GIUSTIZIA INSIEME

Diritto e innovazione

Indagini e giudizio fra regolamento (UE) 2024/1689 (*AI Act*) e legge n. 132 del 2025

di **Giovanni Canzio**

13 gennaio 2026

ABSTRACT

Il contributo analizza l'impatto dell'intelligenza artificiale sul sistema della giustizia penale alla luce del regolamento (UE) 2024/1689 (*AI Act*) e della legge italiana n. 132 del 2025, interrogandosi sulla compatibilità tra l'uso di strumenti algoritmici e i principi costituzionali del giusto processo. Muovendo dal rischio che modelli decisionali automatizzati possano compromettere la dimensione umana, dialogica e garantista dell'accertamento penale, il lavoro ricostruisce il quadro internazionale ed europeo della cosiddetta «sostenibilità digitale», evidenziando l'approccio antropocentrico e *risk-based* adottato dall'Unione europea. Particolare attenzione è dedicata alle pratiche di intelligenza artificiale vietate o ad alto rischio nelle attività di contrasto della criminalità, alle condizioni, limitazioni e deroghe previste dall'*AI Act* e alle ricadute sul regime delle indagini preliminari e della prova nel giudizio. Il

saggio mette in luce l'esigenza di un adeguamento puntuale della disciplina del processo penale interno, volto a garantire trasparenza, controllo umano significativo, contraddittorio effettivo e tutela dei diritti fondamentali. In tale prospettiva, la legge n. 132 del 2025 è esaminata criticamente quale primo passo verso la transizione digitale della giurisdizione, ma anche per le sue lacune regolatorie e per il rischio di una *governance* eccessivamente centralizzata e poco attenta all'indipendenza della magistratura. L'analisi si conclude sottolineando la necessità di un modello collaborativo uomo-macchina, capace di integrare l'innovazione tecnologica senza alterare il paradigma razionalista e garantista del processo penale.

Sommario: 1. **Intelligenza artificiale e giusto processo penale** – 2. **La comunità internazionale e la «*sostenibilità digitale*»** – 3. **Il ruolo dell'Europa e il regolamento (UE) 2024/1689 (*AI Act*)** – **4. Le attività di contrasto della criminalità fra pratiche di IA vietate o ad alto rischio: limitazioni e deroghe** – 5. **Il sistema processuale penale e il modello regolatorio dell'*AI Act*** – 6. **La legge n. 132 del 2025 e la transizione digitale**

1. Intelligenza artificiale e giusto processo penale

Lo statuto epistemico del *giusto processo* penale è disciplinato da una rete di principi costituzionali e regole di procedura, per i quali il rispetto della legalità del procedere e il metodo dialogico nell'acquisizione e valutazione delle prove per l'accertamento della verità, prevalgono sull'obiettivo del risultato, non postulando affatto che l'esito sia «esatto» bensì che sia «giusto», in termini di qualificata probabilità logica e di alta credibilità razionale della soluzione decisoria, il cui discorso giustificativo non è sottratto al sindacato di legalità e logicità.

L'intelligenza artificiale applicata al settore della giurisdizione può definirsi un sistema che, con specifico riguardo all'ambiente della giustizia, acquisisce, ordina e rielabora una enorme massa

di informazioni (*big data*), di tipo giudiziario o giurisprudenziale, per inferire, sulla base di calcoli algoritmici di tipo probabilistico, la migliore soluzione di una questione, in termini predittivi o addirittura decisori.

Si sostiene unanimemente che l'impetuoso e inarrestabile irrompere dei sistemi di intelligenza artificiale, in particolare nell'ambiente della giurisdizione penale, può costituire un serio rischio per il nucleo essenziale del giusto processo e, di conseguenza, per la dignità, i diritti e le libertà fondamentali della persona, quindi per l'efficacia della *rule of law* e della stessa democrazia, poiché i nuovi strumenti digitali (*machina sapiens*) sarebbero in grado non solo di supportare ma anche di sostituire i responsabili protagonisti umani.

Il modello «*forte*» di *decision-making*, pure caratterizzato da un'indubbia forza attrattiva a fronte della crisi di calcolabilità, uniformità e celerità della risposta del sistema giudiziario alla domanda di giustizia, comporterebbe, all'evidenza, il distopico venir meno della dimensione umana del «paradigma indiziario», proprio della tradizione razionalista occidentale**[1]**.

Orbene, siffatta rivoluzione, tecnologica, culturale ed epistemica, nella consapevolezza del pericolo di gravi distorsioni o effetti discriminatori dello stesso calcolo algoritmico (*bias automation*), ha determinato il preoccupato intervento della comunità internazionale, al fine di assicurare che l'implementazione computazionale delle fonti informative nell'amministrazione della giustizia si coniughi con le istanze di sorveglianza e responsabilità umane, sulla base di un modello «*collaborativo*» o «*ibrido*» che veda la complementarità uomo-macchina.

