

PROFILI GIURIDICI DELL'UTILIZZO DELL'INTELLIGENZA ARTIFICIALE NELLA DIDATTICA

Elettra Stradella, *Università di Pisa*, Gabriele Rugani, *Università di Pisa*¹

1. Introduzione

L'introduzione dell'Intelligenza artificiale (IA) nel settore educativo rappresenta una delle sfide più complesse e affascinanti del nostro tempo, in cui innovazione tecnologica e tutela dei diritti fondamentali si intrecciano in modo inedito. La crescente diffusione di sistemi di IA per la valutazione delle competenze, l'orientamento dei percorsi formativi e il supporto all'insegnamento ha reso necessario un quadro normativo capace di bilanciare le opportunità offerte dalla tecnologia con la protezione dei valori costituzionali ed europei, quali eguaglianza, non discriminazione e libertà di apprendimento.

L'Unione europea, con l'adozione del *Regolamento (UE) 2024/1689* (Parlamento europeo, Consiglio dell'Unione europea 2024) – il cosiddetto «*AI Act*» (AIA) – ha fornito per la prima volta un quadro giuridico organico per lo sviluppo e l'impiego di sistemi di IA, classificando le applicazioni ad alto rischio e imponendo requisiti di trasparenza, tracciabilità e supervisione umana. Il settore educativo rientra esplicitamente tra quelli considerati ad alto impatto sociale e ciò rende necessario interrogarsi sulle implicazioni etiche e giuridiche dell'uso di algoritmi nel contesto scolastico e universitario, rispetto al quale, come si vedrà, restano aperti non pochi interrogativi, e le soluzioni offerte non possono considerarsi del tutto soddisfacenti ed esaustive.

Il contributo affronta questo snodo critico, a partire dalle riflessioni condotte e dagli stimoli ricevuti nell'ambito del XXXI Convegno nazionale ILSA, proponendo un'analisi della disciplina europea e della sua capacità di rispondere alle principali preoccupazioni legate all'introduzione dell'IA nella didattica: il rischio di discriminazioni algoritmiche e *bias* nei sistemi di valutazione; la tutela dei dati personali e sensibili degli studenti; l'impatto sull'autonomia didattica dei docenti e sulla libertà di insegnamento; la possibilità di garantire un controllo umano significativo (*human-in-the-loop*) nei processi decisionali automatizzati.

La riflessione si sviluppa lungo tre direttrici: in primo luogo, viene ricostruito il quadro normativo europeo e nazionale di riferimento, con particolare attenzione ai principi del diritto dell'educazione e del diritto antidiscriminatorio. In secondo luogo, si analizzano i rischi e le opportunità derivanti dall'uso dell'IA nella progettazione didattica e nelle modalità di valutazione. Infine, si propone una lettura critica del bilanciamento tra innovazione e diritti fondamentali, con l'obiettivo di delineare raccomandazioni per una *governance* dell'IA in educazione che sia coerente con i valori democratici e inclusivi.

In tal modo, il contributo si colloca nel dibattito contemporaneo sull'«etica dell'IA», offrendo un punto di vista giuridico che ambisce a essere utile non solo agli studiosi e agli operatori del diritto, ma anche ai docenti e a chi progetta tecnologie educative, nella prospettiva di un uso responsabile e sostenibile dell'innovazione tecnologica.

¹ Il presente articolo è stato realizzato nell'ambito del PRIN 2022 «*Artificial intelligence between Generating and tackling gender-based Discriminations*» (AiGeDi), Unità dell'Università di Pisa. Il paragrafo 1 è attribuibile a Stradella, mentre i paragrafi 2-8 sono attribuibili a Rugani.

2. Adozione del *Regolamento* sull'IA

Preliminarmente, si rende necessaria una panoramica generale del *Regolamento* sull'IA (Parlamento europeo, Consiglio dell'Unione europea 2024), con riferimento sia alla procedura che ha portato alla sua adozione, sia al suo contenuto essenziale.

In relazione all'*iter legis*, occorre ricordare che già in data 19 febbraio 2020, all'interno del *Libro bianco sull'intelligenza artificiale. Un approccio europeo all'eccellenza e alla fiducia*, la Commissione europea aveva preannunciato un'iniziativa finalizzata all'adozione di un «quadro normativo europeo chiaro», il quale «rafforzerebbe la fiducia dei consumatori e delle imprese nell'IA e accelererebbe quindi l'adozione della tecnologia» (Commissione europea 2020: 11). Così, a seguito di consultazione pubblica svoltasi dal 19 febbraio al 14 giugno 2020², in data 21 aprile 2021 la Commissione ha presentato la *Proposta di «Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'Intelligenza artificiale (legge sull'Intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione»* (Commissione europea 2021).

