L'INTELLIGENZA ARTIFICIALE GENERATIVA NELLA SCUOLA TRA DIRITTI FONDAMENTALI E NUOVE FORME DI GOVERNANCE*

di Matteo Angeloni **

Sommario. 1. Il diritto all'istruzione e il quadro normativo applicabile. – 2. L'intelligenza artificiale generativa: natura, rischi e opportunità. – 3. L'IA nella scuola: *case study* comparati.

1. Il diritto all'istruzione e il quadro normativo applicabile. Il diritto all'istruzione occupa, nell'architettura costituzionale italiana, una posizione che trascende la dimensione meramente individuale. L'art. 34 Cost. proclama che «la scuola è aperta a tutti» e sancisce l'obbligatorietà e la gratuità dell'istruzione inferiore, delineando una pretesa che non si esaurisce nel mero accesso fisico all'edificio scolastico, ma richiede alla Repubblica di apprestare condizioni materiali, organizzative e strumentali capaci di rendere quella frequenza effettivamente formativa. L'art. 3, co. 2 Cost. affida alla Repubblica il compito di rimuovere gli ostacoli economici e sociali che limitano di fatto la libertà e l'eguaglianza dei cittadini, orientando l'intera materia verso una dimensione sostanziale e non meramente formale. L'art. 33 Cost., a sua volta, presidia la libertà dell'insegnamento e il pluralismo culturale, ponendo la dimensione formativa al riparo da indebite standardizzazioni[1].

La Corte costituzionale ha progressivamente chiarito che il diritto all'istruzione ha natura di diritto fondamentale a prestazione: non basta rimuovere i divieti formali di accesso, occorre predisporre le condizioni positive che rendano tale accesso reale e il percorso effettivo. Con la sent. n. 215/1987 la Corte cost. ha individuato nelle risorse umane, materiali e organizzative, un nucleo indefettibile di garanzie che il legislatore non può comprimere oltre soglie minime costituzionalmente determinate.

Con la sent. n. 275/2016 ha poi statuito che è la garanzia del nucleo essenziale dei diritti ad incidere sul bilancio, e non l'equilibrio di bilancio a condizionare la doverosa erogazione delle prestazioni. Ne segue che il diritto all'istruzione non coincide con l'accesso formale all'edificio scolastico, ma esige partecipazione effettiva, inclusione e condizioni non discriminatorie lungo tutto il percorso, affinché la scuola operi come fattore di emancipazione e di partecipazione democratica, nel quadro di un ordinamento in cui la libertà di insegnamento e il pluralismo culturale trovano garanzia nell'art. 33 Cost[2].

Questa accezione sostanziale del diritto all'istruzione non è rimasta insensibile alla progressiva digitalizzazione dei processi formativi. Le istituzioni scolastiche hanno adottato strumenti informatici per la gestione e il supporto dell'attività didattica (registri elettronici,

* *Sottoposto a referaggio*.

** Dottoranda di ricerca in Società in mutamento: politiche, diritti e sicurezza – Università della Tuscia.

[1] A. Poggi, *L'istruzione come fattore di partecipazione: le nuove sfide della «complessità»*, in *Rivista AIC*, 1, 2024, 408-425.

[2] E. Rossi, P. Addis, F. Biondi Dal Monte, *La libertà di insegnamento e il diritto all'istruzione nella Costituzione italiana*, in *Osservatorio Costituzionale*, 1, 2016, 1-17.

piattaforme di comunicazione scuola-famiglia, ambienti virtuali di apprendimento) che, pur non incidendo direttamente sul contenuto dell'insegnamento, hanno reso la fruizione effettiva del diritto dipendente dalla disponibilità di dispositivi, connessione e competenze digitali. Già in questa prima fase sono emerse le disuguaglianze materiali che l'art. 3, c. 2 Cost. impone di rimuovere[3]. L'esperienza della pandemia da Covid-19 ha reso strutturale questa dipendenza, sollevando questioni di *governance* e di dipendenza istituzionale dai fornitori tecnologici che non si sono dissolte con il ritorno alla didattica in presenza[4]. Con l'introduzione dei sistemi di intelligenza artificiale generativa questo quadro compie un salto qualitativo: tali sistemi non si limitano a mediare la trasmissione di informazioni esistenti, ma generano contenuti nuovi, incidendo direttamente sulla sostanza del processo formativo. È per questa ragione che la cornice costituzionale appena descritta è il parametro concreto rispetto al quale tale introduzione deve essere valutata: qualunque infrastruttura che incida sulle condizioni reali di partecipazione e sui risultati formativi non è neutra rispetto al diritto all'istruzione.

Il quadro normativo che governa l'adozione di sistemi di intelligenza artificiale in ambito educativo si articola su tre livelli, che si presentano qui in ordine cronologico.