2. La comunità internazionale e la «*sostenibilità digitale*»

A partire dal *report* del 15 settembre 2021 dell'Alto Commissario per i Diritti Umani e da quello del 7 agosto 2024 del Segretario Generale (cui ha fatto seguito la risoluzione 79/1 dell'Assemblea Generale: *Global Digital Compact*), fino al più recente contributo del 16 luglio 2025 della *Special Rapporteur* M. Satterthwaite sulla indipendenza di giudici e avvocati, le Nazioni Unite, con riguardo all'avanzare del nuovo *epistemological power* e alla *black box nature* degli *ouputs* algoritmici dell'IA, sia predittiva che generativa, avvertono che, soprattutto nei sistemi di giustizia criminale, l'uso della stessa – oltre a produrre sicuri benefici in funzione efficientista – comporta l'elevato rischio di *potential harms* per il perimetro di tutela della dignità, dei diritti e delle libertà della persona.

Contro i potenziali rischi di *epistemic capture and even techno-authoritarianism* dei sistemi giudiziari è pressante il richiamo delle Linee Guida per l'uso di sistemi di IA nelle corti elaborate dall'UNESCO (Paris, 2025), accompagnato dalle raccomandazioni agli Stati membri di adottare «*clear guidelines, standards and safeguards that comply with international human rights law*». Si chiede il rispetto delle regole del *fair trial* e della *equality of arms* nella *disclosure* e nella contestazione difensiva della *AI generated evidence* prodotta dall'organo di accusa, insieme con la garanzia del diritto di accesso *to a human judge, competent, indipendent, impartial, and a human lawyer*, dei quali va curata una comune formazione culturale e professionale, *in multistakholder consultation and pilot programmes* e in collaborazione con idonei *technologist*.

Si afferma, in definitiva, dalle Nazioni Unite il principio generale per il quale «*use of AI must comply with human rights*».

3. Il ruolo dell'Europa e il regolamento (UE) 2024/1689 (*AI Act*)

Nello scenario di competizione globale che l'era digitale ha aperto fra almeno tre ecosistemi politici nell'innovazione tecnologica e nell'uso delle nuove forme di potere epistemico, le istituzioni europee – a fronte della più recente strategia deregolatoria e mercantile degli Stati Uniti e di quella di stringente dirigismo della Cina – stanno svolgendo un ruolo valoriale di straordinario rilievo per la definizione di un perimetro applicativo, giuridicamente ed eticamente accettabile, degli strumenti di intelligenza artificiale nel sistema di giustizia.

Con le *Linee Guida della Carta etica sull'uso dell'intelligenza artificiale nei sistemi giudiziari e nel loro ambiente*, elaborate il 3 dicembre 2018 dalla Commissione europea per l'efficienza dei sistemi di giustizia (*CEPEJ*), si è rimarcato con vigore lo speciale rilievo che rivestono, a garanzia della dignità, dei diritti e delle libertà fondamentali della persona, della legalità del procedere e della *rule of law* i principi di non discriminazione, trasparenza, imparzialità, equità, comprensibilità dei metodi di elaborazione dei dati, controllabilità dei percorsi di calcolo, qualità e attendibilità del risultato, in un quadro di permanente autonomia e responsabilità del decisore umano.

La stessa Commissione si avvia ad implementare i contenuti della Carta etica, ribadendo (nel *Draft* del 5 dicembre 2025 delle *Linee Guida sull'uso dell'intelligenza artificiale generativa per le corti*) i medesimi principi anche per i *local large language models* (*LLMs*), aggiungendo anzi che

l'approccio all'IA generativa, il cui uso risulta in costante aumento, dovrebbe ispirarsi ai principi di neutralità tecnologica e di sussidiarietà. Nel senso che siffatta tipologia di IA, siccome esposta al rischio di *hallucinations* cognitive prive di supporto reale, andrebbe impiegata solo in caso di inefficienza o ineffettività di soluzioni tecnologiche alternative, previe opportune valutazioni di rischio e di impatto seguite da un monitoraggio continuo.

Sono invero indiscutibili i benefici che possono trarsi nel contesto professionale mediante varie operazioni ausiliarie. Fermo restando l'obbligo di un'adeguata formazione del personale e degli utenti istituzionali, L'IA generativa potrebbe supportare la gestione dei carichi di lavoro, creare archivi giurisprudenziali tematici, identificare schemi argomentativi ricorrenti, rilevare contenziosi analoghi, facilitare l'assegnazione dei casi, generare automaticamente verbali, indici tematici, estrarre dati da documenti non strutturati, compilare *template* con informazioni chiave, fornire con *chatbot* specializzati informazioni procedurali ai cittadini, chiare e accurate, generare bozze di documenti standard, riassunti strutturati, analisi preliminari di atti di impugnazione, sintesi della giurisprudenza pertinente, mappe concettuali delle questioni controverse ecc.