Come noto – ai sensi delle disposizioni del *Trattato sul funzionamento dell'Unione europea* (TFUE), che disciplinano la cosiddetta «procedura legislativa ordinaria», nello specifico gli artt. 289 e 294³ – a seguito della *Proposta* della Commissione europea (2021), era necessario che l'atto fosse approvato nella medesima formulazione da entrambi i «co-legislatori» dell'Unione europea, vale a dire il Consiglio – composto da rappresentanti degli Stati membri a livello ministeriale, ex art. 16 del *Trattato sull'Unione europea* (TUE)⁴ – e il Parlamento europeo, composto invece dai rappresentanti dei cittadini dell'UE come previsto dall'Art. 14 TUE⁵.

In data 6 dicembre 2022, il Consiglio ha adottato il proprio «Orientamento generale» sulla *Proposta*, mentre in data 14 giugno 2023 il Parlamento europeo ha adottato ad ampia maggioranza la propria posizione negoziale⁶. Per semplificare, si potrebbe dire che se il Consiglio è parso più attento a promuovere lo sviluppo dei sistemi di IA e a rendere il testo più vicino alle esigenze delle imprese, il Parlamento europeo è sembrato invece più sensibile alle istanze di protezione dei diritti dei cittadini e dei consumatori (cfr. Rugani 2023c).

Per poter giungere all'approvazione dell'atto, era dunque necessario trovare un punto di incontro, negoziando un testo condiviso. Tale accordo politico tra i due co-legislatori è stato infine raggiunto in data 9 dicembre 2023 ed è stato definito dalla Presidente della Commissione europea Ursula von der Leyen come «un momento storico»⁷. Così, a seguito degli ultimi passaggi formali, in data 13 giugno 2024 è stato finalmente adottato il *Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale* e modifica alcuni precedenti atti UE (Parlamento europeo, Consiglio dell'Unione europea 2024).

² Per approfondimenti sul tema, cfr. Rugani 2023a, 2023b.

³ Cfr., fra le altre, la pagina del sito ufficiale del Consiglio europeo e del Consiglio dell'Unione europea dedicata alla procedura legislativa ordinaria (<https://www.consilium.europa.eu/it/council-eu/decision-making/ordinary-legislative-procedure/>).

⁴ Cfr., fra le altre, la pagina del sito ufficiale del Consiglio europeo e del Consiglio dell'Unione europea dedicata proprio al Consiglio dell'Unione europea (<https://www.consilium.europa.eu/it/council-eu/>).

⁵ Cfr., fra le altre, la pagina del sito ufficiale del Parlamento europeo intitolata «Tutto sul Parlamento» (<https://www.europarl.europa.eu/about-parliament/it>).

⁶ Su quest'ultima in particolare, cfr. fra gli altri, Aszódi 2023; Bracy 2023; Nahra *et al.* 2023.

⁷ Cfr., fra le altre, la pagina del sito ufficiale della Commissione europea dal titolo «Soddisfazione della Commissione per l'accordo politico sulla legge sull'intelligenza artificiale», pubblicata in data 9.12.2023 e aggiornata in data 16.01.2025 (https://ec.europa.eu/commission/presscorner/detail/it/ip_23_6473).

È bene ricordare che lo strumento giuridico in questione, trattandosi per l'appunto di un Regolamento, ai sensi dell'art. 288 TFUE ha portata generale, è obbligatorio in tutti i suoi elementi ed è direttamente applicabile in tutti gli Stati membri⁸. Tuttavia, in base a quanto previsto dal suo art. 113, l'AIA è solo in parte già applicabile nel momento in cui si scrive, mentre in relazione ad altre parti lo diventerà gradualmente, fino al 2 agosto 2027 (data in cui l'intero strumento sarà applicabile).

3. *Regolamento sull'IA: l'approccio basato sul rischio*

Ci si concentrerà ora sul contenuto del *Regolamento* (Parlamento europeo, Consiglio dell'Unione europea 2024)⁹. L'AIA – il cui «Oggetto» e il cui «Ambito di applicazione» sono specificati, rispettivamente, dai suoi artt. 1 e 2 – stabilisce le regole applicabili a un «sistema di IA», che viene definito dall'art. 3, n. 1, come

un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali (Parlamento europeo, Consiglio dell'Unione europea 2024).

