

Didattica e intelligenza artificiale: risvolti etici, problemi di privacy e sorveglianza, manipolazione dei dati^a

Patrizia Natale*

Abstract

L'introduzione dell'intelligenza artificiale (IA) nella didattica sta rivoluzionando l'educazione, offrendo personalizzazione e ottimizzazione gestionale, ma sollevando importanti questioni etiche. Tra queste emergono i rischi legati a privacy, sorveglianza, manipolazione e "bias" algoritmici, che possono generare discriminazioni razziali, di genere, economiche e linguistiche, ampliando le disuguaglianze preesistenti. Inoltre, la gestione dei diversi dati degli studenti, spesso minorenni, raccolti dai programmi di IA richiede trasparenza e sicurezza, nel rispetto di normative come il GDPR e l'AI Act. Per evitare che l'IA rafforzi divari sociali e stereotipi, è fondamentale una progettazione etica, il monitoraggio continuo degli algoritmi e la collaborazione tra scuole, aziende e autorità. Serve integrare nei curricula l'educazione digitale e la consapevolezza critica sull'uso della tecnologia, ricordando che l'IA deve restare uno strumento al servizio dell'apprendimento e non un fine. Solo con una governance solida e condivisa si potrà garantire un'istruzione equa, inclusiva e rispettosa dei diritti di tutti.

Parole Chiave: didattica ed intelligenza artificiale; bias; disuguaglianza algoritmica; protezione dei diritti individuali; intelligenza artificiale; controllo degli algoritmi.

This paper explores the limitations of algorithmic fairness, particularly the “impossibility theorem of fairness”, and discusses how a structural understanding of justice can address the related ethical concerns. After presenting the main models of algorithmic fairness, I argue that they overlook key justice concerns by prioritizing outcome-based metrics and isolating decision-making from broader socio-historical contexts. Furthermore, when base rates differ, it becomes impossible to satisfy more than one fairness metric simultaneously. To address these shortcomings, I propose integrating algorithmic fairness with Iris M. Young’s notion of structural injustice, which accounts for entrenched inequalities rooted in the interplay of behaviors, norms, and institutions. This approach situates algorithms within their broader socio-historical context, emphasizing systemic factors that influence decision-making and perpetuate unjust outcomes. I further contend that a structural perspective assigns algorithms a twofold role, particularly in contentious cases where ethical controversies are at play. First, a diagnostic function: by exposing underlying ethical imbalances and biases, algorithms can highlight critical areas for systemic reforms. Second,

^a Saggio ricevuto in data 10/03/2025 e pubblicato in data 09/12/2025.

* Docente di informatica presso istituto di istruzione superiore, e-mail: pattnatale@gmail.com

they can serve as evaluative tools, enabling the assessment and prioritization of fairness metrics on a case-by-case basis.

Keywords: education and artificial intelligence, bias, algorithmic inequality, protection of individual rights; artificial intelligence, algorithmic control.

1. Introduzione

Le metodologie didattiche stanno subendo profonde trasformazioni con l'introduzione dell'intelligenza artificiale (AI), e continuamente vengono evidenziati i benefici derivanti dall'utilizzo dell'AI per giungere ad una vera individualizzazione dell'educazione rispettando ed esaltando le competenze e conoscenze pregresse, il background culturale le esperienze vissute dai discenti. L'uso dell'AI, infatti, permette la personalizzazione dell'apprendimento e il miglioramento dei processi di gestione del gruppo classe elaborando percorsi personalizzati per ciascun alunno; devono, tuttavia, essere analizzati a fondo i problemi etici legati all'uso dell'AI che fanno riferimento alla sorveglianza, alla privacy e alla possibilità di poter manipolare gli studenti. In riferimento a ciò, come analizza Neil Selwyn nel volume *Should robots replace teachers?*¹, è fondamentale che l'introduzione dell'AI non diventi un processo passivo e tecnocratico, ma una scelta didattica consapevole e sostenuta da riflessioni pedagogiche e critiche. In questo ambito i temi che afferiscono alla protezione dei diritti individuali e alla responsabilità sociale nell'uso delle tecnologie avanzate diventano focali. A livello nazionale troviamo pubblicazioni recenti quali *L'intelligenza artificiale a scuola. Guida per una pratica didattica consapevole*² di Giulia Lorenzoni e *La classe potenziata*³ di Lorenzo Redaelli che propongono un approccio didattico in cui l'AI sia al servizio dell'insegnamento e non viceversa; questi contributi mostrano che l'AI non va introdotta come elemento neutrale ma va collocata in un contesto pedagogico e normativo che ne sostenga l'uso etico, responsabile ed equo.