Il primo è il reg. UE n. 2016/679, noto come GDPR[5]. La scuola opera quale titolare del trattamento dei dati e deve garantire che i dati degli studenti siano raccolti e trattati nel rispetto dei principi di finalità, minimizzazione e limitazione della conservazione. Il contesto scolastico introduce profili di attenzione specifici: il trattamento può coinvolgere dati appartenenti alle categorie particolari di cui all'art. 9 GDPR (informazioni su disabilità, bisogni educativi speciali, fragilità personali), la cui raccolta è consentita soltanto in presenza di condizioni di liceità rinforzata. Laddove gli studenti siano minorenni, si aggiunge una tutela specifica che impone criteri di valutazione del rischio particolarmente rigorosi e il consenso esercitato attraverso i titolari della responsabilità genitoriale[6]. Ogni volta che l'adozione di un sistema di intelligenza artificiale comporta un trattamento ad alto rischio per i diritti degli interessati, l'art. 35 GDPR impone la valutazione d'impatto sulla protezione dei dati; nei casi previsti dall'*AI Act*, a questa si aggiunge la valutazione d'impatto sui diritti fondamentali, che con la prima dovrebbe formare un'unica catena integrata di controllo preventivo.

Il secondo livello è la dottrina della cosiddetta legalità algoritmica, elaborata dalla giurisprudenza amministrativa italiana attraverso una serie di pronunce del Consiglio di Stato a partire dal 2019. Secondo tale dottrina, l'utilizzo di un algoritmo non esonera l'istituzione scolastica dall'obbligo di motivare la propria decisione, di renderla conoscibile e di garantirne la contestabilità: il processo decisionale automatizzato deve restare ricostruibile, le regole operative devono essere conoscibili e la responsabilità dell'esito deve poter essere imputata a un soggetto umano identificabile[7]. Da questo principio discendono tre criteri operativi: la «spiegabilità praticabile», quale capacità di fornire informazioni comprensibili sulle logiche del sistema; la *contestabilità*, intesa come possibilità concreta di riesame umano degli esiti con

[3] S. Dexter, A. Francisco, C. Luke Luna, *Five leading-edge K-12 districts' decision-making processes for EdTech innovations*, in *Journal of Educational Administration*, 59(3), 2021, 352–366; N. Sandu, C. Taylor, Digital learning technologies usage during Covid-19 lockdowns, in British Educational Research Journal, 2025.

[4] F. Cornali, *In platform we trust(?): Concepts, problems and a case study about platform education*, in *European Journal of Education*, 59(4), 2024.

[5] Reg. UE n. 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la dir. UE n. 95/46/CE, in G.U.U.E., L 119, 4 maggio 2016.

[6] Garante per la protezione dei dati personali, parere 4 agosto 2025, n. 454 (parere favorevole sullo schema di decreto del MIM recante le *Linee guida per l'introduzione dell'intelligenza artificiale nelle istituzioni scolastiche*, poi adottate con d.m. 9 agosto 2025, n. 166).

[7] Cons. Stato, sez. VI, sent. n. 2270/2019.

implicazioni rilevanti per gli studenti; e la «responsabilità», quale attribuzione chiara della decisione educativa a un soggetto umano identificabile che ne risponde. Tali criteri si raccordano ai requisiti di trasparenza e supervisione umana dell'*AI Act* e non sono ostacoli all'innovazione, ma condizioni della sua legittimità nel contesto di un'istituzione che esercita funzioni pubbliche incidenti sui diritti fondamentali degli studenti.
Il terzo livello è il reg. UE n. 2024/1689, noto come *AI Act*, prima disciplina organica dei sistemi di intelligenza artificiale nell'ordinamento europeo[8]. Il regolamento adotta una classificazione fondata sul rischio: quanto maggiore è il rischio che un sistema pone per i diritti fondamentali, tanto più stringenti sono gli obblighi imposti a chi lo sviluppa e a chi lo adotta. Per il settore dell'istruzione, l'Allegato III individua tra i sistemi ad alto rischio quelli impiegati per valutare le prestazioni degli studenti, determinarne l'accesso a istituzioni educative, orientarne il percorso formativo o rilevarne i comportamenti durante le verifiche[9]. I sistemi ad alto rischio sono soggetti a obblighi stringenti, ossia valutazione di conformità preventiva, trasparenza verso gli utenti, supervisione umana effettiva e adeguatezza dei dati di addestramento, obblighi che si applicano non solo al produttore, ma anche alla scuola che adotta il sistema in qualità di *deployer*: l'istituzione scolastica non può limitarsi a scegliere uno strumento sul mercato ignorando le condizioni della sua utilizzazione lecita.