Nel disciplinare tali sistemi di IA, l'AIA abbraccia quello che può essere definito come «approccio basato sul rischio» o *risk-based approach*¹⁰: vengono pertanto dettate regole diverse per i sistemi di IA a seconda che essi creino un rischio «inaccettabile», «alto», oppure «basso o minimo». Per le pratiche di IA da cui deriva un rischio inaccettabile, che sono tassativamente elencate, è previsto un radicale divieto (AIA, Capo II, *id est* art. 5). Si tratta di utilizzi impropri dell'IA, che permettono pratiche di manipolazione, sfruttamento e controllo sociale. Tali pratiche sono particolarmente dannose e abusive e vengono per l'appunto vietate poiché contrarie ai valori dell'Unione relativi al rispetto della dignità umana, alla libertà, all'uguaglianza, alla democrazia e allo Stato di diritto e ai diritti fondamentali sanciti dalla Carta, compresi il diritto alla non discriminazione, alla protezione dei dati e alla vita privata, nonché i diritti dei minori (AIA, Considerando 28). Il divieto di pratiche di IA da cui deriva un rischio inaccettabile è già applicabile dal 2 febbraio 2025 (AIA, *ex art.* 113, lett. a).

Vi sono poi i sistemi di IA da cui deriva un rischio alto (AIA, art. 6). Tra questi, figurano due categorie principali: da un lato, i sistemi di IA destinati ad essere utilizzati come componenti di sicurezza di prodotti soggetti a valutazione della conformità *ex ante* da parte di terzi (AIA, art. 6, par. 1); dall'altro, ulteriori sistemi indipendenti che presentano implicazioni principalmente in relazione ai diritti fondamentali (AIA, art. 6, par. 2). Questi ultimi sono esplicitamente indicati nell'Allegato III al *Regolamento* (dedicato proprio ai «Sistemi di IA ad alto rischio, di cui all'art. 6, par. 2») e più precisamente si tratta di sistemi di IA elencati in certi settori particolarmente importanti e delicati.

Orbene, i sistemi di IA ad alto rischio possono essere immessi sul mercato dell'Unione, messi in servizio o utilizzati, ma solo qualora soddisfino stringenti requisiti obbligatori, stabiliti

⁸ Cfr., fra le altre, la pagina del sito ufficiale dell'Unione europea dal titolo «Regolamenti, direttive e altri atti» (https://european-union.europa.eu/institutions-law-budget/law/types-legislation_it).

⁹ Per approfondimenti, cfr. Pehlivan, Forgó, Valcke 2024.

¹⁰ Per approfondimenti, cfr., fra gli altri, De Gregorio, Dunn 2022.

dal Capo III (artt. 6-49) dell'AIA. In particolare, la Sezione 2 del Capo III dell'AIA (artt. 8-15) sancisce i «Requisiti per i sistemi di IA ad alto rischio», tra cui sistemi di gestione dei rischi (art. 9), elevata qualità dei *dataset* (art. 10), documentazione dettagliata (art. 11), conservazione delle registrazioni delle attività (art. 12), trasparenza e informazioni chiare (art. 13), sorveglianza umana (art. 14), accuratezza, robustezza e cybersicurezza (art. 15).

La Sezione 3 del Capo III dell'AIA (artt. 16-27) è poi dedicata a «Obblighi dei fornitori e dei *deployer* dei sistemi di IA ad alto rischio e di altre parti». La differenza tra le due figure è la seguente: il fornitore è la persona fisica o giuridica, l'autorità pubblica, l'agenzia o l'altro organismo che sviluppa un sistema di IA o che lo fa sviluppare e lo immette sul mercato o lo mette in servizio con il proprio nome o marchio, a titolo oneroso o gratuito (art. 3, n. 3); il *deployer*, invece, è la persona fisica o giuridica, l'autorità pubblica, l'agenzia o l'altro organismo che utilizza un sistema di IA sotto la propria autorità, tranne nel caso in cui il sistema di IA sia utilizzato nel corso di un'attività personale non professionale (art. 3, n. 4). A titolo esemplificativo, tra le responsabilità di quest'ultimo rientra, in alcuni casi, l'effettuazione di una valutazione dell'impatto sui diritti fondamentali che l'uso del sistema di IA può produrre (art. 27).

Quanto alla Sezione 4 (artt. 28-39) e alla Sezione 5 (artt. 40-49) del Capo III dell'AIA, esse riguardano, rispettivamente, «Autorità di notifica e organismi notificati» e «Norme, valutazione della conformità, certificati, registrazione».

Invece, i sistemi che non rientrano nelle prime due categorie, individuati dunque in via residuale, presentando rischio basso o minimo, devono solo rispettare in determinati casi alcuni obblighi di trasparenza (AIA, Capo IV, *id est* art. 50).

4. Regolamento sull'IA: modelli di IA per finalità generali, tra cui quelli generativi

Un discorso a parte lo merita il concetto di «modello di IA per finalità generali», definito dall'art. 3, n. 63, dell'AIA come

un modello di IA, anche laddove tale modello di IA sia addestrato con grandi quantità di dati utilizzando l'auto supervisione su larga scala, che sia caratterizzato da una generalità significativa e sia in grado di svolgere con competenza un'ampia gamma di compiti distinti, indipendentemente dalle modalità con cui il modello è immesso sul mercato, e che può essere integrato in una varietà di sistemi o applicazioni a valle, ad eccezione dei modelli di IA utilizzati per attività di ricerca, sviluppo o prototipazione prima di essere immessi sul mercato (Parlamento europeo, Consiglio dell'Unione europea 2024).