2. IA e implicazioni etiche: esempi

L'IA presenta strumenti promettenti per personalizzare l'apprendimento e migliorare l'efficienza educativa, ma solleva anche preoccupazioni etiche, in particolare riguardo all'equità degli algoritmi che analizzano i dati degli studenti. Il rischio di discriminazione, causato da "bias" nei modelli algoritmici, è uno degli aspetti più critici di questa tecnologia. È fondamentale riflettere sulle implicazioni etiche dell'uso dell'IA nella didattica e specialmente su come gli algoritmi possano amplificare o creare nuove disuguaglianze tra gli studenti. I sistemi di IA utilizzati nelle scuole raccolgono una grande quantità di informazioni sugli alunni, comprese le loro conoscenze, lacune, progressi e persino abitudini di studio e dettagli biometrici, a seconda del programma di IA impiegato. Holmes⁴

¹ N. Selwyn, *Should robots replace teachers?*, Polity Press, Cambridge 2019.

² G. Lorenzoni, *L'intelligenza artificiale a scuola. Guida per una pratica didattica consapevole*, Lattes Editori, Torino 2024.

³ L. Redaelli, *La classe potenziata*, Mondadori Università, Milano 2025.

⁴ W. Holmes, M. Bialik, C. Fadel, *Artificial Intelligence in Education: Promises and implications for teaching and learning*, Center for Curriculum Redesign, Boston, 2019.

e Xu & Ouyang⁵ evidenziano come la raccolta di dati, se non supportata da trasparenza e accountability, possa diventare uno strumento sotto il controllo di pochi attori e recare danno all'autonomia educativa. Lorenzoni, in Italia, suggerisce che la progettazione educativa includa attività di alfabetizzazione digitale preventiva, volta a far sì che docenti e studenti comprendano i protocolli di raccolta dei dati e i rischi ad essi associati. È necessaria una particolare attenzione e un elevato grado di cautela nella gestione e nel trattamento dei dati, prevedendo l'adozione di adeguate misure tecniche e organizzative volte a garantire la riservatezza, l'integrità e la disponibilità delle informazioni, nonché a minimizzare i rischi connessi a possibili accessi non autorizzati, alterazioni o perdite dei dati trattati. Gestire adeguatamente la privacy degli studenti è complicato poiché i loro dati sono spesso amministrati da terzi che controllano i database e le piattaforme di IA. Chi ha accesso a questi contenuti? Chi garantisce dai rischi di utilizzo improprio e/o non autorizzato delle informazioni? Le scuole devono bilanciare l'innovazione didattica con il diritto alla riservatezza degli studenti. La situazione è complicata dalla mancanza di normative uniformi tra i vari paesi e dalla scarsa trasparenza delle aziende tecnologiche riguardo all'uso dei dati; non è sempre chiaro quali dati vengano raccolti, il loro scopo e la durata della conservazione, rendendo difficile per le scuole effettuare verifiche accurate. Sarebbe opportuno proporre modelli di “data stewardship” partecipata, in cui le scuole e le famiglie possono controllare in modo trasparente i dati generati dagli studenti. Chi lavora nel settore deve essere consapevole che gli algoritmi, pur essendo progettati per personalizzare l'insegnamento, possono anche influenzare pensieri e comportamenti dei discenti. Gli esperti del settore si interrogano sulla neutralità delle piattaforme educative e su come l'IA possa influenzare le scelte degli studenti in modo inconsapevole. È essenziale affrontare questi interrogativi per prevenire manipolazioni, anche non intenzionali; le ricerche di Boulamwini e Floridi mostrano che il modo in cui gli studenti vedono l'intelligenza artificiale può influenzare quanto si aspettano da sé stessi e quanto si sentono motivati a imparare. Se la percepiscono come superiore o “perfetta”, possono sentirsi scoraggiati o meno coinvolti.

Se gli algoritmi, a causa di pregiudizi, favoriscono determinate aree di studio o carriere, gli studenti potrebbero prendere decisioni future che non riflettono le loro vere aspirazioni, limitando le loro opportunità di crescita personale e professionale. Non esiste un algoritmo equo e neutrale; studi dimostrano che possono contenere pregiudizi, sia consapevoli che inconsapevoli, riflettendo i “bias” presenti nei dati di addestramento, il che può portare a trattamenti ingiusti, specialmente per studenti di minoranze etniche o gruppi svantaggiati. Se un algoritmo non considera adeguatamente le diversità culturali, socio-economiche e di genere, potrebbe ostacolare invece di supportare gli studenti nel superare le difficoltà. I sistemi educativi stanno iniziando ad utilizzare l'IA per raccomandare corsi, monitorare il rendimento degli studenti e/o valutarne le prestazioni. In una situazione come questa, il pregiudizio razziale può influenzare l'accesso alle opportunità educative e modificare la qualità degli strumenti proposti; ad esempio, gli studenti di colore potrebbero essere trattati in modo diverso rispetto ai loro coetanei bianchi a causa di pregiudizi storici che si riflettono nei dati, non giudicando esclusivamente le loro capacità o il loro impegno. Altro problema significativo è la discriminazione di genere: gli algoritmi possono rafforzare stereotipi, orientando gli studenti verso specifici percorsi accademici in base al sesso,

