di copiare e incollare schemi di ricerca generati dall'IA, il che potrebbe ridurre le opportunità di apprendimento per utenti alle prime armi
Esploratore generativo di dati e revisore di letteratura	Potrebbe essere utile in ambiti di ricerca poco strutturati	Raccolta automatica di informazioni, esplorazione di una vasta gamma di dati, proposta di bozze di revisioni bibliografiche e automazione di alcune parti dell'interpretazione dei dati. Trasformazione potenziale: Formatori AI per l'esplorazione dei dati e le revisioni bibliografiche	Iniziando con l'elenco nella Sezione 1.2, valutare se gli strumenti GenAI sono accessibili a livello locale, open source, sottoposti a test rigorosi o convalidati dalle autorità. Inoltre, considerare i vantaggi e le sfide di qualsiasi strumento GenAI specifico e assicuratevi che affronti adeguatamente le esigenze specifiche dell'umanità	Gli utenti devono avere una solida conoscenza delle metodologie e delle tecniche per analizzare i dati	Definizioni progressive dei problemi, dell'ambito dei dati e delle fonti di letteratura, delle metodologie utilizzate per l'esplorazione dei dati e le revisioni bibliografiche, e i risultati attesi e i loro formati	È necessario prestare attenzione alle informazioni fabbricate dall'IA generativa, alla gestione impropria dei dati, alle possibili violazioni della privacy, al profilamento non autorizzato e ai pregiudizi di genere. Bisogna essere all'erta riguardo alla diffusione delle norme dominanti e alla minaccia che queste rappresentano per norme alternative e opinioni plurali.

5.3.2 Generative AI per agevolare l'insegnamento.

Sia l'uso di piattaforme GenIA generali che la progettazione di strumenti GenIA educativi specifici dovrebbero essere progettati per migliorare la comprensione degli insegnanti della loro materia e delle metodologie di insegnamento, anche attraverso la co-progettazione - da parte degli insegnanti insieme all'IA - di piani di lezione, pacchetti di corsi o interi curricula. Gli assistenti insegnanti conversazionali assistiti dalla GenIA o i 'gemelli generativi degli assistenti insegnanti' preaddestrati basati su dati di insegnanti esperti e librerie sono stati testati in alcune istituzioni scolastiche e potrebbero avere un potenziale sconosciuto così come rischi etici inesplorati. I processi di applicazione pratica e ulteriori iterazioni di questi modelli devono ancora essere attentamente valutati attraverso il quadro raccomandato in questa Guida e garantiti dalla supervisione umana, come esemplificato nella Tabella 4.

TABELLA 4: Co-progettare l'uso di GenAI per supportare gli insegnanti e l'insegnamento

Utilizzi potenziali ma non comprovati	Ambiti appropriati di conoscenza o problemi.	Risultati attesi	Strumenti GenAI appropriati e vantaggi comparativi	Requisiti per gli utenti	Metodi pedagogici umani richiesti e esempi di prompt	Rischi possibili
Progettista curricolare o di corsi	Conoscenza concettuale su determinati argomenti didattici e conoscenza procedurale sulle metodologie di insegnamento.	Assistenza nel processo di progettazione del curriculum e delle lezioni, compresa la delineazione o l'estensione delle vedute su aree chiave dell'argomento preso in considerazione, e definizione della struttura del curriculum. Potrebbe anche aiutare gli insegnanti a preparare test ed esami offrendo esempi di domande e criteri di valutazione. Trasformazione potenziale: Curriculum generato dall'IA	Iniziando con l'elenco nella Sezione 1.3, valutate se gli strumenti GenAI sono accessibili a livello locale, open source, sottoposti a test rigorosi o convalidati dalle autorità. Inoltre, considerate i vantaggi e le sfide di qualsiasi strumento GenAI specifico e assicuratevi che affronti adeguatamente le esigenze specifiche dell'umanità.	Gli insegnanti devono comprendere e specificare con attenzione ciò che desiderano che il curriculum, i corsi, le lezioni o i test coprano e raggiungano, se desiderano affrontare conoscenze procedurali o concettuali e quale teoria dell'insegnamento desiderano applicare	Domande rivolte a GenAI riguardo alla suggerimento della struttura e degli esempi di conoscenze fattuali su argomenti, alla proposta di metodi e processi di insegnamento su argomenti o problemi, o alla creazione di pacchetti di corsi o piani di lezione basati su argomenti e formattazione. I progettisti curricolari umani devono verificare la correttezza delle conoscenze fattuali e verificare l'adeguatezza dei pacchetti di corsi suggeriti	Il rischio che GenAI imponga norme dominanti e metodi pedagogici è elevato. Potrebbe involontariamente perpetuare pratiche escludenti a vantaggio di gruppi già ricchi di dati e rafforzare le disuguaglianze nell'accesso a opportunità educative rilevanti e di alta qualità per gruppi con scarsità di dati.