Viceversa, l'uso dell'IA generativa dovrebbe essere radicalmente escluso per la valutazione critica della prova, per l'accertamento dei fatti, per l'apprezzamento della credibilità di una testimonianza, per ogni funzione che esiga il diretto esercizio dell'autorità giudiziaria: «*decision making, legal reasoning and evidentiary assesment should not be delegated to generative AI systems; such systems may never replace the human judgement of the judicial body*».

In questo contesto di *soft law* è intervenuto il regolamento (UE) del Parlamento europeo e del Consiglio 2024/1689 (*AI Act*) che, nella premessa della non breve, dettagliata e obiettivamente pesante architettura, avverte (art. 1, comma 1) che, pur promuovendosi l'innovazione tecnologica, s'intende assicurare la diffusione di una intelligenza artificiale «*antropocentrica, coerente, affidabile ed eticamente valida*», che garantisca un livello elevato di protezione dei diritti fondamentali sanciti dalla Carta dell'Unione, della democrazia e dello stato di diritto contro gli effetti nocivi dei sistemi di IA.

La dignità umana si erge dunque a bussola della innovazione tecnologica e della sostenibilità digitale.

Dei ben 180 *Considerando* del regolamento vanno menzionati alcuni che assumono specifico rilievo per l'ambiente della giurisdizione penale.

(42) In linea con la presunzione d'innocenza dell'imputato la persona fisica dovrebbe sempre essere giudicata in base al suo comportamento effettivo, non al comportamento previsto dall'IA basato unicamente sulla profilazione, sui tratti della personalità o su caratteristiche personali, senza che vi sia un ragionevole sospetto che essa sia coinvolta in un'attività criminosa sulla base di fatti oggettivi e verificabili e senza una valutazione umana. Dovrebbero essere vietate le valutazioni del rischio intese a determinare la probabilità che una persona commetta un reato unicamente sulla base della sua profilazione o dei tratti della personalità o di sue caratteristiche.

(59) Sono classificati ad alto rischio, nella misura in cui il loro uso è consentito, i sistemi di IA destinati ad essere utilizzati nel contesto di azioni delle autorità di contrasto, in cui l'accuratezza, l'affidabilità e la trasparenza risultano particolarmente importanti per evitare impatti negativi sui diritti fondamentali e per garantire la responsabilità, mezzi di ricorso efficaci, un giudice imparziale, la presunzione di innocenza e i diritti della difesa della persona indagata, attesa la difficoltà di ottenere informazioni significative sul funzionamento di tali sistemi e di confutarne i risultati in tribunale.

(61) Alcuni sistemi di IA destinati a essere utilizzati da un'autorità giudiziaria, per assisterla nelle attività di ricerca e interpretazione dei fatti e del diritto, dovrebbero essere classificati ad alto rischio per far fronte a potenziali distorsioni, errori e opacità. Il loro utilizzo può fornire sostegno al potere decisionale finale del giudice, ma non sostituirlo, dovendo esso rimanere un'attività «*a guida umana*»

(170)-(171) Vanno previsti mezzi di ricorso efficaci per le persone sui cui diritti e libertà incide negativamente l'uso dei sistemi di IA. Inoltre, le persone interessate hanno diritto di ottenere una spiegazione chiara qualora la decisione si basi principalmente sugli *output* di sistemi di IA ad alto rischio e incida significativamente sui diritti fondamentali.

Il regolamento adotta un approccio metodologico *risk based*, basato cioè sulla variabile rischio, secondo una scala graduale di obblighi e limiti di applicazione degli strumenti: dal divieto di utilizzo all'utilizzo condizionato dall'esistenza di un alto rischio, fino al rischio moderato e a quello minimo e accettabile. Laddove il rischio sia alto, debbono essere in ogni caso assicurate la trasparenza del funzionamento (art. 13), la supervisione e la sorveglianza dell'uomo (art. 14), l'accuratezza, la robustezza e la cibersicurezza (art. 15), la valutazione d'impatto sui diritti fondamentali (art. 27), il diritto alla spiegazione dei processi decisionali (art. 86).

Al regolamento ha fatto seguito il 5 settembre 2024 a Vilnius la Convenzione quadro del Consiglio di Europa su «*AI and Human Rights, Democracy and the Rule of Law*» (*CETS* 225), aperta alla

firma e alla ratifica sia degli Stati membri che di altri Stati non europei e della stessa Unione Europea. Al fine di mitigare i rischi di un impatto negativo dei sistemi di IA, il primo accordo internazionale in materia conferma la dimensione antropocentrica del regolamento per i profili inerenti al rispetto della dignità umana e dell'autonomia individuale, alla trasparenza e alla sorveglianza, alle esigenze di *accountability* e responsabilità, ai criteri di uguaglianza e non discriminazione, alla protezione della *privacy* e dei *personal data*, alla affidabilità, qualità, robustezza e sicurezza, all'approccio basato sul rischio e alla previsione di misure rimediali contro l'abuso.