2. L'intelligenza artificiale generativa: natura, rischi e opportunità. Per valutare le implicazioni giuridiche dell'intelligenza artificiale generativa nella scuola occorre chiarire, in termini non tecnici ma giuridicamente rilevanti, che cosa la distingua dagli strumenti digitali ordinari e dai sistemi algoritmici predittivi già in uso in campo educativo. Un registro elettronico trasmette dati esistenti; un sistema predittivo produce una classificazione o una stima di probabilità sulla base di variabili predefinite dal progettista. Un sistema di intelligenza artificiale generativa fa qualcosa di qualitativamente diverso: «genera» contenuti nuovi, quali testi, spiegazioni, retroazioni ed elaborati, che non preesistevano all'interazione, e lo fa sulla base di una richiesta dell'utente senza che né il produttore né la scuola possano predeterminarne l'esito. È questa capacità generativa, non la sofisticazione tecnologica in sé, a giustificare una riflessione giuridica specifica: perché tali sistemi incidono direttamente sulla sostanza del processo formativo, sulle valutazioni che ne derivano e, per questa via, sull'effettività del diritto all'istruzione nella sua accezione costituzionale.
La prima implicazione giuridica strutturale riguarda l'affidabilità della fonte. I modelli di intelligenza artificiale generativa producono risposte statisticamente plausibili rispetto ai materiali su cui sono stati formati, ma la coerenza stilistica dell'*output* non garantisce l'accuratezza dei contenuti: il sistema non attinge a una base di conoscenza verificata, ma produce testi la cui plausibilità formale è indipendente dalla loro corrispondenza al vero[10]. Il fenomeno delle cosiddette «allucinazioni», vale a dire la produzione di affermazioni false o di citazioni inesistenti con lo stesso tono assertivo delle informazioni corrette, non è una disfunzione accidentale ma una conseguenza strutturale di questo meccanismo. In ambito scolastico le ricadute giuridiche sono dirette: l'integrità degli elaborati diventa difficilmente verificabile, e il processo valutativo perde la sua funzione costituzionale di certificare le competenze effettive dello studente quando il testo consegnato non riflette le conoscenze di

[8] Reg. UE n. 2024/1689 del Parlamento europeo e del Consiglio, del 13 giugno 2024, che stabilisce regole armonizzate sull'intelligenza artificiale, in G.U.U.E., L 2024/1689, 12 luglio 2024.
[9] Reg. UE n. 2024/1689, cit., Allegato III, punto 3, lett. a), b), c) e d).
[10] S. Grassini, *Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT*, *in Educational Settings*, in *Education Sciences*, 13(7), 2023, 692.

chi lo ha prodotto[11].
La seconda implicazione riguarda la distorsione sistematica. I materiali su cui i modelli vengono formati rispecchiano inevitabilmente i pregiudizi e gli squilibri rappresentativi della produzione scritta disponibile: le voci sovrarappresentate sono quelle di lingue europee e nordamericane, di cultura accademica, di condizioni socioeconomiche elevate. Il sistema riproduce, e in taluni casi amplifica, tali squilibri nei contenuti che genera, con il risultato che le valutazioni e le raccomandazioni prodotte possono incorporare forme di discriminazione sistematica nei confronti di studenti appartenenti a gruppi sottorappresentati o a contesti culturali non egemoni[12]. In termini giuridici, questa caratteristica strutturale si traduce in un rischio di violazione del principio di non discriminazione: il *bias* incide sulle opportunità educative in modo opaco, rendendo difficile per lo studente identificare l'errore e contestarlo efficacemente, con una compressione del diritto di difesa che l'art. 22 GDPR e il principio di legalità algoritmica intendono scongiurare.

La terza implicazione è l'opacità del processo generativo. A differenza di un manuale scolastico, che esplicita le proprie fonti e la propria struttura argomentativa, e a differenza di un docente, che può essere interrogato sulle ragioni delle proprie valutazioni, il sistema non rende visibile il percorso che conduce all'output prodotto. In ambito scolastico questa opacità ha due conseguenze giuridiche dirette: compromette lo sviluppo delle capacità critiche che sono tra gli obiettivi fondamentali dell'istruzione, poiché lo studente riceve un prodotto finito senza poterlo valutare o discutere consapevolmente; e ostacola la supervisione umana effettiva che l'*AI Act* impone come condizione di legittimità per i sistemi ad alto rischio, poiché un docente che non comprende come il sistema ha prodotto un esito non è in grado di esercitare un controllo sostanziale né di assumersene la responsabilità[13].
La quarta implicazione riguarda il trattamento dei dati prodotti nell'interazione. Quando uno studente utilizza un sistema di intelligenza artificiale, l'interazione genera e conserva contenuti e metadati che possono consentire inferenze sulle sue difficoltà e condizioni personali. In molti casi tali dati vengono riutilizzati per finalità commerciali o per il perfezionamento del modello, senza che l'utente ne abbia piena consapevolezza. La scuola che adotta un sistema il quale trasferisce dati degli studenti a un operatore esterno non cessa di essere titolare del trattamento: non può sottrarsi alla responsabilità di verificare la conformità di tale trasferimento al quadro normativo, né può aggirare il principio di finalità, in base al quale i dati raccolti nell'ambito di un'attività didattica non possono essere destinati a usi diversi da quelli originari[14].
Sarebbe tuttavia riduttivo descrivere l'intelligenza artificiale generativa esclusivamente come un fattore di rischio. I quattro profili appena descritti sono le variabili che una *governance* responsabile deve presidiare, non argomenti per il rifiuto della tecnologia: la stessa tecnologia, adeguatamente governata, può costituire un potente fattore per ridurre le disparità nell'istruzione. Il nesso con la cornice costituzionale è diretto: là dove la scuola non è strutturalmente in grado di offrire sostegno personalizzato agli studenti in difficoltà, un sistema di intelligenza artificiale ben governato può colmare parzialmente questa lacuna,