Chatbot generativo come assistente didattico	Conoscenza concettuale in diverse aree in problemi ben strutturati	Fornire supporto personalizzato, rispondere a domande ed individuare risorse. Trasformazione potenziale: Gemelli generativi degli assistenti degli insegnanti	Iniziando con l'elenco nella Sezione 1.2, valutate se gli strumenti GenAI sono accessibili a livello locale, open source, sottoposti a test rigorosi o convalidati dalle autorità. Inoltre, prendete in considerazione i vantaggi e le sfide di ogni particolare strumento GenAI e assicuratevi che affronti adeguatamente le specifiche esigenze umane	Supporta gli insegnanti ma si rivolge direttamente agli studenti, il che richiede che gli studenti abbiano una conoscenza precedente, abilità e competenze metacognitive sufficienti per verificare gli output dell'IA generativa e rilevare eventuali informazioni errate. Di conseguenza, potrebbe essere più adatto agli studenti dell'istruzione superiore	Richiede agli insegnanti di comprendere chiaramente i problemi, monitorare la conversazione e aiutare gli studenti a verificare risposte dubbie fornite dall'IA generativa	In base alle attuali capacità dei modelli GenAI, le istituzioni educative devono garantire la supervisione umana delle risposte fornite dagli strumenti GenAI, rimanendo all'erta del rischio di disinformazione. Ciò potrebbe anche limitare l'accesso degli studenti alla guida e al supporto umani, ostacolando lo sviluppo di una solida relazione insegnante-studente, il che è particolarmente preoccupante per i bambini.
--	--	--	---	--	--	--

5.3.3 La GenIA generativa come coach 1:1 per l'acquisizione autonoma di competenze fondamentali.

Sebbene il pensiero di ordine superiore e la creatività stiano attirando sempre più attenzione nella definizione degli obiettivi di apprendimento, non c'è dubbio sull'importanza delle competenze fondamentali nello sviluppo psicologico dei bambini e nella progressione delle competenze. Tra un ampio spettro di abilità, queste competenze fondamentali includono l'ascolto, la pronuncia e la scrittura di una lingua madre o straniera, nonché la numerazione di base, l'arte e la programmazione. Promuovere l'apprendimento attraverso l'azione pratica e l'esperienza diretta non dovrebbe essere considerato un metodo pedagogico obsoleto; piuttosto, dovrebbe essere rinvigorito e potenziato con tecnologie GenIA per favorire l'allenamento autonomo degli studenti alle nuove competenze fondamentali. Se guidati da principi etici e pedagogici, gli strumenti GenIA hanno il potenziale per diventare coach 1:1 per questa pratica autonoma, come illustrato nella Tabella 5.

TABELLA 5: Co-progettare l'uso di GenAI come coach 1:1 per l'acquisizione autonoma di competenze fondamentali nell'apprendimento delle lingue e delle arti

Utilizzi potenziali ma non comprovati	Ambiti appropriati di conoscenza o problemi.	Risultati attesi	Strumenti GenAI appropriati e vantaggi comparativi	Requisiti per gli utenti	Metodi pedagogici umani richiesti e esempi di prompt	Rischi possibili
---------------------------------------	--	------------------	--	--------------------------	--	------------------