La Commissione europea, preso atto dello scenario geopolitico della competitività tecnologica e della obiettiva complessità dell'intero quadro normativo digitale dell'UE (fra cui l'*AI Act*), ha peraltro presentato il 19 novembre 2025 una proposta di regolamento, mirata a una significativa semplificazione dell'apparato regolatorio (*Digital Omnibus* e *AI Office*), così da assicurarne una più standardizzata e centralizzata *governance* e una più flessibile *compliance*. E, nel segno di una vocazione costituzionale per un efficiente bilanciamento fra innovazione tecnologica e presidio dei diritti fondamentali, il 20 novembre ha presentato il *Digital Justice Package 2030* sulla strategia per la transizione digitale dei sistemi giudiziari europei, diretto a rafforzare l'impiego dell'IA nelle corti e la specifica formazione di magistrati e personale.

4. Le attività di contrasto della criminalità fra pratiche di IA vietate o ad alto rischio: limitazioni e deroghe

Nel Capo II, art. 5, del regolamento sono elencati gli specifici divieti di uso della IA, accompagnati, peraltro, da talune deroghe ed eccezioni.

Sono vietate (art. 5, par. 1, lett. *a-h*) le pratiche di un sistema di IA:

- che utilizza tecniche subliminali o manipolative o ingannevoli;
- che è mirato alla valutazione o classificazione delle persone sulla base del loro comportamento o di caratteristiche personali, inferendone un punteggio sociale pregiudizievole (*social scoring*);
- che effettua valutazioni o previsioni del rischio che una persona commetta un reato unicamente sulla base della profilazione della stessa (*profiling*), a meno che non sia utilizzato a sostegno della valutazione del suo coinvolgimento in un'attività criminosa, basata «su fatti oggettivi e verificabili direttamente connessi a un'attività criminosa»;

- che crea o espande database di riconoscimento facciale mediante *scraping* di immagini facciali ricavate da internet
- che deduce le emozioni di una persona nel luogo di lavoro, salvo che per motivi medici o di sicurezza;
- che utilizza sistemi di categorizzazione biometrica che classificano le persone in merito a razza, convinzioni politiche, sindacali, religiose, filosofiche, sessuali, a meno che i dati biometrici non siano acquisiti legalmente nelle «attività di contrasto».

È in particolare vietato (lett. *h*) l'uso di «*sistemi di identificazione biometrica remota in tempo reale in spazi accessibili al pubblico*», finalizzati alla identificazione della persona mediante il confronto dei dati biometrici con quelli contenuti in una banca dati. A meno che e nella misura in cui tale uso non sia strettamente necessario per la ricerca mirata di specifiche vittime di sottrazione, tratta o sfruttamento di esseri umani o di persone scomparse, per la prevenzione di una minaccia specifica, sostanziale e imminente per la vita o l'incolumità fisica delle persone o di una minaccia reale e attuale o prevedibile di un attacco terroristico, per la localizzazione o identificazione di una persona sospettata di avere commesso un reato ai fini dell'indagine o dell'esercizio dell'azione penale o dell'esecuzione di una sanzione penale per i gravi crimini elencati nell'Allegato II, punibili con una pena privativa della libertà della durata massima di almeno quattro anni.

Quanto alle «*attività di contrasto*» (par. 5, comma 2), svolte dalle «*autorità di contrasto*» – competenti in materia di prevenzione, indagine, accertamento o perseguimento di reati o esecuzione di sanzioni penali, incluse la salvaguardia e la prevenzione contro le minacce alla sicurezza pubblica -, l'uso è consentito solo per confermare «*l'identità*» della persona interessata, dovendosi peraltro tenere conto della gravità, probabilità e entità del danno che sarebbe causato in caso di mancato uso del sistema e delle conseguenze dell'uso del sistema per i diritti e le libertà delle persone interessate. Vanno, inoltre, rispettate le condizioni «necessarie e proporzionate» in relazione all'uso, che è autorizzato solo se l'autorità di contrasto ha completato una «valutazione d'impatto sui diritti fondamentali» e ha registrato il sistema nella banca dati UE; «in situazioni d'urgenza debitamente giustificate», è possibile usare tali sistemi senza la registrazione, che dovrà essere completata senza indebito ritardo.

L'uso di un siffatto sistema a fini di contrasto (par. 3) «è subordinato a un'autorizzazione preventiva rilasciata da un'autorità giudiziaria o da un'autorità amministrativa indipendente», la cui decisione è «vincolante» ed è rilasciata su richiesta motivata e in conformità alle regole dettagliate del diritto nazionale; l'autorità competente rilascia l'autorizzazione solo se ha accertato, sulla base di prove oggettive o indicazioni chiare, che l'uso è «necessario e

proporzionato» al conseguimento degli obiettivi consentiti, restando «limitato a quanto strettamente necessario»; in una situazione d'urgenza debitamente giustificata, è possibile usare il sistema senza autorizzazione che va tuttavia richiesta al più tardi entro 24 ore, ma, se l'autorizzazione non è concessa, l'uso è interrotto e gli *output* sono eliminati.