⁵ W. Xu, F. Ouyang, *The application of AI technologies in STEM education: A systematic review from 2011 to 2021*, «International Journal of STEM Education», 9, 2022.

piuttosto che alle loro reali attitudini. Ad esempio, un sistema che consiglia corsi di matematica e scienze agli studenti maschi e corsi di arte alle studentesse potrebbe limitare le opportunità per le ragazze di sviluppare competenze scientifiche, mentre i ragazzi potrebbero essere esclusi da aree che richiedono sensibilità umanistica. Tale “bias”, se non gestito adeguatamente, può perpetuare ruoli di genere tradizionali, limitando la libertà di scelta e l’equità nell’istruzione. Per questo motivo, è fondamentale che l’introduzione dell’IA nelle scuole non avvenga in modo isolato o puramente tecnico. Lorenzoni e Rivoltella, ad esempio, sostengono che l’IA vada inserita in percorsi didattici che non limitino l’apprendimento a un rapporto studente-macchina, ma che includano momenti di riflessione collettiva, valutazione critica, confronto tra pari e discussione etica e sociale, stimolando un apprendimento consapevole e partecipato.

Un ulteriore elemento critico riguarda il divario digitale, che rischia di amplificare le disuguaglianze nell’ambito dell’istruzione supportata dall’IA. Gli studenti provenienti da famiglie a basso reddito, infatti, potrebbero non disporre di dispositivi adeguati o di connessioni Internet sufficientemente stabili e veloci, limitando così la loro possibilità di accedere e utilizzare le risorse educative basate sull’intelligenza artificiale. Questa situazione potrebbe tradursi in forme di discriminazione indiretta, penalizzando tali studenti nella valutazione delle loro performance non a causa di competenze o risultati inferiori, ma per la carenza di reali opportunità di partecipazione alle attività online. Quando gli algoritmi analizzano i dati degli studenti, spesso non considerano le differenze individuali o le esperienze di vita ma solo i dati pregressi a cui hanno avuto accesso.

È dimostrato da numerosi studi che l’uso distorto dei dati e i “bias” nei sistemi di intelligenza artificiale possono avere effetti negativi in vari contesti, riproducendo e amplificando le disuguaglianze esistenti. Prendendo in considerazione, ad esempio, il modo in cui lavora l’algoritmo COMPAS, utilizzato nel sistema giuridico statunitense per prevedere il rischio di recidiva, si evidenzia come i pregiudizi tecnologici possano avere impatti concreti; COMPAS⁶, infatti, ha mostrato risultati problematici, in particolare per la discriminazione razziale, penalizzando in modo sproporzionato le persone di colore. Un’analisi ha rivelato che il sistema classificava erroneamente gli imputati neri come ad alto rischio di recidiva, mentre sottovalutava il rischio per gli imputati bianchi. Questo esempio dimostra come i sistemi automatizzati possano rinforzare disuguaglianze preesistenti, riflettendo i “bias” storici dei dati utilizzati. Il caso di Amazon⁷, invece, riguarda l’uso dell’IA nel reclutamento. Nel 2017-2018, l’azienda ha abbandonato un algoritmo progettato per selezionare candidati a causa di un “bias” di genere. Addestrato su curriculum di candidati passati, il sistema rifletteva la predominanza maschile nelle posizioni tecniche, favorendo i candidati maschi e scartando le donne, anche con qualifiche superiori. Questo episodio solleva interrogativi su come gli algoritmi possano riprodurre e amplificare disuguaglianze pregresse. In un contesto educativo, un algoritmo che valuta le performance degli studenti potrebbe facilmente commettere errori simili, discriminando le studentesse in settori dominati da uomini o viceversa. Sempre parlando di disuguaglianze di genere è

⁶ J. Angwin, J. Larson, S. Mattu, L. Kirchner, *Machine Bias: There’s software used across the country to predict future criminals. And it’s biased against blacks*, «ProPublica», 23 May 2016, <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

⁷ Reuters, *Amazon ditched AI recruiting tool that favored men for technical jobs*, «The Guardian», 11 October 2018, <https://www.theguardian.com/technology/2018/oct/10/amazon-hiring-ai-gender-bias-recruiting-engine>.

facile constatare la presenza di rappresentazioni stereotipate (e discriminatorie) con una semplice navigazione su Internet utilizzando motori di ricerca. Ad esempio, una ricerca di immagini associate alla parola “donna” mostra principalmente donne giovani e attraenti, mentre una ricerca di contenuti per la stessa parola chiave porta a risultati legati a cosmetici, moda o abbigliamento, se invece cerchiamo “uomo” il numero di contenuti relativi alla moda e le immagini di uomini giovani e attraenti è inferiore.