Coach di competenze linguistiche 1:1.	Apprendimento delle lingue, compresa la pratica conversazionale	<p>Coinvolgere gli studenti nella pratica conversazionale per aiutarli a migliorare le competenze di ascolto, parola e scrittura, offrendo feedback, correzioni e modellizzazione della lingua madre o della lingua straniera. Aiutare gli studenti a migliorare le loro competenze di scrittura.</p> <p>Trasformazione potenziale: Tutorial linguistici 1:1 a livello principiante</p>	Iniziando con l'elenco nella Sezione 1.3, valutate se gli strumenti GenAI sono accessibili a livello locale, open source, sottoposti a test rigorosi o convalidati dalle autorità. Inoltre, prendete in considerazione i vantaggi e le sfide di qualsiasi particolare strumento GenAI e assicuratevi che affronti adeguatamente le specifiche esigenze umane	<p>Potrebbe essere stabilito un limite di età per le conversazioni indipendenti, considerando le risposte culturalmente insensibili o inadatte all'età fornite dai sistemi GenAI.</p> <p>Lo studente deve avere una motivazione intrinseca iniziale per impegnarsi in una conversazione con un sistema AI; dovrebbe inoltre essere in grado di adottare un approccio critico alle suggerimenti del GenAI e verificarne l'accuratezza</p>	<p>Quando si utilizzano piattaforme GenAI generali, gli insegnanti possono guidare gli studenti nell'interagire con gli strumenti GenAI per richiedere feedback per il miglioramento, la correzione della pronuncia o esempi di scrittura.</p> <p>Ad esempio: Coinvolgimi in una conversazione nella lingua [x], aiutandomi a migliorare continuamente. Suggestiscimi alcune idee per aiutarmi a scrivere su [argomento x].</p>	È necessario rimanere vigili riguardo a linguaggio culturalmente insensibile o inesatto dal punto di vista contestuale e l'involontaria perpetuazione di stereotipi o pregiudizi culturali. Senza adeguate strategie pedagogiche per stimolare le motivazioni intrinseche degli studenti, potrebbe limitare la creatività e l'originalità dei bambini, portando a scritture standardizzate. Potrebbe inoltre limitare le opportunità di interazioni nella vita reale, di espressione pluralistica e di pensiero critico.
Coach artistico 1:1	Competenze tecniche in ambiti artistici come musica e disegno	<p>Fornire supporto personalizzato, rispondere a domande ed individuare risorse.</p> <p>Trasformazione potenziale: Insegnante d'arte 1:1 a livelli introduttivi</p>	Iniziando con l'elenco nella Sezione 1.3.2, valutate se gli strumenti GenAI sono accessibili a livello locale, open source, sottoposti a test rigorosi o convalidati dalle autorità. Inoltre, prendete in considerazione i vantaggi e le sfide di qualsiasi particolare strumento GenAI e assicuratevi che affronti adeguatamente le specifiche esigenze umane	Gli studenti devono avere alcuni obiettivi iniziali per la creazione di opere d'arte o musica, una comprensione fondamentale degli elementi chiave del dominio dell'arte o della musica e capacità di base per analizzare le opere d'arte o le composizioni musicali	<p>Gli insegnanti dovrebbero chiedere agli studenti di confrontare le tecniche artistiche degli strumenti AI con le proprie creazioni. Gli insegnanti devono incoraggiare gli studenti a sviluppare e applicare la propria immaginazione e creatività, che GenAI non può sostituire.</p> <p>Esempio di stimolo: Suggestiscimi alcune idee per ispirarmi a creare un'immagine su [argomenti/idee].</p>	Potrebbe esporre i bambini a contenuti inappropriati o offensivi, il che potrebbe violare il loro diritto alla protezione e al benessere. Gli strumenti GenAI aumentano il rischio di impedire agli studenti di sviluppare la loro immaginazione e creatività

Coach 1:1 per la programmazione o l'aritmetica	Conoscenze e competenze di programmazione concettuale a livello introduttivo. Potrebbe essere applicato anche all'apprendimento della matematica di base	Supportare l'apprendimento a proprio ritmo di conoscenze e competenze di base di programmazione, individuare errori nel codice degli studenti e fornire feedback immediato, e adattare le risposte alle domande. Trasformazione potenziale: Insegnante di programmazione 1:1 a livello introduttivo	Iniziando con l'elenco nella Sezione 1.3, valutate se gli strumenti GenAI sono accessibili a livello locale, open source, sottoposti a test rigorosi o convalidati da autorità. Inoltre, considerate i vantaggi e le sfide di qualsiasi particolare strumento GenAI e assicuratevi che affronti adeguatamente le specifiche esigenze umane.	Trova e definisci un problema e progetta algoritmi per risolvere il problema rimangono gli aspetti centrali dell'apprendimento della programmazione e del coding. Gli studenti devono avere la motivazione intrinseca per utilizzare il coding, insieme a una conoscenza e abilità di base nell'uso del linguaggio di programmazione	I docenti dovrebbero trasmettere conoscenze e promuovere competenze di base e ispirare gli studenti a utilizzare il pensiero computazionale e la programmazione per risolvere problemi, anche attraverso la codifica collaborativa. Esempio di stimolo: Suggestisci alcune idee insolite per la codifica	L'accuratezza dei feedback e delle suggerimenti rimane un problema, poiché GenAI non sarà sempre corretto. C'è un alto rischio che gli strumenti GenAI impediscano agli studenti di sviluppare competenze e abilità nel pensiero computazionale e nell'abilità di trovare e definire problemi significativi per la programmazione
--	--	---	---	--	--	---

5.3.4 Generative AI per agevolare l'inchiesta o l'apprendimento basato su progetti

Se non utilizzati con uno scopo ben preciso per facilitare il pensiero di ordine superiore o la creatività, gli strumenti GenAI tendono a incoraggiare il plagio o a produrre uscite superficiali simili a "simulazioni stocastiche". Tuttavia, dato che i modelli GenAI sono stati addestrati su dati su larga scala, hanno il potenziale per agire come avversari in dialoghi socratici o come assistenti di ricerca nell'apprendimento basato su progetti. Questi potenziali possono però essere sfruttati solo attraverso processi di progettazione dell'istruzione/apprendimento che mirano a stimolare la capacità di riflettere criticamente, analizzare, sintetizzare, valutare e applicare le conoscenze in modi che vanno oltre la memorizzazione o la comprensione superficiale, come illustrato nella Tabella 6.