Un tipico esempio è rappresentato dai grandi modelli di IA generativi, i quali «consentono una generazione flessibile di contenuti, per esempio sotto forma di testo, audio, immagini o video, che possono prontamente rispondere a un'ampia gamma di compiti distinti» (AIA Considerando 99).

Un esempio perfetto di IA per finalità generali, e segnatamente di IA generativa, è rappresentato dal noto *ChatGPT*, utilizzato al giorno d'oggi da milioni di persone: si tratta di un modello addestrato su quantità di dati senza precedenti e capace di sostenere conversazioni straordinariamente varie, che spaziano dalla scrittura di poesie alla valutazione di temi scolastici; in una vasta gamma di contesti e casi d'uso, è sufficiente porre una domanda in linguaggio naturale per ottenere una risposta fluida e convincente. Si potrebbero però citare anche altri esempi, come *Gemini* e *DeepSeek*.

Sul punto, occorre ricordare come, in realtà, i modelli di IA di cui sopra non fossero stati affatto presi in considerazione nell'originaria *Proposta* della Commissione europea (2021). Ciò non deve sorprendere: nel 2021, l'IA generativa non era ancora al centro dell'attenzione come pochi anni dopo. È stato poi il Parlamento europeo, con la sua posizione negoziale del 2023, a indirizzare il dibattito verso tale tema.

Nel disciplinare l'IA per finalità generali (e quindi anche quella generativa) vi è tuttavia una rilevante difficoltà: essa mette seriamente in discussione l'intero approccio basato sul rischio abbracciato dall'AIA. Secondo tale logica, infatti, un sistema di IA viene classificato all'interno delle categorie di rischio sulla base della *finalità* prevista dal fornitore. Ma l'IA per finalità generali, come dice l'espressione stessa, non ha una finalità predefinita: non è dunque il fornitore, ma l'utente a determinare come il sistema sarà impiegato; gli utenti finali, tuttavia, sono fondamentalmente esclusi dall'ambito di applicazione dell'AIA (Helberger, Diakopoulos 2023).

All'interno del *Regolamento*, tale problematica non è stata affrontata facendo rientrare l'IA per finalità generali all'interno della categoria del rischio alto, come aveva suggerito taluno, bensì creando una categoria *ad hoc*. Quest'ultima è disciplinata all'interno del Capo V (artt. 51-56), dedicato proprio ai «Modelli di IA per finalità generali», il quale stabilisce per i fornitori di tali modelli alcuni obblighi, specialmente di trasparenza e rispetto del diritto d'autore (AIA, art. 53). Sono tuttavia previsti ulteriori obblighi qualora i modelli di cui sopra presentino un «rischio sistemico», cioè

un rischio specifico per le capacità di impatto elevato dei modelli di IA per finalità generali, avente un impatto significativo sul mercato dell'Unione a causa della sua portata o di effetti negativi effettivi o ragionevolmente prevedibili sulla salute pubblica, la sicurezza, i diritti fondamentali o la società nel suo complesso, che può propagarsi su larga scala lungo l'intera catena del valore»¹¹ (Parlamento europeo, Consiglio dell'Unione europea 2024).

In tali casi, i fornitori devono, in particolare, valutare e attenuare i rischi; tenere traccia e riferire circa gli eventuali incidenti gravi; effettuare test; garantire un livello adeguato di cybersicurezza (AIA, art. 55)¹².

5. *Regolamento* sull'IA: disposizioni che disciplinano la didattica (i sistemi vietati)

Fatta una tale panoramica, è ora possibile concentrarsi sulle disposizioni del *Regolamento* (Parlamento europeo, Consiglio dell'Unione europea 2024) concernenti la didattica. Occorre infatti evidenziare che, alla luce dell'importanza e della delicatezza di tale ambito, determinati utilizzi dell'IA nella didattica – ritenuti inaccettabili – vengono radicalmente vietati, mentre altri vengono qualificati come ad alto rischio e, conseguentemente, sottoposti ai relativi obblighi¹³.

Partendo dai sistemi a rischio inaccettabile, l'art. 5, par. 1, lett. f, dell'AIA vieta espressamente l'immissione sul mercato, la messa in servizio per tale finalità specifica o l'uso di sistemi di IA per «inferire le emozioni» di una persona fisica non solo sul luogo di lavoro, ma anche «nell'ambito degli istituti di istruzione tranne laddove l'uso del sistema di IA sia destinato a

¹¹ Cfr. art. 3, n. 65.

¹² Cfr., fra le altre, la pagina del sito ufficiale della Commissione europea dal titolo «Legge sull'IA» (<https://digital-strategy.ec.europa.eu/it/policies/regulatory-framework-ai>).