Studi sulla discriminazione derivante dall'uso dell'IA sono stati effettuati principalmente negli Stati Uniti poiché da più tempo utilizzano questi tipi di sistemi. Tra i vari esempi rilevanti, basato su studi americani, troviamo il sistema di concessione di prestiti adottato da diverse banche e finanziarie. Certi algoritmi di valutazione del credito penalizzavano minoranze etniche e residenti in aree svantaggiate⁸. Questi modelli, progettati per stimare la capacità di rimborso dei clienti, presentavano evidenti distorsioni razziali, poiché si basavano su dati storici influenzati da pratiche discriminatorie del settore bancario. Le informazioni utilizzate mostravano disuguaglianze nell'accesso ai finanziamenti tra gruppi etnici, con approvazioni inferiori per persone di colore, anche a parità di punteggio creditizio rispetto ai richiedenti bianchi. Tale esempio dimostra chiaramente come gli algoritmi predittivi possano aggravare disparità preesistenti e consolidare stereotipi sociali ed economici.

Anche in ambito sanitario possiamo dedurre che l'uso di algoritmi possa comportare il rischio di “bias”, se, come è successo in Amazon, tali sistemi vengono addestrati su dati storici che riflettono squilibri preesistenti. Poiché nella medicina e nella sanità i ruoli di responsabilità sono stati storicamente occupati in maggioranza da uomini, un algoritmo di selezione o pianificazione potrebbe replicare e rafforzare questa disparità, penalizzando le donne e i gruppi sottorappresentati. Oltre al versante lavorativo, la letteratura mostra che “bias” emergono anche nell'erogazione delle cure: alcuni algoritmi clinici hanno infatti sottostimato i bisogni dei pazienti neri⁹, riducendo l'accesso a programmi di assistenza. Ne deriva che l'introduzione non controllata dell'IA in sanità rischia di accentuare le disuguaglianze invece di colmarle.

3. Utilizzo dell'IA in contesti educativi

Nel contesto scolastico, diversi studi internazionali hanno evidenziato criticità legate agli algoritmi di correzione automatica, specialmente quelli basati su intelligenza artificiale per valutare test scritti. Un esempio riguarda gli esami di ammissione universitaria negli Stati Uniti, dove sistemi di valutazione automatizzata tendevano a penalizzare studenti con background linguistici diversi o provenienti da scuole non anglofone. Gli errori rilevati dagli algoritmi non derivavano da carenze nei contenuti, ma piuttosto da differenze stilistiche o linguistiche legate alla cultura di origine, gli studenti non madrelingua anglofoni vedevano ridursi le opportunità di accesso poiché non si conformavano agli standard linguistici con cui gli algoritmi erano stati addestrati.

⁸ R. Browne, M. Sigalos, *A.I. has a discrimination problem. In banking, the consequences can be severe*, «CNBC», 23 June 2023, <https://www.cnbc.com/2023/06/23/ai-has-a-discrimination-problem-in-banking-that-can-be-devastating.html>.

⁹ K. Manke, *Widely used health care prediction algorithm biased against black people*, «UC Berkeley News», 23 June 2023, <https://news.berkeley.edu/2019/10/24/widely-used-health-care-prediction-algorithm-biased-against-black-people/>.

Vanno segnalate, inoltre, criticità nell'uso scolastico di sistemi di IA per monitorare il comportamento degli alunni attraverso dati biometrici o interazioni digitali. Diverse analisi hanno messo in luce come questi strumenti possano sfavorire studenti con disturbi dell'attenzione o modalità di apprendimento atipiche. Ad esempio, software per il controllo dell'attenzione rischiano di classificare in modo scorretto studenti con ADHD come poco motivati o disinteressati, semplicemente perché utilizzano strategie diverse di concentrazione. Questo tipo di “bias” rischia di generare valutazioni distorte delle capacità e risposte educative non adeguate alle reali esigenze degli studenti.

Le scuole italiane che hanno iniziato ad utilizzare sistemi basati su algoritmi per personalizzare i percorsi formativi, monitorare i risultati e proporre contenuti su misura in base alle abitudini di studio degli alunni, hanno effettuato un'attenta progettazione senza la quale queste tecnologie rischiano di amplificare le disuguaglianze?