[11] M.A. Cardona, R.J. Rodríguez, K. Ishmael, *Artificial Intelligence and the Future of Teaching and Learning: Insights and Recommendations*, 2023 (ERIC n. ED631097).
[12] O.O. Ayeni *et al.*, *AI in education: A review of personalized learning and educational technology*, in *GSC Advanced Research and Reviews*, 18(2), 2024, 261-271.
[13] F. Laurenzi, A. Fontana, G. Camera, A. Barone, V. Bellantone, A. Canuti, G. Malgieri, F.P. Patti, O. Pollicino, S. Quattrocolo, D. Simeoli, P. Sirena (a cura di), *Intelligenza artificiale nella pubblica amministrazione: opportunità e rischi. Atti del convegno*, Milano, 8 febbraio 2024, 2024.
[14] P. Martena, S.D. Riccardis, op. cit.

rimuovendo un ostacolo di fatto all'eguaglianza nel senso dell'art. 3, co. 2 Cost[15]. Può rendere disponibili spiegazioni in lingue diverse per studenti di recente immigrazione, abbattendo le barriere linguistiche che limitano l'effettività del diritto all'istruzione per una delle categorie tradizionalmente più vulnerabili[16]. Può liberare il docente da attività ripetitive, consentendogli di dedicare più tempo alle dimensioni della relazione educativa che richiedono presenza umana e che nessun sistema automatico può sostituire[17]. Può, infine, diventare lo strumento attraverso cui gli studenti sviluppano capacità critiche, imparando a interrogare, verificare e discutere le risposte del sistema anziché accettarle passivamente: un esercizio che corrisponde esattamente alle competenze che l'istruzione è chiamata a coltivare[18]. Come si vedrà attraverso i casi concreti che seguono, è la qualità della *governance*, e non la sofisticazione del sistema, a determinare se l'intelligenza artificiale serve il diritto all'istruzione o lo erode.

3. L'IA nella scuola: case study comparati. I profili sin qui delineati acquistano concretezza giuridica quando si esaminano esperienze già in atto in diversi sistemi educativi. I casi che seguono, tratti da contesti scolastici e universitari internazionali inclusi paesi extra-UE, documentano sia le condizioni alle quali l'intelligenza artificiale rischia di ledere il diritto all'istruzione, sia quelle alle quali può rafforzarne l'effettività con risultati misurabili. Poiché alcune esperienze provengono dalla formazione superiore, è opportuno precisare che il reg. UE n. 2024/1689 e il reg. UE n. 2016/679 si applicano uniformemente a tutti i livelli istituzionali in cui sistemi di intelligenza artificiale siano impiegati per finalità valutative o formative. Ogni caso porta con sé un argomento per la tesi: la legittimità dell'intelligenza artificiale nella scuola dipende dalla qualità della *governance* che la sorregge, non dalla tecnologia adottata in sé.

Il primo profilo di rischio, quello della distorsione algoritmica e della sua ricaduta sul principio di eguaglianza, è documentato sistematicamente dalla letteratura sui sistemi predittivi in ambito educativo. La rassegna di Baker e Hawn ha rilevato che gli effetti di genere nei modelli di previsione del successo e dell'abbandono scolastico risultano altamente inconsistenti a seconda del contesto: in alcuni studi il sistema penalizza sistematicamente le studentesse, in altri gli studenti maschi, e in entrambi i casi la distorsione emerge in modo opaco, senza che gli utenti del sistema ne siano consapevoli. La stessa architettura algoritmica, addestrata su popolazioni diverse o applicata in contesti socioculturali differenti, può produrre effetti discriminatori opposti, rendendo impossibile trasferire una valutazione di equità da un distretto a un altro senza verifiche specifiche[19]. Questo risultato è fondamentale: un sistema predittivo può superare le soglie di accuratezza tecnica richieste e al contempo produrre effetti sistematicamente svantaggiosi per determinati gruppi di studenti, in violazione del principio di non discriminazione. La valutazione di equità non è dunque un adempimento aggiuntivo rispetto alla valutazione di conformità tecnica: è una componente necessaria della valutazione d'impatto che l'*AI Act* impone per i sistemi ad alto rischio, e il suo esito non è prevedibile a priori né trasferibile da un contesto all'altro senza verifiche specifiche.