¹³ Per approfondimenti sull'uso dell'IA nella didattica, cfr., fra gli altri, Chen, Chen, Lin 2020.

essere messo in funzione o immesso sul mercato per motivi medici o di sicurezza» (Parlamento europeo, Consiglio dell'Unione europea 2024).

Sul punto, occorre infatti ricordare che sussistono serie preoccupazioni in merito alla base scientifica dei sistemi di IA volti a identificare o inferire emozioni, in particolare perché l'espressione delle emozioni varia notevolmente in base alle culture e alle situazioni e persino in relazione a una stessa persona. Tra le principali carenze di tali sistemi figurano la limitata affidabilità, la mancanza di specificità e la limitata generalizzabilità. Pertanto, i sistemi di IA che identificano o inferiscono emozioni o intenzioni di persone fisiche sulla base dei loro dati biometrici possono portare a risultati discriminatori e possono essere invasivi dei diritti e delle libertà delle persone interessate.

Considerando lo squilibrio di potere in contesti quali il lavoro o, per l'appunto, l'istruzione, combinato con la natura invasiva di tali sistemi, questi ultimi potrebbero determinare un trattamento pregiudizievole o sfavorevole di talune persone fisiche o di interi gruppi di persone fisiche (AIA, Considerando 44). Dalle circostanze di cui sopra discende, evidentemente, la necessità di un radicale divieto di tali sistemi nei contesti in questione.

6. Regolamento sull'IA: disposizioni che disciplinano la didattica (sistemi ad alto rischio)

Ancor più rilevanti, quantomeno dal punto di vista quantitativo, sono poi gli utilizzi dell'IA nella didattica ritenuti ad alto rischio.

Come meglio illustrato sopra, ai sensi dell'Allegato III del *Regolamento* (Parlamento europeo, Consiglio dell'Unione europea 2024), i sistemi di IA ad alto rischio *ex art.* 6, par. 2 – cioè sistemi che presentano implicazioni principalmente in relazione ai diritti fondamentali – sono quelli elencati in uno dei settori indicati dall'Allegato medesimo. Orbene, tra tali settori figura anche l'istruzione e la formazione professionale, e, in particolare, sono elencati i seguenti sistemi di IA:

- a) i sistemi di IA destinati a essere utilizzati per determinare l'accesso, l'ammissione o l'assegnazione di persone fisiche agli istituti di istruzione e formazione professionale a tutti i livelli;
- b) i sistemi di IA destinati a essere utilizzati per valutare i risultati dell'apprendimento, anche nei casi in cui tali risultati sono utilizzati per orientare il processo di apprendimento di persone fisiche in istituti di istruzione o formazione professionale a tutti i livelli;
- c) i sistemi di IA destinati a essere utilizzati per valutare il livello di istruzione adeguato che una persona riceverà o a cui potrà accedere, nel contesto o all'interno di istituti di istruzione o formazione professionale a tutti i livelli;
- d) i sistemi di IA destinati a essere utilizzati per monitorare e rilevare comportamenti vietati degli studenti durante le prove nel contesto o all'interno di istituti di istruzione e formazione professionale a tutti i livelli (Parlamento europeo, Consiglio dell'Unione europea 2024).

In buona sostanza, da un lato, il *Regolamento* riconosce che la diffusione dei sistemi di IA nell'istruzione è importante per promuovere un'istruzione e una formazione digitali di alta qualità, nonché per consentire a tutti i discenti e gli insegnanti di acquisire e condividere le competenze e le abilità digitali necessarie, compresa l'alfabetizzazione mediatica e il pensiero critico, per partecipare attivamente all'economia, alla società e ai processi democratici. Per tale ragione, non è quindi previsto un divieto *tout court*. Dall'altro lato, tuttavia, l'AIA non trascura il

fatto che i sistemi di IA utilizzati nell'istruzione o nella formazione professionale, in particolare quelli di cui sopra, possono determinare il percorso d'istruzione e professionale della vita di una persona e quindi possono incidere sulla sua capacità di garantire il proprio sostentamento.

Pertanto, se progettati e utilizzati in modo inadeguato, tali sistemi possono essere particolarmente intrusivi e violare il diritto all'istruzione e alla formazione, nonché il diritto alla non discriminazione, e perpetuare modelli storici di discriminazione, per esempio nei confronti delle donne, di talune fasce di età, delle persone con disabilità o delle persone aventi determinate origini razziali o etniche o un determinato orientamento sessuale (AIA, Considerando 56). Da ciò deriva la necessità di classificare simili sistemi di IA come ad alto rischio, con la conseguente applicazione dei rigorosi obblighi meglio esposti (cfr. par. 3).