Alcune ricerche recenti¹⁰ hanno messo in evidenza come i sistemi di apprendimento adattivo e le piattaforme scolastiche basate su intelligenza artificiale possano rafforzare le disuguaglianze già presenti. Se gli algoritmi non sono progettati con attenzione, rischiano di favorire studenti già avvantaggiati o con maggiori risorse, mentre chi parte da situazioni di svantaggio riceve un supporto meno mirato. In questo modo, l'uso dell'IA in ambito educativo, invece di ridurre i divari, può contribuire ad ampliarli, con effetti rilevanti sia nella scuola che nell'università. Algoritmi che non considerino la varietà di bisogni e contesti degli studenti, rischiano di escludere proprio coloro che richiedono più attenzione. Anche nelle università sono stati adottati strumenti predittivi di IA per individuare chi potrebbe abbandonare gli studi e avere quindi bisogno di un supporto; tali strumenti sono basati su dati storici come voti, presenze e attività extra. Queste analisi, spesso, riflettono pregiudizi legati all'origine sociale ed etnica, segnalando come “a rischio”, con maggiore frequenza, gli studenti provenienti da famiglie meno abbienti o da minoranze, replicando disuguaglianze già radicate nel sistema educativo. Tra i principali rischi c'è il “bias” di genere, quando gli algoritmi riflettono stereotipi educativi già presenti nei dati storici, e quello razziale, come dimostrato dagli errori dei sistemi di riconoscimento facciale, meno precisi con i volti delle persone di colore. Se applicate anche a tecnologie educative, come il riconoscimento delle emozioni o il monitoraggio degli studenti, queste distorsioni potrebbero causare valutazioni scorrette, penalizzando alcuni gruppi e compromettendo l'equità dei processi educativi, influenzando negativamente l'accesso alle opportunità di apprendimento.

Un ulteriore esempio riguarda l'uso dell'intelligenza artificiale per valutare e finanziare le scuole. Alcuni algoritmi hanno favorito istituti in zone benestanti, penalizzando quelli situati in aree svantaggiate; basi di dati che ignorano fattori esterni, come il livello di risorse disponibili o il supporto sociale, hanno contribuito ad accentuare il divario tra scuole ricche e povere, riducendo ulteriormente il sostegno proprio dove sarebbe più necessario. Ad esempio lo studio *Algorithmic Bias in Education*¹¹ evidenzia come i sistemi educativi basati sull'IA possano perpetuare disuguaglianze, specialmente in

¹⁰ W. Strielkowski, T. Veinbender, L. Volkova, O. Garanina, *AI-driven adaptive learning for sustainable educational development*, «Sustainable Development», vol. 33, n. 2, 2025. <https://doi.org/10.1002/sd.3221>; I. Molenaar, R. F. Kizilcec, B. Chen, *Unveiling the shadows: Beyond the hype of AI in education*, «Heliyon», vol. 10, n. 5, 2024, e30696.

¹¹ R.S. Baker, A. Hawn, *Algorithmic Bias in Education*, «OSF», preprint 2021, <https://osf.io/preprints/edrxiv/pbmvgz>.

contesti rurali o con risorse limitate. Le scuole in aree svantaggiate potrebbero non avere infrastrutture adeguate per implementare efficacemente l'IA, portando a risultati meno accurati e a decisioni che non rispondono alle reali esigenze degli studenti.

4. *Necessità del controllo degli algoritmi*

I casi esposti dimostrano come i pregiudizi algoritmici incidano anche sull'istruzione, minacciando l'uguaglianza di opportunità. Per evitare che l'IA aggravi le disuguaglianze, è indispensabile progettare sistemi trasparenti, equi e inclusivi, vigilando attentamente sui dati utilizzati e correggendo eventuali distorsioni. Per questo l'introduzione dell'IA nelle scuole deve essere accompagnata da riflessioni etiche serie, evitando che strumenti pensati per supportare l'apprendimento diventino causa di esclusione. È essenziale che ogni algoritmo sia costantemente controllato e aggiornato, garantendo parità di trattamento per tutti, senza distinzione di origine sociale, etnica o di genere. È fondamentale che insegnanti e studenti sviluppino competenze critiche nei confronti delle tecnologie basate sull'intelligenza artificiale, per poter riconoscere e segnalare tempestivamente eventuali ingiustizie o distorsioni algoritmiche. Come osservano Williamson e Piattoeva¹², l'uso crescente di dati e algoritmi nelle pratiche scolastiche rischia di produrre decisioni opache, non sempre comprensibili da chi le subisce. In quest'ottica, la formazione digitale non deve limitarsi all'aspetto tecnico-operativo, ma includere anche la capacità di comprendere le logiche sottese ai sistemi di IA, per individuarne i limiti e le implicazioni etiche e sociali.

Servono regole chiare e rigorose, trasparenza sull'uso dei dati e un'informazione completa rivolta a famiglie e studenti sugli scopi e sui meccanismi degli algoritmi. Secondo Veale e Binns¹³, i sistemi automatizzati, soprattutto se impiegati in ambiti delicati come l'educazione, devono essere progettati per garantire accountability e possibilità di contestazione. Questo significa che le decisioni algoritmiche devono essere comprensibili, verificabili e, se necessario, modificabili.