Sul medesimo piano dei sistemi predittivi, un secondo caso riguarda la Open University britannica. Nel 2017 Kuzilek, Hlosta e Zdrahal hanno pubblicato l'*Open University Learning*

[15] M.A. Cardona, R.J. Rodríguez, K. Ishmael, op. cit.; O.O. Ayeni *et al.*, op. cit.
[16] S. Pasta, *IA ed educazione. Oltre il determinismo algoritmico*, in V. Corrado, S. Pasta (a cura di), *Intelligenza artificiale e sapienza del cuore*, Brescia, 2024, 147-163.
[17] S. Leaton Gray, op. cit.; F. Cornali, op. cit.
[18] S. Grassini, op. cit.; K. McConvey, S. Guha, A. Kuzminykh, op. cit.
[19] R.S. Baker, A. Hawn, *Algorithmic Bias in Education*, in *International Journal of Artificial Intelligence in Education*, 32, 2022, 1052-1092.

Analytics Dataset (OULAD), raccogliendo dati relativi agli anni accademici 2013–2014: l'analisi condotta su oltre 32.000 studenti iscritti a corsi scientifici e di scienze sociali ha evidenziato il peso determinante assunto dai dati demografici in assenza di sufficienti dati di interazione con la piattaforma virtuale di apprendimento[20]. In pratica, quando uno studente si era appena iscritto e non aveva ancora prodotto attività registrabili, il sistema tendeva a classificarlo sulla base di variabili come l'area geografica di provenienza, il titolo di studio dei genitori o la disabilità dichiarata, piuttosto che sulla base delle sue capacità effettive di apprendimento: il rischio, in altri termini, era di classificare uno studente in base alla propria origine piuttosto che alle proprie competenze. In termini giuridici, il caso illustra la distinzione, rilevante ai fini dell'art. 22 GDPR, tra un sistema che assiste il docente nell'individuazione di studenti che potrebbero beneficiare di supporto aggiuntivo, compatibile con il principio di supervisione umana effettiva, e un sistema che produce una classificazione automatica vincolante sul percorso dello studente, soggetta invece a garanzie rinforzate.

Il problema del consenso sostanziale al trattamento algoritmico trova una concreta illustrazione in un'esperienza condotta alla ESCP Business School di Parigi nel marzo 2026, nell'ambito dell'insegnamento rivolto a studenti di Executive MBA e di Master in Management. Tali studenti, abituati nella propria attività professionale a utilizzare sistemi algoritmici per la valutazione del personale, sono stati sottoposti essi stessi alla valutazione di un sistema di intelligenza artificiale configurato come assistente didattico e denominato Ember (sviluppato a partire da ChatGPT), incaricato di attribuire punteggi a presentazioni e memorandum nel rispetto dei protocolli etici del *syllabus*, inclusa la conformità al GDPR e all'*AI Act*[21]. La reazione osservata in aula rivela una tensione giuridicamente significativa: l'impiego di sistemi algoritmici è percepito come legittimo quando si controlla il sistema e se ne beneficia, e come inaccettabile quando ci si trova nella posizione di chi ne è oggetto. Questa asimmetria rivela che il consenso al trattamento automatizzato è frequentemente procedurale ma non sostanziale: il soggetto non possiede una comprensione adeguata delle logiche operative del sistema né delle possibili distorsioni. Il GDPR esige all'art. 13 un'informativa che renda comprensibili la logica e le conseguenze del trattamento automatizzato; l'*AI Act* impone ai sistemi ad alto rischio un livello di trasparenza che consenta agli utenti di interpretare l'*output* e utilizzarlo adeguatamente. Il caso illustra perché queste disposizioni non siano adempimenti burocratici ma condizioni di legittimità sostanziale: in assenza di una comprensione reale del funzionamento del sistema, la contestazione dell'esito valutativo, condizione posta come indefettibile dalla dottrina della legalità algoritmica, diviene di fatto impossibile per lo studente.

Che la trasparenza strutturale possa essere un valore progettuale e non solo un obbligo normativo lo illustra un workshop intensivo della durata di tre giorni, condotto nel 2026 presso l'Università di Oxford nell'ambito dell'Oxford Academy of English, dedicato all'insegnamento della scrittura accademica. Lo strumento didattico impiegato era progettato con vincoli generativi espliciti, dichiarati agli studenti come scelte pedagogiche deliberate: il sistema non può produrre testi completi né riscrivere elaborati altrui, ma guida il processo attraverso domande socratiche, segnala lacune argomentative e fornisce retroazione diagnostica senza offrire correzioni preformulate. L'impossibilità di generare contenuto al posto dello studente era presentata in aula come una precisa scelta educativa finalizzata a

[20] J. Kuzilek, M. Hlosta, Z. Zdrahal, *Open University Learning Analytics dataset*, *in Scientific Data*, 4, 2017, art. 170171; R.S. Baker, A. Hawn, op. cit.