Spetta allo Stato membro autorizzare l'uso di sistemi di identificazione biometrica remota in tempo reale in spazi accessibili al pubblico a fini di attività di contrasto, entro i limiti e alle condizioni di cui ai par. 1 lett. *h*), 2 e 3, stabilendo all'uopo «*regole dettagliate*» in conformità del diritto dell'Unione (par. 5).

Quanto ai sistemi di IA ad alto rischio, sono considerati tali dall'art. 6, par. 2 del Capo III del Regolamento, quelli elencati nell'Allegato III, fra i quali rilevano, per i sistemi giudiziari, i settori intitolati: *6. Attività di contrasto* e *8. Amministrazione della giustizia e processi democratici*.

Per le «*attività di contrasto*» di cui al punto 6, sono considerati i sistemi di IA destinati ad essere utilizzati dalle autorità di contrasto: a) per determinare il rischio per una persona di diventare vittima di reati; b) come poligrafi e strumenti analoghi; c) per valutare l'affidabilità degli elementi probatori nel corso delle indagini o del perseguimento di reati; d) per determinare il rischio di commissione del reato o di recidiva in relazione a una persona, non solo sulla base della profilazione o per valutare i tratti della personalità o il comportamento criminale pregresso; e) per effettuare la profilazione delle persone nel corso delle indagini, dell'accertamento e del perseguimento di reati.

Per l'«*amministrazione della giustizia*» di cui al punto 8, sono considerati i sistemi di IA destinati ad essere usati da un'autorità giudiziaria per assistenza nella ricerca e nell'interpretazione dei fatti e del diritto e nell'applicazione della legge a una serie concreta di fatti o nella risoluzione alternativa delle controversie, con l'avvertenza che «*The use of AI tools can support the decision-making power of judges or judicial indipendence, but should not replace it: the final decision-making must remain a human-driven activity*».

Ai sensi dell'art. 6, par. 3, un sistema di IA di cui all'allegato III non è considerato ad alto rischio se non presenta un rischio significativo di danno per i diritti fondamentali delle persone, anche nel senso di non influenzare materialmente il risultato del processo decisionale. Il che si verifica quando il sistema di IA è destinato: a) ad eseguire un compito procedurale limitato; b) a migliorare il risultato di un'attività umana completata; c) a rilevare schemi decisionali o deviazioni da schemi decisionali precedenti e non è finalizzato a sostituire o influenzare la valutazione umana precedentemente completata senza un'adeguata revisione umana; d) a

eseguire un compito preparatorio per una valutazione pertinente ai fini dei casi d'uso di cui all'allegato III. Un sistema di IA di cui all'allegato III è sempre considerato ad alto rischio qualora effettui la profilazione di persone.

I sistemi di gestione ad alto rischio debbono rispettare una serie di condizioni di esercizio e requisiti, sia di tipo tecnico o amministrativo, per gli aspetti dell'accuratezza, robustezza, cibersicurezza e responsabilità, sia attinenti, in materia di attività di contrasto, alla trasparenza e alla fornitura di informazioni ai *deployer* (art. 13), alla sorveglianza umana (art. 14) e alla valutazione d'impatto sui diritti fondamentali (art. 27).

Inoltre, qualsiasi persona che abbia motivo di ritenere che vi sia stata una violazione delle disposizioni del regolamento può presentare – almeno – un reclamo alla competente autorità di vigilanza del mercato (art. 85) e qualsiasi persona oggetto di una decisione adottata dal *deployer* sulla base dell'*output* di un sistema di IA ad alto rischio elencato nell'allegato III, ad eccezione di quelli al punto 2, e che incida significativamente su tale persona per avere un impatto negativo sui diritti fondamentali, ha il diritto di ottenere dal *deployer* spiegazioni chiare sul ruolo del sistema nella procedura decisionale e sui principali elementi della decisione adottata (art. 86).

5. Il sistema processuale penale e il modello regolatorio dell'*AI Act*

La regolamentazione sovranazionale, con la previsione di clausole aperte, deroghe e limitazioni per le pratiche vietate e per quelle ad alto rischio e di specifiche prescrizioni in materia di «*attività di contrasto*» da parte delle «*autorità di contrasto*», pretende il necessario adeguamento dell'ordinamento processuale interno, che va puntualmente e urgentemente armonizzato.