A ciò si aggiunge la necessità di garantire precise tutele sulla privacy e sul diritto alla gestione dei propri dati personali. Il *White Paper on Artificial Intelligence*¹⁴ pubblicato dalla Commissione Europea sottolinea l'importanza di garantire agli utenti – soprattutto se minori – la possibilità di conoscere come vengono utilizzati i propri dati, oltre al diritto di rettificarli o cancellarli. In linea con questa posizione, Livingstone, Stoilova e Nandagiri¹⁵ evidenziano come i diritti digitali dei bambini siano spesso ignorati anche nei contesti scolastici, e che sia fondamentale promuovere una cultura della responsabilità anche nella progettazione delle piattaforme educative.

¹² B. Williamson, N. Piattoeva, *Objectivity as standardization in data-scientific education policy, technology and governance. Learning*, «Media and Technology», vol. 43, n. 3, pp. 1-13, 2018, <https://doi.org/10.1080/17439884.2018.1556215>.

¹³ M. Veale, M. Van Kleek, R. Binns, *Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making*, «Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems ACM», 2018, <https://dl.acm.org/doi/pdf/10.1145/3173574.3174014>.

¹⁴ European Commission, *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust*, Brussels 2020, <https://ec.europa.eu>.

¹⁵ M. Stoilova, S. Livingstone, R. Nandagiri, *Digital by default: Children's capacity to understand and manage online data and privacy*, «Media and Communication», vol. 8, n. 4, 2020, pp. 197-207, <https://doi.org/10.17645/mac.v8i4.3407>.

Solo combinando competenze critiche, regole trasparenti e garanzie solide sulla privacy sarà possibile costruire un ambiente educativo in cui l'IA sia uno strumento di supporto equo e responsabile, capace di valorizzare le potenzialità di tutti senza riprodurre o amplificare disuguaglianze preesistenti.

I sistemi di IA non raccolgono solo dati anagrafici, ma anche aspetti sensibili come comportamenti, preferenze di studio e dati biometrici. Questa mole di informazioni, spesso elaborata senza un controllo trasparente da parte degli utenti finali, pone seri interrogativi etici, soprattutto nel contesto scolastico dove gli studenti sono soggetti vulnerabili, soprattutto a causa della giovane età. I dati generati quotidianamente nelle interazioni educative possono essere utilizzati non solo per personalizzare l'apprendimento, ma anche per sorvegliare, valutare e classificare gli alunni in base a parametri non sempre chiari né scientificamente fondati¹⁶.

In particolare, le tecnologie di tracciamento dell'attenzione, il riconoscimento facciale e le piattaforme adattive rischiano di esporre gli studenti a una continua analisi comportamentale, potenzialmente lesiva della loro autonomia e del diritto alla riservatezza¹⁷. Inoltre, il trattamento automatizzato di tali dati può consolidare stereotipi e discriminazioni preesistenti se non viene costantemente monitorato da un punto di vista etico e sociale¹⁸.

Solo un approccio condiviso e responsabile tra scuole, istituzioni e aziende può assicurare un equilibrio tra innovazione tecnologica e tutela dei diritti degli studenti. Come sottolinea il *White Paper on Artificial Intelligence* della Commissione Europea¹⁹, le applicazioni di IA nel contesto scolastico devono rispettare principi fondamentali come la trasparenza, la non discriminazione, la protezione dei dati e la supervisione umana. È essenziale che i processi decisionali automatizzati siano accompagnati da meccanismi di controllo accessibili e da una governance partecipata, in cui le famiglie, i docenti e gli studenti stessi possano esercitare un ruolo attivo.

5. Normativa attuale e supervisione degli strumenti di IA

Per garantire un uso realmente etico dell'intelligenza artificiale nella scuola, è fondamentale considerare leggi attuali, innovazioni digitali e responsabilità educative. Il Regolamento Generale sulla Protezione dei Dati (GDPR) stabilisce che scuole e aziende devono raccogliere solo le informazioni indispensabili per scopi didattici, evitando dati superflui, anche se questi potrebbero migliorare la personalizzazione dei percorsi formativi. Le scuole, quindi, devono informare chiaramente studenti, famiglie e insegnanti sull'uso dell'IA, specificando finalità, dati coinvolti e funzionamento degli algoritmi. Allo stesso modo, le aziende produttrici di software educativi devono garantire la conformità alle

¹⁶ B. Williamson, N. Piattoeva, *Objectivity as standardization in data-scientific education policy, technology and governance in Learning*, «Media and Technology», vol. 44, n. 1, 2019, pp. 64-76, <https://doi.org/10.1080/17439884.2018.1556215>.