TABELLA 6: Co-progettare utilizzi di GenAI per facilitare l'apprendimento basato sull'inchiesta o sui progetti

Utilizzi potenziali ma non comprovati	Ambiti appropriati di conoscenza o problemi.	Risultati attesi	Strumenti GenAI appropriati e vantaggi comparativi	Requisiti per gli utenti	Metodi pedagogici umani richiesti e esempi di prompt	Rischi possibili
---------------------------------------	--	------------------	--	--------------------------	--	------------------

Sfida socratica	Problemi poco strutturati.	<p>Coinvolgere gli studenti in un dialogo che ricorda l'interrogazione socratica delle conoscenze pregresse, portando alla scoperta di nuove conoscenze o a una comprensione più profonda.</p> <p>Trasformazione potenziale: Opponente socratico 1:1.</p>	<p>Iniziando con l'elenco nella Sezione 1.3, valutate se gli specifici strumenti GenAI sono accessibili a livello locale, open source, sottoposti a test rigorosi e convalidati dalle autorità. Inoltre, prendete in considerazione i vantaggi e le sfide di qualsiasi strumento GenAI specifico e assicuratevi che affronti adeguatamente le specifiche esigenze umane</p>	<p>Lo studente deve aver raggiunto l'età che gli consente di condurre conversazioni indipendenti con gli strumenti GenAI. Gli studenti devono avere conoscenze e abilità pregresse per verificare se gli argomenti e le informazioni presentate sono corretti</p>	<p>Gli insegnanti possono aiutare a preparare un elenco di domande progressivamente più approfondite come esempi per gli studenti da adattare in prompt. Gli studenti possono anche iniziare con un prompt generale come 'Coinvolgimi in un dialogo socratico per aiutarmi a assumere una prospettiva critica su [argomento x]' e quindi approfondire gradualmente il dialogo attraverso prompt sempre più raffinati.</p>	<p>Gli attuali strumenti GenAI potrebbero generare risposte simili o standard che limitano l'esposizione degli studenti a punti di vista diversi e prospettive alternative, portando a un effetto eco-camera e ostacolando lo sviluppo del pensiero indipendente</p>
-----------------	----------------------------	---	---	---	---	--

<p>Consulente per l'apprendimento basato su progetti</p>	<p>Problemi di ricerca poco strutturati in scienze o studi sociali</p>	<p>Supportare la creazione di conoscenza aiutando gli studenti a condurre l'apprendimento basato su progetti. Ciò include il GenAI che svolge un ruolo simile a quello del consulente di ricerca descritto nella Tabella 3.</p> <p>Trasformazione potenziale: Coach per l'apprendimento basato su progetti 1:1.</p>	<p>Iniziando con l'elenco nella Sezione 1.3, valutate se gli strumenti GenAI sono accessibili a livello locale, open source, sottoposti a test rigorosi o convalidati dalle autorità. Inoltre, considerate i vantaggi e le sfide di qualsiasi strumento GenAI specifico e assicuratevi che affronti adeguatamente le specifiche esigenze umane</p>	<p>Gli studenti potrebbero agire come ricercatori junior nella pianificazione e nell'attuazione dell'apprendimento basato su progetti. Gli studenti devono essere abbastanza grandi per l'uso indipendente delle piattaforme GenAI. Gli studenti devono avere la motivazione e la capacità di partecipare a attività di apprendimento basate su progetti autodirette, in modo che non siano tentati di copiare passivamente e incollare le risposte fornite dagli strumenti GenAI.</p>	<p>Gli insegnanti guidano gli studenti a chiedere a GenAI di fornire idee di base per la definizione dei problemi di ricerca, come suggerito in 4.3.1. Gli studenti da soli e/o in gruppo utilizzano gli strumenti GenAI per condurre revisioni della letteratura, raccogliere e elaborare dati e creare relazioni</p>	<p>Gli studenti senza una solida conoscenza pregressa e la capacità necessaria per verificare l'accuratezza delle risposte potrebbero essere ingannati dalle informazioni fornite dagli strumenti GenAI. Ciò potrebbe anche limitare le discussioni e le interazioni degli studenti con i pari e ridurre le opportunità di apprendimento collaborativo, potenzialmente danneggiando il loro sviluppo sociale</p>
--	--	---	--	--	--	--

5.3.5 Intelligenza Artificiale Generativa per sostenere gli studenti con bisogni speciali.

Teoricamente, i modelli GenIA hanno il potenziale per aiutare gli studenti con disabilità uditive o visive. Le pratiche emergenti includono sottotitoli o didascalie abilitate dalla GenIA per studenti sordi o con problemi di udito e descrizioni audio generate dalla GenIA per studenti ipovedenti. I modelli GenIA possono anche convertire il testo in voce e viceversa per consentire alle persone con disabilità visive, uditive o del linguaggio di accedere ai contenuti, fare domande e comunicare con i loro pari. Tuttavia, questa funzione non è ancora stata sfruttata su larga scala. Secondo l'indagine menzionata in precedenza condotta dall'UNESCO nel 2023 sull'uso dell'IA nei sistemi educativi governativi, solo quattro paesi (Cina, Giordania, Malaysia e Qatar) hanno segnalato che le loro agenzie governative avevano convalidato e raccomandato strumenti assistiti dall'IA per sostenere l'accesso inclusivo per gli studenti con disabilità (UNESCO, 2023c).