Con particolare riguardo all'ambiente della giustizia penale, i valori costituzionali e convenzionali del giusto processo (l'indipendenza dei magistrati, la terzietà e imparzialità del giudice, la presunzione di innocenza, la parità delle armi, il contraddittorio, la motivazione della decisione nel rispetto del criterio dell' «al di là di ogni ragionevole dubbio», il controllo impugnatorio) e di tutela dei diritti fondamentali della persona pretendono la fissazione di principi e regole chiare, per un verso, circa la dislocazione e l'equilibrio dei poteri fra polizia giudiziaria, pubblico ministero e giudice nella fase delle indagini preliminari, nonché, per altro verso, circa la validità e utilizzabilità probatoria degli atti investigativi supportati da sistemi di intelligenza artificiale, l'ammissibilità e la valutazione della prova digitale nel giudizio,

l'effettività della garanzia di partecipazione dialettica di tutte le parti prima alla perimetrazione del ragionamento probatorio e poi al sindacato impugnatorio della decisione.

Il legislatore nazionale è chiamato innanzitutto a dislocare, con «*regole dettagliate*» e mediante scelte anche di tipo valoriale, i poteri e le funzioni di accertamento spettanti alla competente autorità nella fase delle indagini preliminari, assurta – com'è noto – a indiscusso baricentro anche mediatico del processo, riportando in equilibrio i rapporti fra la polizia giudiziaria, il pubblico ministero e il giudice, mediante la disciplina di appositi interventi ordinatori (le cd. « *finestre di giurisdizione*») del giudice per le indagini o dell'udienza preliminare.

Invero, laddove l'attività di contrasto si debba concretizzare nell'utilizzo di strumenti di intelligenza artificiale vietati o ad alto rischio, potenzialmente pervasivi dei diritti fondamentali e delle libertà individuali, la previsione di specifiche deroghe e prescrizioni rinvia ad operazioni logiche di apprezzamento di fatti e circostanze e di bilanciamento di valori più propriamente riservate alla figura del giudice, anziché a quella del pubblico ministero, il quale resta "*parte*" , benché pubblica. Si pensi alla definizione dei requisiti di stretta necessità e proporzione nell'uso del sistema di IA, della valutazione d'impatto sui diritti fondamentali e sulle libertà delle persone, della procedura autorizzatoria, dell'apprezzamento dell'urgenza, della probabilità e entità del danno da mancato uso, del criterio di sussidiarietà suggerito per l'uso dei dispositivi di IA generativa ecc. L'ingresso nel processo di una siffatta prova algoritmica sembra dunque giustificare, ai fini della affidabile ricostruzione dei fatti e della utile prospettazione dell'accusa nel giudizio, l'articolazione di un anticipato e flessibile contraddittorio, almeno cartolare, fra le parti dinanzi a un giudice, in merito alla legalità di atti investigativi supportati da sistemi di intelligenza artificiale.

Parimenti, con riguardo alla fase del giudizio, in funzione sia dell'ammissione che della valutazione della prova algoritmica, sembrano insufficienti i tradizionali criteri enunciati, per la *scientific evidence*, dalla Corte Suprema statunitense nella nota sentenza *Daubert*[2].

Neppure risulta utile rinviare il controllo di attendibilità di questa speciale prova al contraddittorio fra le parti «*sulla*» prova, cioè quando essa sia stata già ammessa e acquisita, mentre sarebbe più efficace prevedere un filtro preliminare di ammissibilità, accompagnato da un agile contraddittorio «*per*» la prova, così da evitare che entrino nel dibattito informazioni non sorrette da una previa validazione di attendibilità o utilizzabilità[3]. Un filtro, dunque, a maglie ben più strette rispetto a quello previsto dall'art. 190, comma 1, c.p.p. che, ai fini dell'ammissione della prova in genere, si limita a selezionare negativamente solo «*le prove vietate dalla legge e*

quelle che manifestamente sono superflue o irrilevanti», e inoltre assistito da un significativo rafforzamento del contraddittorio anticipato «*per*» la prova.

Quanto alla valutazione della prova digitale e agli schemi del ragionamento probatorio del giudice, che costituisce la piattaforma degli eventuali rimedi impugnatori, appare senz'altro dirimente l'insegnamento della Corte Suprema del Wisconsin, che, con la sentenza pronunciata nel *leading case* «Loomis» (Wisconsin S.C., State v. Loomis, 881, Wis. 2016), già avvertiva, in tema di giustizia predittiva, che il software COMPAS, comunemente utilizzato per misurare il rischio di recidivanza, «*… should be always constitute merely one tool available to a Court, that need to be confirmed by additional sound information …*». Sembra chiaro il riferimento al criterio della non esclusività del risultato algoritmico, che va accuratamente riscontrato – corroborato – dagli ulteriori e diversi elementi di prova acquisiti nel procedimento, al fine di sterilizzare gli effetti perversi del cd. *confirmation bias*. Il giudice dovrebbe pertanto integrare le stime probabilistiche della prova algoritmica con le altre informazioni probatorie funzionali all'accertamento dei fatti e delle circostanze, nel rispetto del nucleo essenziale dei valori del giusto processo.