7. *Regolamento sull'IA: problematiche dell'IA nella didattica (rischi di discriminazione)*

Da quanto esposto, appare chiaro come *Regolamento* (Parlamento europeo, Consiglio dell'Unione europea 2024) – attraverso la radicale proibizione di alcuni sistemi intollerabili e la previsione di stringenti condizioni per altri sistemi ritenuti comunque pericolosi – cerchi di eliminare, o quantomeno mitigare, i principali rischi derivanti dall'utilizzo dell'IA nella didattica. Tra tali rischi spiccano indubbiamente quelli di discriminazione (che infatti sono già stati citati più di una volta). Non bisogna infatti dimenticare che – come dimostrato da analisi statistiche (World Economic Forum 2018; UNESCO 2020), studi accademici (tra cui il progetto «Gender Shades» sulle tecniche di riconoscimento facciale¹⁴), nonché note vicende (si pensi a quella dell'algoritmo progettato da *Amazon* per il reclutamento¹⁵) – l'IA può produrre effetti discriminatori ai danni delle donne, nonché di vari gruppi minoritari e vulnerabili (cfr., p. es. Nardocci 2021): ci si riferisce a persone, per esempio, appartenenti a minoranze etnico-razziali, con diverse abilità, non eterosessuali, oppure «non cisgender»¹⁶.

Particolarmente penalizzati risultano gli individui che presentano contestualmente più fattori di discriminazione, per esempio una donna appartenente a una minoranza etnico-razziale. Come ormai noto, si parla in questi casi di «discriminazioni intersezionali». Tale concetto, di origine nord-americana (Crenshaw 1989), è stato successivamente preso in considerazione anche dalle Nazioni Unite (CEDAW 2010), nonché dalle istituzioni dell'Unione europea (Parlamento europeo 2022)¹⁷.

Tornando all'IA, le cause per cui essa può produrre gli effetti discriminatori di cui sopra possono risiedere in diversi fattori¹⁸. Tra questi, il più significativo è senz'altro il *data training*, ovvero la raccolta e la selezione dei dati, poiché più essi sono parziali o frutto di pregiudizi, più il modello ne rifletterà la parzialità, agendo in modo discriminatorio. A questo proposito si usa

¹⁴ Per approfondimenti sulle tecniche di riconoscimento facciale, cfr. Buolamwini, Gebru 2018; Buolamwini 2019.

¹⁵ Cfr., fra gli altri, Dustin 2018; Lauret 2019.

Secondo la definizione dell'Enciclopedia Treccani, con l'espressione «cisgender» si intende una «persona nella quale collimano il sesso biologico e l'identità di genere». Pertanto, «non cisgender» è una persona in cui non vi è tale corrispondenza (https://www.treccani.it/enciclopedia/cisgender_%28altro%29/).

¹⁷ In particolare, con la Risoluzione del 6 luglio 2022 dedicata proprio alla «discriminazione intersezionale», il Parlamento europeo ha definito la medesima come «una situazione nella quale intervengono diversi motivi di discriminazione che interagiscono tra loro, per esempio il genere con altri motivi di discriminazione quali razza, colore della pelle, status etnico o socioeconomico, età, orientamento sessuale, identità ed espressione di genere, caratteristiche sessuali, caratteristiche genetiche, religione o credo, nazionalità, status di residente, provenienza da un contesto migratorio o disabilità, tra gli altri, in modo tale da essere indissociabili e da produrre tipologie specifiche di discriminazione» (Parlamento europeo 2022).

¹⁸ Sull'argomento, cfr., fra gli altri, Barocas, Selbst 2016; Borgesius 2018; Zuddas 2020.

parlare di regola del «*Garbage In, Garbage Out*» (GIGO)¹⁹ (Resta 2019; Simoncini 2019), vi sono però anche altri fattori.

In ogni caso, indipendentemente dalle specifiche cause, tale circostanza può avere gravi ripercussioni in diversi ambiti, come quello lavorativo, medico, giudiziario e bancario; tra questi vi è anche il settore dell'istruzione e della formazione professionale, dove gli algoritmi possono perpetuare – o persino amplificare – le discriminazioni esistenti, finendo per rafforzare le disuguaglianze sociali a danno di alcuni segmenti della popolazione. Donne o persone appartenenti a minoranze potrebbero, per esempio, essere escluse da istituzioni educative e formative o essere valutate in modo più sfavorevole.