¹⁷ UNICEF, *Policy guidance on AI for children*, UNICEF Office of Research – Innocenti, 2021, <https://www.unicef.org/innocenti/media/1341/file/UNICEF-Global-Insight-policy-guidance-AI-children-2.0-2021.pdf>.

¹⁸ M. Veale, M. Van Kleek, R. Binns, *Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making*, «CHI 2018 Papers», <https://dl.acm.org/doi/pdf/10.1145/3173574.3174014>.

¹⁹ European Commission, *White Paper on Artificial Intelligence*, cit.

norme e ai principi etici, così da consentire agli utenti di esercitare i propri diritti in modo consapevole. La protezione dei dati richiede soluzioni concrete, come crittografia avanzata, autenticazione sicura e aggiornamenti costanti, oltre alla formazione del personale scolastico nella gestione di informazioni sensibili. Solo così si possono ridurre errori e rischi, anche considerando le minacce informatiche. È essenziale che docenti e operatori scolastici comprendano l'importanza della sicurezza digitale e agiscano con attenzione.

Anche la fase di progettazione degli algoritmi è un tema cruciale: trascurare i valori etici potrebbe portare alla creazione di strumenti inadatti e dannosi per la scuola, capaci di generare discriminazioni o percorsi scolastici non adeguati. Gli sviluppatori devono quindi seguire criteri di responsabilità e imparzialità, ma resta aperta la questione su chi debba controllare e garantire tale imparzialità. Devono farlo le singole scuole quando decidono di adottare le piattaforme dell'IA, avvalendosi di esperti del settore, anche esterni, o l'amministrazione centrale validando, attraverso apposite commissioni di comprovati esperti del settore, i diversi strumenti di IA da utilizzare in ambito educativo lasciando agli istituti solo il compito di scegliere quale piattaforma “approvata” utilizzare? Se si volesse interpretare nel modo più ampio il concetto di “libertà di insegnamento” dovrebbero farlo le singole scuole, ma queste, anche consorziate, hanno a disposizione le risorse necessarie per far ricorso a professionisti veramente esperti del settore? Una risposta, per ora, non c'è, ma è chiaro che, in questo settore così delicato, non si può risolvere il problema affidando a personale interno all'istituzione scolastica, magari formato con un corso on line, il compito di effettuare una valutazione minuziosa di tutti gli aspetti degli algoritmi utilizzati e che impatteranno sulla qualità dell'educazione impartita alle future generazioni.

Diversi studiosi hanno messo in luce quanto sia importante affrontare con attenzione il rapporto tra intelligenza artificiale e scuola. La filosofa Vallor²⁰, ad esempio, sottolinea che non basta costruire algoritmi “etici per definizione”: ciò che conta davvero è come questi strumenti vengono utilizzati nei contesti concreti, perché il loro impatto dipende anche dalle regole, dalle relazioni sociali e dagli obiettivi educativi. Vallor propone quindi un approccio che coinvolga non solo informatici, ma anche insegnanti, giuristi e sociologi, per garantire un uso responsabile della tecnologia fin dalla fase di progettazione.

Anche la sociologa Ruha Benjamin²¹ invita a non affidarsi ciecamente alla tecnologia, soprattutto quando si tratta di decisioni che riguardano le persone, afferma che è necessario creare forme di controllo trasparenti e partecipate, simili a quelle usate nella medicina per valutare i trattamenti sperimentali. Se riportiamo queste riflessioni in ambito scolastico potremmo dire che le scuole dovrebbero coinvolgere anche studenti, famiglie e comunità locali nella scelta e nella valutazione degli strumenti di IA.

Anche altri studiosi si muovono su questa linea e propongono una vera e propria “governance” dell'intelligenza artificiale in ambito scolastico. Per loro non è realistico pensare che le scuole, da sole, possano controllare strumenti così complessi, né che tutto debba essere deciso centralmente dai ministeri. La soluzione, dicono, sta in un lavoro di squadra: con standard condivisi, verifiche indipendenti e più formazione per i docenti, così che possano capire davvero come funzionano gli algoritmi che usano.

²⁰ S. Vallor, *Technology and the virtues: A philosophical guide to a future worth wanting*, Oxford University Press, Oxford 2016.

²¹ R. Benjamin, *Race after technology: Abolitionist tools for the new Jim code*, Polity Press, Cambridge 2019.

Infine, lo studioso Paul Prinsloo²² mette al centro la responsabilità nella gestione dei dati scolastici. Ricorda che l'autonomia didattica è un valore importante, ma non può trasformarsi in solitudine decisionale. Quando si introducono tecnologie che incidono sul percorso formativo degli studenti, serve una vera cultura della cura, che tenga conto dei diritti, dei bisogni e delle differenze di tutti.