C'è anche una tendenza verso gli sviluppi dei modelli GenIA addestrati per supportare gli studenti nell'uso delle proprie lingue, comprese lingue minoritarie e indigene, per apprendere e comunicare. Ad esempio, PaLM 2, la next-generation LLM di Google, è addestrata su dati multilingue paralleli che coprono centinaia di lingue sotto forma di coppie di testo di origine e di destinazione. L'inclusione di dati multilingue paralleli è progettata per migliorare ulteriormente la capacità del modello di comprendere e generare testo multilingue (Google, 2023b).

Fornendo traduzioni in tempo reale, parafrasi e correzione automatica, gli strumenti GenAI hanno il potenziale per aiutare gli studenti che utilizzano lingue minoritarie a comunicare idee e migliorare la loro collaborazione con coetanei provenienti da diverse realtà linguistiche. Tuttavia, ciò avverrà su vasta scala solo con una progettazione mirata.

Infine, è stato anche suggerito che i sistemi GenAI hanno il potenziale per effettuare diagnosi basate su conversazioni, identificando problemi psicologici o emotivi-sociali, così come difficoltà di apprendimento. Tuttavia, rimangono poche prove che questo approccio sia efficace o sicuro, e comunque qualsiasi diagnosi al riguardo richiederebbe l'interpretazione da parte di professionisti qualificati.

TABELLA 7: Co-progettazione di usi di GenAI per supportare gli studenti con bisogni speciali

Utilizzi potenziali ma non comprovati	Ambiti appropriati di conoscenza o problemi.	Risultati attesi	Strumenti GenAI appropriati e vantaggi comparativi	Requisiti per gli utenti	Metodi pedagogici umani richiesti e esempi di prompt	Rischi possibili
Diagnosi conversazionale delle difficoltà di apprendimento	Questo potrebbe essere utile per gli studenti che affrontano difficoltà di apprendimento causate da problemi psicologici, sociali o emotivi.	Usando l'interazione in linguaggio naturale per identificare le esigenze degli studenti con problemi psicologici, sociali o emotivi o con difficoltà di apprendimento, al fine di fornire loro un supporto o un'istruzione adeguati. Trasformazione potenziale: Consulente primario 1:1 per studenti con problemi sociali o emotivi o con difficoltà di apprendimento	Oltre agli strumenti GenAI generali, cerca chatbot basati su GenAI. Valuta se sono accessibili localmente, open source, rigorosamente testati o validati dalle autorità. Considera inoltre *i vantaggi e le sfide di uno strumento GenAI specifico, e *assicurati che affronti adeguatamente le esigenze umane specifiche.	Gli insegnanti o gli specialisti che lavorano con questo gruppo di studenti dovranno assicurarsi che i consigli principali suggeriti dal sistema GenAI siano accurati.	Gli insegnanti o i facilitatori devono creare ambienti confortevoli per coinvolgere lo studente in una conversazione al fine di diagnosticare problemi psicologici, sociali o emotivi, o difficoltà di apprendimento.	Potrebbero erroneamente diagnosticare erroneamente le sfide specifiche dello studente, portando alla fornitura del supporto sbagliato.
Utilizzi potenziali ma non comprovati	Ambiti appropriati di conoscenza o problemi.	Risultati attesi	Strumenti GenAI appropriati e vantaggi comparativi	Requisiti per gli utenti	Metodi pedagogici umani richiesti e esempi di prompt	Rischi possibili