6. La legge n. 132 del 2025 e la transizione digitale

Il regolamento (UE) 2024/1689 avverte in premessa gli Stati membri circa l'esigenza di procedere a una «*governance systems at Union and National level*». Nei Considerando (8) e (26) ammonisce che è necessaria l'adozione di un quadro giuridico dell'Unione, proporzionato ed efficace, che istituisca «*regole armonizzate*» e «*vincolanti*», avvalendosi di un approccio basato sul rischio che può essere generato dal sistema di IA, definito in modo chiaro, e garantendo allo stesso tempo un elevato livello di protezione degli interessi riconosciuti e tutelati dal diritto dell'Unione.

I principi e le regole dettati dalla fonte sovranazionale per la *governance* delle innovazioni tecnologiche nei sistemi giudiziari hanno dunque bisogno di essere concretamente attuati, secondo una efficace e credibile strategia nazionale, per cui occorre prestare speciale attenzione sia alle politiche legislative di armonizzazione del diritto interno che alle relative prassi interpretative e applicative delle corti.

La legge n. 132 del 23 settembre 2025, recante *Disposizioni e deleghe al Governo in materia di intelligenza artificiale*, premesso che «Le disposizioni della presente legge si interpretano e si applicano conformemente al regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, del 13 giugno 2024» (art. 1, comma 2) e ribadita nei principi generali la dimensione

antropocentrica dell'uso della strumentazione tecnologica, nel rispetto dei diritti fondamentali e delle libertà costituzionali (art. 3), contiene, tra l'altro, anche alcune norme dirette a disciplinare l'uso dell'IA nel sistema giudiziario, anche per quello criminale.

L'art. 15 riconosce innanzitutto che «è sempre riservata al magistrato ogni decisione sull'interpretazione e sull'applicazione della legge, sulla valutazione dei fatti e delle prove e sull'adozione dei provvedimenti» (comma 1). Restano così salvaguardati il nucleo centrale della funzione giurisdizionale e il principio di responsabilità umana ultima, mentre l'impiego dell'IA è limitato alle aree di supporto e assistenza del decisore. La stessa norma centralizza poi nel Ministero della giustizia sia la disciplina degli impieghi dei sistemi di intelligenza artificiale «per l'organizzazione dei servizi relativi alla giustizia, per la semplificazione del lavoro giudiziario e per le attività amministrative accessorie» (comma 2), sia il potere autorizzatorio per «la sperimentazione e l'impiego dei sistemi di intelligenza artificiale negli uffici giudiziari ordinari, sentite le Autorità nazionali di cui all'articolo 20 [AgID e ACN]» (comma 3), sia infine l'elaborazione di linee programmatiche e la promozione di specifiche attività didattiche nella «*formazione digitale di base e avanzata*» dei magistrati e del personale amministrativo (comma 4).

La disposizione di delega dell'art. 24, nel rinviare a «uno o più decreti legislativi per l'adeguamento della normativa nazionale al regolamento (UE) 2024/1689» (art. 1, comma 1), secondo i principi e criteri direttivi del comma 2, avverte che «il Governo è altresì delegato ad adottare entro dodici mesi dalla data di entrata in vigore della presente legge uno o più decreti legislativi per adeguare e specificare la disciplina dei casi di realizzazione e di impiego illeciti di sistemi di intelligenza artificiale» (comma 3).

E però, i principi e criteri direttivi – fra i quali quelli riguardanti la «previsione di un'apposita disciplina per l'utilizzo di sistemi di intelligenza artificiale per l'attività di polizia» (comma 2, lett. h), la «regolazione dell'utilizzo dei sistemi di intelligenza artificiale nelle indagini preliminari, nel rispetto delle garanzie inerenti al diritto di difesa e ai dati personali dei terzi, nonché dei principi di proporzionalità, non discriminazione e trasparenza» (comma 5, lett. e), la «modifica, a fini di coordinamento e di razionalizzazione del sistema, della normativa sostanziale e processuale vigente, in conformità ai principi e ai criteri enunciati nelle lettere a), b), c) d) ed e)» (comma 5, lett. f) – sembrano invero ancora troppo generici rispetto all'obiettivo di effettiva armonizzazione del sistema processuale penale interno.

Non si rinvengono, infatti, prescrizioni puntuali per i decreti attuativi della delega in merito ai poteri, ai divieti, alle condizioni, ai limiti e alle procedure dirette ad acquisire, utilizzare e

valutare le informazioni generate dai sistemi di intelligenza artificiale, anche generativa, sia nelle indagini che nel giudizio.