Da quanto detto finora, si evince che, in ogni caso, studenti e famiglie devono essere parte attiva del processo decisionale, conoscendo i rischi e le potenzialità dell'IA e contribuendo alla scelta degli strumenti utilizzati. Una corretta informazione e un dialogo costante tra scuole, istituzioni e aziende tecnologiche sono indispensabili per proteggere i diritti degli studenti e costruire un ambiente educativo sicuro e moderno. L'intelligenza artificiale applicata alla didattica si trova di fronte a un delicato equilibrio: da un lato, il principio del GDPR impone di limitare i dati raccolti, dall'altro, la personalizzazione didattica richiede grandi quantità di informazioni. Questo crea una sfida complessa, soprattutto per le scuole con poche risorse e competenze digitali limitate, che potrebbero persino rinunciare all'uso dell'IA per rispettare le normative. Oltre a regole precise, servono politiche pubbliche che supportino concretamente le scuole, tramite investimenti in tecnologia e formazione. Insieme al GDPR, normative come l'AI Act puntano a promuovere un'intelligenza artificiale affidabile, sicura ed etica, tutelando diritti fondamentali e interessi pubblici, tuttavia, le procedure per rispettare queste regole, rischiano di rallentare l'adozione dell'IA nelle scuole. Questa complessità, d'altra parte, può trasformarsi in un'occasione per cercare soluzioni nuove, capaci di garantire un utilizzo responsabile e sicuro dell'intelligenza artificiale nella didattica.

6. Possibili soluzioni

Le scuole potrebbero adottare in modo sistematico la valutazione d'impatto sulla protezione dei dati (DPIA), uno strumento prezioso per garantire che le informazioni personali siano gestite nel pieno rispetto delle normative vigenti. Questo processo tiene conto non solo dei rischi tecnici, ma anche delle implicazioni etiche, sociali e didattiche delle scelte compiute. Un approccio del genere favorisce uno sviluppo tecnologico che coniuga efficienza e responsabilità. Tuttavia, per adeguarsi davvero al GDPR, è necessario un cambiamento culturale all'interno delle scuole, che coinvolga in maniera attiva dirigenti, insegnanti, studenti e famiglie nella gestione consapevole dei dati. L'obiettivo non è solo rispettare le regole, ma anche costruire un clima di fiducia attorno all'uso degli strumenti digitali nella didattica. Le scuole, però, non operano isolate, ma fanno parte di un sistema più ampio, con forti implicazioni sociali ed economiche. È quindi fondamentale che partecipino attivamente alla definizione di politiche future, individuando anche eventuali disuguaglianze legate all'uso dell'IA e collaborando per ridurle. Se un domani fosse richiesto di valutare l'impatto sociale delle tecnologie usate in classe, la scuola avrebbe un ruolo centrale nell'evidenziare disparità e criticità.

In ambito educativo piattaforme affidabili, senza pregiudizi o discriminazioni, sono fondamentali per garantire standard minimi di sicurezza; le autorità di vigilanza dovrebbero garantire i controlli e monitorare il rispetto delle norme oltre ad offrire indicazioni pratiche alle scuole. In questo contesto complesso, la collaborazione tra scuole, enti di controllo e

²² P. Prinsloo, S. Slade, *Ethics and Learning Analytics: Charting the (Un)Charted*, in C. Lang, G. Siemens, A. Wise, D. Gašević (a cura di), *Handbook of Learning Analytics*, Society for Learning Analytics and Research, 2017.

aziende tecnologiche diventa indispensabile, soprattutto per affrontare sfide delicate come l'uso di dati biometrici o sistemi di riconoscimento facciale per la gestione della sicurezza scolastica. Anche se queste tecnologie possono apportare numerosi vantaggi, vanno sempre valutate attentamente sotto il profilo etico. Inoltre, è essenziale che gli studenti, ma anche i genitori e i docenti, acquisiscano consapevolezza dei rischi legati all'IA e sviluppino le competenze necessarie per muoversi in una società digitale; tutti gli stakeholders devono essere protagonisti attivi del dibattito sull'uso della tecnologia e non semplici utilizzatori passivi, per questo è fondamentale integrare nei programmi scolastici l'educazione alla protezione dei dati, alla sicurezza informatica e all'uso critico delle tecnologie digitali. La tecnologia, infatti, deve restare uno strumento e non diventare il fine ultimo della didattica. Senza una riflessione attenta, si rischia di perdere il controllo e di generare nuovi problemi.

In definitiva, l'introduzione dell'IA nella scuola può offrire grandi opportunità, ma solo se accompagnata da solide basi etiche e giuridiche. Norme come il GDPR non devono essere viste come ostacoli, ma come garanzie fondamentali per proteggere i diritti di tutti e costruire un sistema educativo capace di far crescere persone preparate e consapevoli, pronte ad affrontare le sfide del futuro.