<p>Strumenti di accessibilità basati sull'intelligenza artificiale</p>	<p>Permette agli studenti con disabilità uditive o visive di accedere a una gamma più ampia di contenuti, migliorando così la qualità del loro apprendimento.</p>	<p>Soddisfare le esigenze di accesso degli studenti e supportare l'acquisizione di conoscenze disciplinari fornendo sottotitolazione abilitata da GenAI e/o interpretazione in lingua dei segni per contenuti audio o video e descrizioni audio per testi o altro materiale visivo.</p> <p>Trasformazione potenziale: Ausili linguistici personalizzati 1:1 basati su AI</p>	<p>Oltre agli strumenti GenAI generali, cerca generatori di sottotitoli e descrizioni audio basati su AI pertinenti e affidabili.</p> <p>Valuta se sono accessibili localmente, open source, rigorosamente testati o validati dalle autorità.</p> <p>Considera inoltre</p> <ul style="list-style-type: none"> *i vantaggi e le sfide di qualsiasi strumento GenAI specifico, e *assicurati che affronti adeguatamente le esigenze umane specifiche. 	<p>Gli insegnanti o i facilitatori devono aiutare gli studenti ad accedere e imparare a utilizzare gli strumenti GenAI. Devono inoltre assicurarsi che gli output degli strumenti supportino effettivamente questi studenti e non rafforzino le sfide e i pregiudizi che affrontano.</p>	<p>È necessario testare l'accessibilità di piattaforme o strumenti per identificare e risolvere i problemi di accessibilità prima che vengano utilizzati. Gli strumenti GenAI possono solo fornire accesso ai contenuti, quindi gli insegnanti e i facilitatori dovrebbero concentrarsi sul miglioramento della qualità dell'apprendimento e del benessere sociale. Gli insegnanti e i facilitatori devono insegnare agli studenti a creare prompt vocali o testuali in base alle loro capacità.</p>	<p>Le didascalie o le descrizioni audio prodotte dalle piattaforme GenAI che non sono progettate specificamente per supportare la vista o l'udito sono spesso inaccurate e possono fuorviare gli studenti con bisogni speciali. Questi strumenti potrebbero inavvertitamente rafforzare i pregiudizi esistenti.</p>
--	---	--	---	--	--	---

<p>Amplificatore generativo per studenti marginalizzati</p>	<p>Potrebbe essere utile per gli studenti provenienti da contesti linguistici o culturali minoritari per esprimere e amplificare le loro voci, per partecipare online e per condurre studi sociali collaborativi.</p>	<p>Fornire traduzioni in tempo reale, parafrasi e correzione automatica della scrittura per supportare gli studenti dei gruppi marginalizzati a utilizzare le proprie lingue per comunicare con i loro coetanei di background linguistici diversi.</p> <p>Trasformazione potenziale: LLM inclusivi per studenti marginalizzati</p>	<p>Un esempio specifico da considerare è PaLM 2. Valuta se gli strumenti GenAI sono accessibili localmente, open source, rigorosamente testati o validati dalle autorità. Considera inoltre i vantaggi e le sfide di qualsiasi strumento GenAI specifico e assicurati che affronti adeguatamente le esigenze umane specifiche.</p>	<p>Gli studenti dovrebbero avere conoscenze o opinioni significative sull'argomento della conversazione o dello studio collaborativo. Devono essere in grado di dare contributi responsabili e di evitare il linguaggio d'odio.</p>	<p>Gli insegnanti dovrebbero progettare studi e compiti di scrittura per gli studenti su temi sociali o culturali, oppure organizzare seminari online o collaborazioni interculturali per stimolare gli studenti a generare idee e condividere opinioni.</p>	<p>È necessario identificare e correggere gli errori nelle traduzioni e nelle parafrasi dell'AI che possono causare incomprensioni interculturali. Questo uso può fornire opportunità agli studenti marginalizzati di amplificare le loro voci, ma non toccherà la causa principale della povertà di dati e pertanto non potrà decolonizzare gli strumenti di AI.</p>
---	---	--	--	---	--	---

6. La GenIA e il futuro dell'istruzione e della ricerca.

Le tecnologie GenIA stanno ancora rapidamente evolvendo e probabilmente avranno un profondo impatto sull'istruzione e sulla ricerca, ma i possibili sviluppi sono ancora da comprendere appieno. Le sue implicazioni a lungo termine per l'istruzione e la ricerca necessitano comunque di immediata attenzione e di ulteriori approfondimenti.

6.1 Questioni etiche inesplorate

Gli strumenti GenIA sempre più sofisticati solleveranno ulteriori preoccupazioni etiche che devono essere esaminate in dettaglio. Oltre alle Sezioni 2 e 3, sono necessarie analisi più approfondite e orientate al futuro per rivelare e affrontare questioni etiche inesplorate da almeno le seguenti cinque prospettive:

- **Accesso ed equità:** i sistemi GenIA nell'istruzione potrebbero esacerbare le disparità esistenti nell'accesso alla tecnologia e alle risorse educative, approfondendo ulteriormente le disuguaglianze.
- **Connessione umana:** i sistemi GenIA nell'istruzione potrebbero ridurre l'interazione tra esseri umani e gli aspetti critici dell'apprendimento legati alle emozioni sociali.
- **Sviluppo intellettuale umano:** i sistemi GenIA nell'istruzione potrebbero limitare l'autonomia e l'agire degli studenti fornendo soluzioni predefinite o restringendo la gamma di esperienze di apprendimento possibili. Il loro impatto a lungo termine sullo sviluppo intellettuale dei giovani studenti deve essere indagato.
- **Impatto psicologico:** i sistemi GenIA che simulano le interazioni umane possono avere effetti psicologici sconosciuti sugli studenti, sollevando preoccupazioni sul loro sviluppo cognitivo e benessere emotivo e sulla potenziale manipolazione.
- **Bias nascosti e discriminazione:** con lo sviluppo e l'applicazione di sistemi GenIA sempre più sofisticati nell'istruzione, è probabile che vengano generati nuovi bias e forme di discriminazione basate sui dati di addestramento e sui metodi utilizzati dai modelli, che possono portare a output sconosciuti e potenzialmente dannosi.