L'AIA, attraverso le previsioni meglio illustrate sopra, ha dunque il rilevante merito di cercare di contrastare tale fenomeno. I sistemi che presentano i rischi di discriminazione più gravi rientrano infatti tra quelli vietati. I sistemi da cui possono derivare pericoli comunque elevati, invece, sono sottoposti ad obblighi, numerosi dei quali finalizzati proprio a ridurre i rischi di discriminazione (cfr., fra gli altri, Meding 2025): tra questi, soprattutto quelli concernenti i dati e la loro *governance*, che impongono pratiche tra cui un esame atto a valutare le possibili distorsioni suscettibili di incidere sulla salute e sulla sicurezza delle persone, di avere un impatto negativo sui diritti fondamentali o di comportare discriminazioni vietate dal diritto dell'Unione (AIA, art. 10, par. 2, lett. f), nonché misure adeguate per individuare, prevenire e attenuare le possibili distorsioni di cui sopra (AIA, art. 10, par. 2, lett. g); nonché quelli in materia di accuratezza, robustezza e cybersicurezza, che prescrivono *inter alia* di sviluppare sistemi di IA in modo tale da eliminare o ridurre il più possibile il rischio di *output* potenzialmente distorti (AIA, art. 15, par. 4).

8. Conclusioni: principali sfide relative all'impiego dell'IA nella didattica

Il presente contributo, dopo aver svolto una panoramica generale sia dell'*iter legis* che ha portato all'adozione *Regolamento* (Parlamento europeo, Consiglio dell'Unione europea 2024), sia del contenuto dell'AIA, si è soffermato più nello specifico sulle disposizioni di quest'ultimo dedicate all'impiego dell'IA nella didattica. In particolare, è stato salutato con favore il fatto che gli utilizzi dell'IA in tale settore da cui derivano i maggiori rischi, specialmente di discriminazione, sono stati o radicalmente proibiti, oppure sottoposti a precise regole, varie delle quali finalizzate proprio a mitigare il pericolo di discriminazioni. Tuttavia, non si può negare come vi siano aspetti concernenti l'impiego dell'IA nella didattica che non sono stati presi sufficientemente in considerazione e che richiederebbero un intervento normativo più approfondito. Tra questi, vi è l'utilizzo dell'IA generativa.

Nel campo dell'istruzione, infatti, si assiste a un uso crescente di sistemi di IA generativa, funzionali a fornire strumenti di apprendimento personalizzato, generazione di contenuti ed efficienza amministrativa. Essi sono utilizzati sia dagli studenti che dagli insegnanti per creare saggi, spiegare concetti o simulare problemi reali. Tuttavia, la loro adozione comporta rischi significativi che richiedono attenzione. Una delle preoccupazioni è l'uso improprio, come il fatto che gli studenti si affidino ai contenuti generati dall'IA per evitare di apprendere o creino

¹⁹ La regola fa riferimento al fatto che la qualità dei risultati di un sistema dipende dalla qualità dei dati inseriti. Se i dati forniti sono errati o di bassa qualità («spazzatura in ingresso»), si otterranno risultati altrettanto sbagliati o scadenti («spazzatura in uscita»), anche se il sistema funziona perfettamente.

output non etici, come informazioni distorte o false. Un altro rischio è la *privacy* dei dati, poiché gli strumenti di IA generativa spesso elaborano grandi quantità di informazioni personali, esponendo potenzialmente i dati sensibili a violazioni o usi impropri. Inoltre, la mancanza di trasparenza sul funzionamento di questi sistemi complica l'attribuzione di responsabilità: se si verificano errori, come contenuti inaccurati o pregiudizi non intenzionali, le salvaguardie per affrontarli sono limitate.

Orbene, è lecito dubitare del fatto che l'AIA affronti tali problemi in modo sufficientemente efficace. Come già esposto, da un lato, l'IA generativa non rientra nella categoria dell'alto rischio: pur essendo dettate per i modelli di IA per finalità generali delle norme *ad hoc* dal Capo V, esse di certo non raggiungono il grado di pervasività degli obblighi di cui al Capo III. Dall'altro lato, gli utenti finali – nel caso della didattica, tanto docenti quanto studenti – che sono quelli che determinano le finalità per cui i sistemi in questione saranno impiegati, sono fondamentalmente esclusi dall'ambito di applicazione dell'AIA.

In estrema sintesi, il *Regolamento* sull'IA rappresenta indubbiamente un primo passo nel senso di eliminare, o quantomeno disciplinare, alcuni degli utilizzi più pericolosi dell'IA nella didattica, soprattutto quelli da cui possono derivare effetti discriminatori. Al contempo, alcune questioni, soprattutto quella dell'IA generativa, non sono state sufficientemente prese in considerazione: un nuovo intervento in materia, nel prossimo futuro, è senz'altro da ritenersi auspicabile, prevedendo sul punto una disciplina più completa e obblighi più incisivi, o prendendo eventualmente in considerazione anche l'ipotesi di responsabilizzare maggiormente gli utenti finali.