[21] V.M. Lima, *The (Un)Ethical Double Standards of the "ChatGPT's Ember AI-Enhanced Student Assessment"*, in L.-D. Benyayer (a cura di), *Proceedings of the AI in Higher Education Summit 2026*, ESCP Business School / ESCPTech Institute, Parigi, 17-18 marzo 2026, 54-57.

preservare l'autorialità di chi apprende[22]. I risultati mostrano una trasformazione misurabile: l'ansia da valutazione si trasforma in consapevolezza autonoma dei propri punti di forza e di debolezza argomentativa, con uno studente che riferisce di aver progredito da una padronanza autopercepita del 10–15% a una stimata del 40–50%. Un sistema che per costruzione non può produrre il lavoro al posto dello studente soddisfa strutturalmente il requisito di supervisione umana effettiva dell'*AI Act*; e questa scelta di *design* è già, in sé, una decisione di *governance*: la *compliance* non viene inseguita a valle, ma incorporata a monte nella struttura dello strumento. Una risposta istituzionale al problema strutturale della valutazione nell'era dell'intelligenza artificiale – ossia la difficoltà strutturalmente irrisolvibile con misure puramente procedurali di distinguere l'elaborato autentico dall'*output* generato automaticamente – è offerta dalla riforma avviata dalla BI Norwegian Business School con il proprio «Exam Redesign Project», lanciato a partire dall'anno accademico 2025–2026 in corso di sperimentazione pilota in corsi selezionati. Il progetto articola la risposta in tre pilastri: un accertamento in ambiente controllato delle conoscenze concettuali di base (denominato *Assessment Center*), che il docente utilizza come termine di confronto per le prestazioni successive; una difesa orale digitale (condotta via Zoom in dieci minuti) convocata per un campione selezionato secondo un modello a rischio, comprendente una quota estratta casualmente (5-10% della coorte) e una individuata sulla base di discrepanze nell'elaborato consegnato; e una valutazione del processo formativo in tempo reale, che misura indicatori come la capacità dello studente di interagire criticamente con i materiali durante l'esecuzione dei compiti, anziché limitarsi a valutare il prodotto finale[23]. Dal punto di vista dell'*AI Act*, il terzo pilastro produce valutazioni con effetti sul risultato finale, ricadendo nell'Allegato III con i conseguenti obblighi di valutazione d'impatto e supervisione umana effettiva. L'elemento che distingue questa architettura da una delega acritica alla tecnologia è la distribuzione delle funzioni: l'automazione rileva, il giudizio umano decide. È questa ripartizione a soddisfare il requisito di effettività che l'*AI Act* pone per la supervisione dei sistemi ad alto rischio.

Il profilo dell'eguaglianza sostanziale trova invece concreta illustrazione in un progetto condotto tra il 2022 e il 2023 in un'area suburbana di Baku, in Azerbaigian, caratterizzata da tassi di diploma inferiori alla media nazionale e da una disponibilità limitata di insegnanti qualificati. Lo studio, condotto su 207 studenti di terza e quarta classe, ha analizzato l'impatto della piattaforma Khan Academy – un ambiente di apprendimento adattivo e gratuito che personalizza gli esercizi in base al livello dello studente e utilizza meccanismi di riconoscimento progressivo dei risultati raggiunti – sulla motivazione e sui risultati scolastici: il 68,6% degli studenti ha dichiarato un incremento dei propri voti, e l'analisi ha confermato una correlazione significativa tra l'ottenimento dei riconoscimenti progressivi e la persistenza nell'uso dello strumento[24]. Dal punto di vista del diritto all'istruzione, il caso illustra come le piattaforme digitali adattive possano svolgere una funzione di «democratizzazione cognitiva» in contesti a basse risorse, rendendo accessibili materiali di qualità laddove mancano le condizioni strutturali per garantirli attraverso i canali ordinari. L'accesso alla tecnologia diviene in questo contesto una condizione concreta di eguaglianza sostanziale nel senso

[22] V.M. Lima, *The (Un)Ethical Double Standards of the «ChatGPT's Ember AI-Enhanced Student Assessment»*, in L.-D. Benyayer (a cura di), *Proceedings of the AI in Higher Education Summit 2026*, Atti del convegno, Parigi, 17-18 marzo 2026, Parigi, 2026, 54-57.

[23] Ø.F. Landfald, *Navigating the 'Wicked Problem' of AI Assessment: A Case Study of an Exam Redesign*, in L.-D. Benyayer (a cura di) *Proceedings of the AI in Higher Education Summit 2026*, 17–18 marzo 2026, 41.

[24] A. Karimov, M. Saarela, T. Kärkkäinen, *The impact of online educational platform on students' motivation and grades: the case of Khan Academy in the under-resourced communities*, in M. Feng, T. Käser, P. Talukdar (a cura di), *Proceedings of the 16th International Conference on Educational Data Mining*, 2023.

dell'art. 3, c. 2 Cost.: uno strumento di rimozione degli ostacoli di fatto che impediscono la piena partecipazione al processo formativo. Occorre peraltro segnalare che il ricorso a tali piattaforme solleva, in parallelo, la questione del trattamento massivo di dati personali di studenti minorenni affidati a operatori privati e quella del divario digitale: senza un intervento pubblico che garantisca le infrastrutture di base (connettività e dispositivi), l'efficacia della tecnologia adattiva rischia di replicare piuttosto che ridurre le disuguaglianze di partenza.