6.2 Copyright e proprietà intellettuale

L'emergere della GenIA sta cambiando rapidamente il modo in cui opere scientifiche, artistiche e letterarie vengono create, distribuite e consumate. La copia, la distribuzione o l'uso non autorizzato di opere protette da copyright senza il permesso del titolare dei diritti viola i loro diritti esclusivi e può comportare conseguenze legali. L'addestramento di modelli GenIA è stato accusato di violare il copyright: uno dei casi recenti, la canzone generata dall'IA che presenta 'Drake' e 'The Weeknd' (Abel Tesfaye), ha raggiunto milioni di ascoltatori prima di essere rimossa a causa di una disputa sui diritti d'autore (Coscarelli, 2023). Mentre i nuovi quadri regolamentari intendono richiedere ai fornitori di GenIA di riconoscere e proteggere la proprietà intellettuale dei proprietari dei contenuti utilizzati dal modello, è sempre più difficile determinare la proprietà e l'originalità della straordinaria quantità di opere generate. Questa mancanza di tracciabilità solleva preoccupazioni non solo sulla protezione dei diritti dei creatori e sulla garanzia di una giusta compensazione per i loro contributi intellettuali, ma introduce anche sfide nei contesti educativi su come l'output degli strumenti GenIA possa essere utilizzato in modo responsabile. Ciò potrebbe avere profonde implicazioni per il sistema di ricerca.

6.3 Fonti di contenuto e apprendimento

Gli strumenti GenIA stanno cambiando il modo in cui il contenuto didattico può essere generato e fornito. In futuro, il contenuto generato attraverso conversazioni tra esseri umani e intelligenza artificiale potrebbe diventare una delle principali fonti di produzione di conoscenza. E' probabile che

questo metta ulteriormente in crisi il rapporto che gli studenti hanno con un contenuto educativo basato su risorse certificate, libri di testo e programmi creati e avvalorati da esseri umani. L'aspetto autorevole del testo GenIA potrebbe trarre in inganno giovani studenti che non hanno sufficienti conoscenze pregresse per riconoscere inaccurately o per interrogarlo in modo efficace. C'è poi da chiedersi se lo studio che si basa su contenuti non convalidati debba essere riconosciuto come 'apprendimento'.

La conseguente concentrazione su informazioni di seconda mano aggregate potrebbe anche ridurre le opportunità degli studenti di costruire conoscenze attraverso metodi provati come percepire direttamente ed esperire il mondo reale, imparare dal tentativo ed errore, eseguire esperimenti empirici e sviluppare il buon senso. Potrebbe anche minacciare la costruzione sociale della conoscenza e la promozione dei valori sociali attraverso pratiche collaborative in classe.

6.4 Risposte omogenee rispetto a produzioni diverse e creative.

La GenIA restringe le narrazioni plurali in quanto le risposte generate tendono a rappresentare e rafforzare punti di vista dominanti. La conseguente omogeneizzazione delle conoscenze limita il pensiero pluralistico e creativo. La crescente dipendenza degli insegnanti e degli studenti dagli strumenti GenIA per cercare suggerimenti potrebbe portare alla standardizzazione e alla conformità delle risposte, indebolendo il valore del pensiero indipendente e dell'indagine autodiretta. La potenziale omogeneizzazione dell'espressione in pezzi scritti e opere d'arte può limitare l'immaginazione, la creatività e le prospettive alternative delle espressioni degli studenti. I fornitori di GenIA e gli insegnanti devono considerare fino a che punto EdGPT potrebbe essere sviluppato e utilizzato per promuovere la creatività, la collaborazione, il pensiero critico e altre abilità di pensiero di ordine superiore.

6.5 Ripensare la valutazione e gli obiettivi di apprendimento

Le implicazioni della GenIA per la valutazione vanno ben oltre le preoccupazioni immediate riguardo agli studenti che copiano negli elaborati scritti. Dobbiamo fare i conti con il fatto che la GenIA può produrre articoli e saggi relativamente ben organizzati e opere d'arte impressionanti, e può superare alcuni esami basati sulla conoscenza in determinate materie. Pertanto, è necessario ripensare cosa dovrebbe essere esattamente appreso e a quali scopi, e come l'apprendimento deve essere valutato e convalidato. La discussione critica tra insegnanti, responsabili delle politiche, studenti e altre parti interessate deve prendere in considerazione le seguenti quattro categorie di risultati di apprendimento:

Valori: I valori necessari per garantire il design e l'uso centrato sull'essere umano della tecnologia sono centrali nel ripensare gli obiettivi di apprendimento e la loro valutazione nell'era digitale. Nel riesaminare lo scopo dell'istruzione, i valori che informano il modo in cui la tecnologia si rapporta all'istruzione dovrebbero essere resi espliciti. È attraverso questa lente normativa che gli obiettivi di apprendimento e la loro valutazione e convalida devono essere aggiornati in modo continuativo per rispondere all'uso sempre più diffuso della tecnologia, compresa l'IA, nella società.