Riferimenti bibliografici

- Aszódi N. 2023. *EU Parliament vote on AI Act: Lawmakers chose to protect people against harms of AI systems*. «Algorithmwatch.org».
- Commissione europea 2020. *Libro bianco della Commissione europea sull'intelligenza artificiale. Un approccio europeo all'eccellenza e alla fiducia*. Bruxelles, 19.02.2020, COM(2020) 65 final.
- Commissione europea 2021. *Proposta di «Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'Intelligenza artificiale (legge sull'Intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione»*. Bruxelles, 21.4.2021. COM(2021) 206 final. 2021/0106(COD).
- Barocas, S., Selbst, A. D. 2016. *Big data disparate impact*. «California Law Review», 104: 671-732.
- Borgesius, F. Z. 2018. *Discrimination, Artificial Intelligence, and Algorithmic Decision-Making*. Strasbourg. Council of Europe Publications.
- Bracy, J. 2023. *European Parliament Vote Pushes AI Act Significant Step Forward*. «Iapp.org».
- Buolamwini, J. 2019. *Hearing on: Artificial Intelligence: Societal and Ethical Implications*. Washington DC. United States House Committee on Science, Space and Technology.
- Buolamwini, J., Gebu, T. 2018. *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*. «Proceedings of Machine Learning Research», 81: 77-91.
- CEDAW (United Nations Committee on the Elimination of Discrimination against Women) 2010. *General recommendation No. 28 on the Core Obligations of States Parties Under Article 2 of the Convention on the Elimination of All Forms of Discrimination against Women*. CEDAW/C/GC/28.

- Chen, L., Chen, P., Lin, Z. 2020. *Artificial Intelligence in Education: A Review*. «IEEE Access», 8: 75264-75278.
- Crenshaw, K. 1989. *Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics*. «University of Chicago Legal Forum», 1: 139-167.
- De Gregorio, G., Dunn, P. 2022. *The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age*. «Common Market Law Review», 59(2): 473-500.
- Dustin, J. 2018. *Amazon Scraps Secret AI Recruiting Tool that Showed Bias Against Women*. «Reuters».
- Helberger, N., Diakopoulos, N. 2023. *ChatGPT and the AI Act*. «Internet Policy Review».
- Lauret, J. 2019. *Amazon's Sexist AI Recruiting Tool: How Did It Go So Wrong?*. «Becominghuman.ai».
- Meding, K. 2025. *It's Complicated. The Relationship of Algorithmic Fairness and Non-Discrimination Regulations in the EU AI Act*. «Arxiv.org», 2501.12962v6: 1-23.
- Nahra, K. J, Braun, M., Benizri, I., Jessani, A. A. 2023. *European Parliament Adopts Negotiating Position on the AI Act* «WilmerHale Privacy and Cybersecurity Law Blog».
- Nardocci, C. 2021. *Intelligenza artificiale e discriminazioni*. «Gruppo di Pisa», 3: 9-60.
- Parlamento europeo 2022. *Risoluzione del Parlamento europeo del 6 luglio 2022 sulla discriminazione intersezionale nell'Unione europea: la situazione socioeconomica delle donne di origine africana, mediorientale, latinoamericana e asiatica (2021/2243(INI))*. Strasburgo. Parlamento Europeo.
- Parlamento europeo, Consiglio dell'Unione europea 2024. *Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, del 13 giugno 2024, che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n. 300/2008, (UE) n. 167/2013, (UE) n. 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull'intelligenza artificiale)*. GU Serie L 2024/1689, 12.7.2024.
- Pehlivan, C. N., Forgó N., Valcke, P. 2024. *The EU Artificial Intelligence (AI) Act: A Commentary*. Alphen aan den Rijn. Wolters Kluwer.
- Resta, G. 2019. *Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza*. «Politica del diritto», 2: 199-236.
- Rugani, G. 2023a. *La consultazione dei portatori di interessi nell'elaborazione degli atti dell'Unione Europea in materia di intelligenza artificiale: il caso dell'alleanza europea per l'IA*. «Osservatorio sulle fonti», 2: 407-432.
- 2023b. *Potentialities and Margins for Improvement of the European AI Alliance, an Example of Participatory Democracy in the Field of AI at EU Level*. «Athena – Critical Inquiries in Law, Philosophy and Globalization», 3.2: 135-156.
- 2023c. *The European Parliament's Position on the AI Act: a human rights-oriented approach and an attempt to regulate generative AI systems*. In AA.VV. (a cura di), *Quaderni AISDUE – 2/2023*. Napoli. Editoriale Scientifica: 215-223.
- Simoncini, A. 2019. *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*. «BioLaw Journal – Rivista di BioDiritto», 1: 63-89.
- UNESCO (United Nations Educational, Scientific and Cultural Organization) 2020. *Artificial Intelligence and Gender Equality*. Paris. UNESCO.
- World Economic Forum 2018. *The Global Gender Gap Report*. Cologny/Geneva.
- Zuddas, P. 2020. *Intelligenza artificiale e discriminazioni*. «Consulta Online. Liber amicorum per Pasquale Costanzo»: 1-18.