Anche sul piano della spiegabilità e della tutela dell'autonomia decisionale dello studente, i sistemi di regolazione dinamica della difficoltà degli esercizi basati sulla Teoria della Risposta all'Item (IRT) offrono un esempio istruttivo. L'IRT è un modello matematico che, a differenza della valutazione tradizionale basata sul numero di risposte corrette, stima simultaneamente due fattori: la competenza dello studente e la difficoltà intrinseca di ciascuna domanda. Su questa base, la regolazione dinamica della difficoltà – un meccanismo analogo a quello impiegato in molti videogiochi moderni, in cui il livello di sfida si adatta in tempo reale alle prestazioni del giocatore – propone allo studente esercizi progressivamente più complessi quando va bene e più accessibili quando incontra difficoltà, con l'obiettivo di mantenerlo costantemente nella propria zona di sviluppo prossimale. Sperimentati nell'insegnamento della teoria dei grafi (una branca della matematica discreta che studia le relazioni tra oggetti attraverso reti di nodi e connessioni), questi sistemi mostrano miglioramenti documentati: 12,5 unità di difficoltà in media tra pre-test e post-test per il gruppo sottoposto a regolazione dinamica, rispetto a 9,6 per il percorso predefinito e 6,9 per gli esercizi in ordine casuale[25]. Poiché l'algoritmo decide autonomamente quale esercizio proporre, modificando le condizioni del percorso formativo, esso rientra nella nozione di decisione automatizzata rilevante ai fini dell'art. 22 GDPR: la scuola che lo adotta è tenuta a informare adeguatamente gli interessati e a garantire il diritto di intervento umano. Il sistema è tuttavia progettato per segnalare le situazioni di stallo come indicatore per l'intervento diretto del docente, realizzando quel modello di supervisione umana a valle dell'algoritmo che risponde al requisito di effettività dell'*AI Act*.

Infine, sul piano della spiegabilità, è istruttivo il caso di un sistema di raccomandazione dei corsi universitari basato su un algoritmo di tipo *k-nearest neighbor* – letteralmente «i *k* vicini più prossimi». Il principio di funzionamento è intuitivo: il sistema individua, tra tutti gli studenti già laureati, quelli con il profilo accademico più simile a quello dello studente in esame (i «vicini» appunto), e suggerisce i percorsi che quella popolazione ha seguito con successo. Sperimentato in tre corsi di laurea per ridurre il rischio di abbandono, il sistema indicava sistematicamente agli studenti in difficoltà di concentrarsi su un numero di esami inferiore rispetto a quello in cui tendevano a immatricolarsi[26]. La raccomandazione può essere motivata mostrando il percorso reale di studenti con profilo comparabile, soddisfacendo in modo strutturale, e non soltanto formale, il requisito di spiegabilità praticabile. Inoltre, invece di limitarsi a classificare lo studente come «a rischio», con il pericolo di produrre un effetto di profezia auto-avverante, il sistema propone alternative positive, potenziando la capacità decisionale dello studente anziché comprimerla.

I casi esaminati mostrano che la variabile determinante non è la sofisticazione tecnologica del sistema adottato, ma la qualità della *governance* che lo sorregge: la trasparenza strutturale, la spiegabilità praticabile, la supervisione umana effettiva e la consapevolezza che la decisione educativa resta sempre in capo a un soggetto umano identificabile che ne risponde. È in

[25] A. Schütt, T. Huber, I. Aslan, E. *André, Fast dynamic difficulty adjustment for intelligent tutoring systems with small datasets*, in M. Feng, T. Käser, P. Talukdar (a cura di), *Proceedings of the 16th International Conference on Educational Data Mining*, 2023.

[26] K. Wagner, A. Merceron, P. Sauer, N. Pinkwart, *A course recommender system built on success to support students at risk in higher education*, in *Journal of Educational Data Mining*, 16(1), 2024, 330-364.

Issn 2421-0528

questa consapevolezza, e non nel rifiuto o nell'adozione acritica della tecnologia, che risiede la condizione di legittimità dell'intelligenza artificiale nella scuola.