Conoscenze e competenze fondamentali: Anche nei settori delle competenze in cui gli strumenti GenIA possono fare meglio degli esseri umani, gli studenti avranno comunque bisogno di solide conoscenze e competenze fondamentali. L'alfabetizzazione di base, la numerazione di base e le competenze di base in ambito scientifico rimarranno fondamentali per l'istruzione futura. La portata e la natura di queste competenze fondamentali dovranno essere regolarmente rivisitate per riflettere gli ambienti sempre più ricchi di IA in cui viviamo.

Abilità di pensiero di ordine superiore: Gli obiettivi di apprendimento dovranno includere le competenze necessarie a sostenere il pensiero di ordine superiore e la risoluzione dei problemi basati sulla collaborazione tra esseri umani e intelligenza artificiale e sull'uso delle produzioni generate da GenIA. Queste possono includere la comprensione dei ruoli delle conoscenze fattuali e concettuali nel fondare il pensiero di ordine superiore e la valutazione critica dei contenuti generati dall'IA.

Competenze professionali necessarie per lavorare con l'IA: Nei settori in cui l'IA può fare meglio degli esseri umani e automatizza le unità di compiti, gli studenti devono avere nuove competenze che consentano loro di sviluppare, gestire e lavorare con gli strumenti GenIA. La ridefinizione degli obiettivi di apprendimento e della valutazione educativa dovrà riflettere le competenze professionali richieste per i nuovi posti di lavoro creati dall'IA.

6.6. Processi di pensiero

La prospettiva più fondamentale sulle implicazioni a lungo termine della GenIA per l'istruzione e la ricerca riguarda ancora la relazione complementare tra l'azione umana e le macchine. Una delle domande chiave è se gli esseri umani possano effettivamente cedere ai sistemi di intelligenza artificiale processi di pensiero e acquisizione di competenze di base, concentrandosi invece sulle competenze di pensiero di ordine superiore basate sulle produzioni fornite dall'IA.

La scrittura, ad esempio, è spesso associata alla strutturazione del pensiero. Con la GenIA, invece di partire da zero per pianificare gli obiettivi, la portata e la struttura di un insieme di idee, gli esseri umani possono ora partire da un'outline ben strutturata fornita dalla GenIA. Alcuni esperti hanno bollato l'uso della GenIA per generare testi in questo modo come "scrivere senza pensare" (Chayka, 2023). Man mano che queste nuove pratiche assistite dalla GenIA diventano più ampiamente adottate, i metodi consolidati per l'acquisizione e la valutazione delle competenze di scrittura dovranno adattarsi. Una possibilità per il futuro è che l'apprendimento della scrittura possa concentrarsi sulla costruzione di competenze nella pianificazione e nella composizione di prompt, nella valutazione critica delle produzioni della GenIA, nel pensiero di ordine superiore, così come nella co-scrittura basata sulle outline della GenIA.

Osservazioni conclusive

Dal punto di vista di un approccio centrato sull'essere umano, gli strumenti AI dovrebbero essere progettati per estendere o potenziare le capacità intellettuali e sociali umane, e non per minarle, entrare in conflitto con loro o sostituirle. Da tempo si prevede che gli strumenti AI possano essere integrati ulteriormente come parte integrante dei mezzi a disposizione degli esseri umani per supportare l'analisi e l'azione per un futuro più inclusivo e sostenibile.

Affinché l'AI diventi una parte affidabile della collaborazione tra esseri umani e macchine - a livello individuale, istituzionale e di sistema - l'approccio centrato sull'essere umano, informato dalla Raccomandazione dell'UNESCO del 2021 sull'Etica dell'AI, deve essere ulteriormente sviluppato e attuato in base alle specifiche caratteristiche delle tecnologie emergenti come la GenIA. Solo in questo modo possiamo garantire che la GenIA diventi uno strumento affidabile per ricercatori, insegnanti e studenti.

Mentre la GenIA dovrebbe essere utilizzata per servire l'istruzione e la ricerca, tutti noi dobbiamo essere consapevoli del fatto che la GenIA potrebbe anche cambiare i sistemi consolidati e le loro basi in questi settori. La trasformazione dell'istruzione e della ricerca che potrebbe essere innescata dalla GenIA, se esiste, dovrebbe essere rigorosamente esaminata e guidata da un approccio centrato sull'essere umano. Solo in questo modo possiamo garantire che il potenziale dell'AI, in particolare, e tutte le altre categorie di tecnologie utilizzate nell'istruzione in senso più ampio, migliorino le capacità umane per costruire un futuro digitale davvero inclusivo.