Inoltre, vanno rimarcate le significative criticità riguardanti l'indifferenza per la sostenibilità economica e finanziaria delle pur necessarie innovazioni tecnologiche, confermata dalla consueta clausola di invarianza finanziaria (art. 27), e la eccessiva frammentazione delle competenze fra il Ministero della giustizia e le Agenzie governative AgiD e ACN. E, soprattutto, va sottolineato il clamoroso, mancato coinvolgimento del Consiglio Superiore della Magistratura nel complessivo processo di *governance* del fenomeno di digitalizzazione della giurisdizione, che appare disegnato secondo uno schema di tipo ministeriale, nonostante l'evidente interferenza dello stesso nell'esercizio concreto e responsabile dell'attività giudiziaria, con il conseguente impatto sull'indipendenza dei magistrati**[4]**.

L'inedita sfida, culturale e costituzionale, per la tenuta dei valori del giusto processo e per la protezione dei diritti fondamentali della persona, esige prontezza e coerenza del processo legislativo nazionale di armonizzazione, che sia mirato a legittimare il nuovo modello epistemico **[5]**, sulla base di un serio dialogo multidisciplinare e una equilibrata interazione fra i saperi e le operazioni logiche tradizionalmente affidate al giudice e alle parti e le evidenze digitali della prova algoritmica. Tutto ciò restando nel solco dei valori di trasparenza, controllo umano significativo, autonomia e responsabilità del decisore, fallibile come ogni essere umano ma controllabile secondo legge e ragione.

Il disciplinamento, etico e giuridico, del fenomeno ad opera del regolamento e della Convenzione quadro sembra consentire di reggere l'urto dei pervasivi e nuovi strumenti tecnologici, senza che risultino alterate – almeno finora – l'equità e l'efficacia del sistema di giustizia al servizio delle persone e le garanzie del tradizionale paradigma razionalistico della giurisdizione penale, in favore, viceversa, dell'affermarsi di uno spregiudicato potere economico-finanziario e geopolitico.

Gli esiti della sfida, tuttavia, non sono affatto scontati.

[1] Cfr., volendo, G. CANZIO, *Intelligenza artificiale e processo penale*, in AA.VV., *Prova scientifica e diritto penale*, a cura di G. Canzio e L. Luparia Donati, Walter Kluwer – CEDAM, Milano, III ed., 2025, p. 999 ss.

[2] Secondo Daubert v. Merrel Dow Pharmaceuticals, Inc., 509 US 579 (1993), il giudice deve vagliare l'affidabilità di una teoria o un metodo e di una *expert witness's scientific testimony*, ai fini della loro ammissibilità come prova scientifica nel processo, per i seguenti profili: la controllabilità mediante esperimenti; la falsificabilità mediante test di smentita con esito negativo; la *peer review* della comunità scientifica di riferimento; la conoscenza della percentuale di errore dei risultati; infine, il criterio ausiliario della generale accettazione da parte della comunità degli esperti. Criteri, questi, sostanzialmente condivisi e anzi arricchiti dalla Corte di cassazione italiana (Sez. IV, 17/09/2010, n. 43786, Cozzini; Sez. Un., 11/09/2002, n. 30328, Franzese).

[3] Nel sistema processuale vigente si rinviene la disposizione dell'art. 189 c.p.p. sulla cd. prova atipica, per cui l'apprezzamento di rilevanza, non superfluità e concreta idoneità della prova «ad assicurare l'accertamento dei fatti» – senza che ne resti pregiudicata «la libertà morale delle persone» per il divieto di perizia criminologica ex art. 220, comma – è rimesso al vaglio preliminare del giudice. Questi, dopo avere sentito le parti sulle modalità di assunzione della prova, provvede all'ammissione con ordinanza, fissando le regole per la corretta applicazione dei metodi e delle procedure tecniche di acquisizione della stessa. Nella Relazione al Progetto preliminare del nuovo codice di procedura penale del 1989 (p. 60), riguardante la portata dell'art. 189, si afferma: «È sembrato che una norma così articolata possa evitare eccessive restrizioni ai fini dell'accertamento della verità, tenuto conto del continuo sviluppo tecnologico che estende le frontiere dell'investigazione, senza mettere in pericolo le garanzie difensive».

[4] Sul tema, cfr. il contributo di F. DAL CANTO e S. ROVELLI, *Il modello costituzionale di ordinamento giudiziario e la sfida della digitalizzazione della giurisdizione: il ruolo del Ministero della Giustizia e del Consiglio Superiore della Magistratura prima e dopo la legge italiana sull'intelligenza artificiale*, in *Quest. Giust.*, 9/12/2025.

[5] Cons. il recente intervento di Papa Leone XIV a favore di "*un nuovo umanesimo nell'era digitale*", in *Corriere della sera*, 8 novembre 2025. Fra i tanti contributi dottrinali, cfr. N. LIPARI, *Il diritto del nuovo millennio tra giurisdizionalizzazione ed algoritmo*, in *Accademia*, 2024, p. 9 ss. Merita sicuro apprezzamento la «*Carta dei valori*» elaborata dall'Unione delle Camere Penali Italiane e presentata il 17 gennaio 2025.