Abstract. L'integrazione dell'intelligenza artificiale nei contesti scolastici e della formazione superiore pone questioni giuridiche che non si esauriscono nel piano della protezione dei dati, ma investono l'effettività del diritto all'istruzione nella sua accezione costituzionale moderna. Il contributo muove dalla lettura del diritto all'istruzione come diritto fondamentale a prestazione (alla luce delle sent. nn. 215/1987 e 275/2016 della Corte cost.), che costituisce la cornice entro cui valutare qualunque infrastruttura incidente sulle condizioni reali di partecipazione al processo formativo. Su tale fondamento si innesta il quadro normativo applicabile, presentato in ordine cronologico: il reg. UE 2016/679 (GDPR), che impone alla scuola obblighi di titolarità del trattamento, valutazione d'impatto e tutela rinforzata dei dati dei minori; la dottrina della legalità algoritmica elaborata dal Consiglio di Stato (sentt. nn. 2270/2019, 8472–8474/2019, 881/2020), da cui discendono i tre criteri operativi di spiegabilità praticabile, contestabilità e responsabilità; il reg. UE 2024/1689 (*AI Act*), che classifica come ad alto rischio i sistemi impiegati per valutare le prestazioni degli studenti o orientarne il percorso. Il secondo paragrafo distingue i sistemi di intelligenza artificiale generativa dagli strumenti digitali ordinari e dai sistemi predittivi, analizzandone i profili di rischio (inaffidabilità strutturale della fonte, distorsione sistematica, opacità del processo generativo, trattamento dei dati prodotti nell'interazione) e le opportunità in termini di personalizzazione, inclusione e sviluppo delle competenze critiche, con attenzione al nesso con l'art. 3, c. 2 Cost. Il terzo paragrafo verifica empiricamente questi profili attraverso *case study* comparati tratti da contesti internazionali, inclusi paesi extra-UE: la rassegna sistematica di Baker e Hawn e l'analisi del *dataset* Open University (OULAD, 2017) documentano i rischi di discriminazione algoritmica nei sistemi predittivi; l'esperienza ESCP Ember (Parigi, marzo 2026) illustra il problema del consenso sostanziale al trattamento algoritmico; il *workshop* di scrittura accademica di Oxford (2026) dimostra come la trasparenza strutturale possa essere incorporata nel design già a livello progettuale; l'iniziativa della BI Norwegian Business School propone un'architettura valutativa a tre pilastri come risposta istituzionale al problema dell'autenticità degli elaborati; i sistemi IRT, la piattaforma Khan Academy in Azerbaigian e l'algoritmo KNN illustrano le condizioni alle quali l'automazione produce effetti positivi e compatibili con il principio di eguaglianza sostanziale. Il contributo conclude sostenendo che la legittimità dell'IA nella scuola dipende dalla qualità della *governance* che la sorregge: solo una *governance* verificabile, fondata sulla spiegabilità praticabile e sulla supervisione umana effettiva, consente all'intelligenza artificiale di servire l'educazione democratica anziché eroderne le garanzie.

Abstract. The integration of artificial intelligence in school and higher education settings raises legal questions that go beyond data protection and concern the effectiveness of the right to education in its modern constitutional meaning. The paper starts from the understanding of the right to education as a fundamental right to positive provision (as clarified by the Italian Constitutional Court in decisions n. 215/1987 and n. 275/2016), which constitutes the constitutional frame against which any infrastructure affecting the real conditions of participation in education must be assessed. On this foundation, the paper reconstructs the applicable regulatory framework in chronological order: the GDPR (reg. UE 2016/679), imposing on schools obligations as data controllers, impact assessment requirements, and enhanced protection for minors' data; the doctrine of algorithmic legality developed by the Italian Council of State (decisions nn. 2270/2019, 8472– 8474/2019, 881/2020), from which three operational criteria derive: practicable explainability,

Issn 2421-0528

contestability, and accountability; and the AI Act (reg. UE 2024/1689), classifying as high-risk those systems used to assess student performance or guide learning pathways. The second section distinguishes generative AI systems from ordinary digital tools and predictive algorithms, analysing their risk profiles (structural unreliability of output, systematic bias, opacity of the generative process, and processing of interaction data) and their opportunities in terms of personalisation, inclusion, and development of critical competences, with attention to the constitutional link with substantive equality. The third section tests these profiles through comparative case studies from international contexts, including non-EU countries: the systematic review by Baker and Hawn and the Open University Learning Analytics Dataset analysis (OULAD, 2017) document the risks of algorithmic discrimination in predictive systems; the ESCP Ember experience (Paris, March 2026) illustrates the problem of substantive consent to algorithmic assessment; the Oxford academic writing workshop (2026) demonstrates how structural transparency can be embedded in tool design at the design stage itself; the BI Norwegian Business School initiative proposes a three-pillar assessment architecture as an institutional response to the authenticity problem; and the IRT systems, Khan Academy in Azerbaijan, and the KNN algorithm illustrate the conditions under which automation produces outcomes compatible with the principle of substantive equality. The paper concludes that the legitimacy of AI in education depends on the quality of the governance surrounding it: only verifiable governance, grounded in practicable explainability and effective human oversight, can make artificial intelligence serve democratic education rather than undermine its guarantees.

Parole chiave. Intelligenza artificiale generativa – *AI Act* – GDPR – Legalità algoritmica – *governance* scolastica – Diritto all'istruzione – Diritti fondamentali.

Key words. Generative artificial intelligence – *AI Act* – GDPR – Algorithmic legality – school governance – Right to education – Fundamental rights.