ATTUALITÀ

PIETRO MIRTO

Giustizia predittiva e processo penale: l'algoritmo alla prova del «giusto processo»

Il contributo analizza l'impatto dirompente dell'intelligenza artificiale sul processo penale, superando la logica della mera rassegna normativa per approdare a una proposta ricostruttiva fondata sul 'nuovo garantismo tecnologico'. Attraverso l'esame critico dell'AI Act e della L. 132/2025, l'Autore delinea un modello procedurale di 'discovery coatta' e introduce la figura del 'Perito Algoritmico' come garante della falsificabilità della prova. L'obiettivo è riaffermare la centralità del libero convincimento del giudice inteso come obbligo di dissenso motivato dal calcolo probabilistico, preservando la riserva di umanità contro la deriva verso una legalità puramente statistica.

Predictive Justice and Criminal Procedure: The Algorithm on Trial for a «Fair Trial»

The article analyzes the disruptive impact of artificial intelligence on criminal proceedings, moving beyond a mere regulatory overview to propose a reconstructive model based on "new technological due process" Through a critical examination of the AI Act and Law No. 132/2025, the author o*utlines a procedural model of "mandatory discovery" and introduces the figure of the "Algorithmic Expert" as a guarantor of the falsifiability of evidence. The objective is to reaffirm the centrality of the judge's "free evaluation of evidence," interpreted as a duty to provide a reasoned dissent from probabilistic calculations, thereby preserving the "human reserve" against a shift toward purely statistical legality.*

1. *Premessa. L'ineludibile irruzione dell'algoritmo nel perimetro penalistico*

L'avanzamento tecnologico e l'emergere dei cosiddetti sistemi di intelligenza artificiale[1] hanno catalizzato un fenomeno di portata rivoluzionaria che, pur prospettando massimizzazione dell'efficienza e ottimizzazione funzionale, solleva con crescente urgenza la necessità di esaminare l'impatto che un impiego non adeguatamente regolamentato di tali tecnologie può avere sui diritti fondamentali degli individui coinvolti. In ambito penalistico e processuale pena-

[1] Il termine venne coniato nel 1955 dal matematico MCCARTHY; per una definizione normativa attuale, si veda la Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari adottata dalla CEPEJ nel 2018.

listico, ampia dottrina[2] ha da tempo segnalato le possibili distorsioni derivanti dall'impiego di sistemi automatizzati nel procedimento penale. Ci si interroga, in estrema sintesi, in che misura un sistema standardizzato possa realmente garantire neutralità e imparzialità[3], o se l'integrazione di strumenti di IA nei processi decisionali non contribuisca piuttosto a cristallizzare e amplificare pratiche discriminatorie già radicate nel sistema.

2. Algoritmi predittivi e risk assessment tools*: tra potenzialità tecniche ed esigenze di garanzia.* L'indagine sulla natura degli algoritmi predittivi impone, in via preliminare, una chiarificazione terminologica e strutturale circa il c.d. "tessuto costitutivo" dell'intelligenza artificiale[4]. Tali sistemi si fondano su reti neurali artificiali che, ispirandosi ai modelli biologici, si articolano su una pluralità di strati (*layers*): un *input layer*, deputato alla ricezione dei dati grezzi; una serie di *hidden layers*, ove avviene l'elaborazione e la trasformazione delle informazioni; e, infine, un *output layer* che fornisce la risposta definitiva[5]. Il passaggio cruciale risiede nel *machine learning* (apprendimento automatico), processo mediante il quale la macchina non si limita ad eseguire istruzioni statiche, ma "impara" dai dati di addestramento, costruendo modelli matematici capaci di reperire connessioni non note a priori ai progettisti stessi. Nel perimetro della giustizia penale, tali strumenti assumono la denominazione di *risk assessment tools*[6]. La loro funzione non è quella di accertare un fatto nel senso storico-fenomenico del termine, bensì quella di formulare una previsione probabilistica su un comportamento futuro del soggetto (tipicamente la recidiva). Si assiste, dunque, ad una frizione ontologica: mentre il processo penale tradizionale è orientato all'indietro (*backward-looking*), l'algoritmo predittivo è proiettato in avanti (*forward-looking*), sostituendo la certezza processuale del fatto commesso con una "pericolosità statistica" fondata su indici di de-

---

[2] *Si vedano, tra gli altri,* CANZIO, *Intelligenza artificiale, algoritmi e giustizia penale, in www.sistemapenale.it, 2021*; BACCARI-FELICIONI, *La decisione penale tra intelligenza emotiva ed intelligenza artificiale, Padova, 2023.*

[3] *Sul punto v.* AIROLDI-GAMBETTA, *Sul mito della neutralità algoritmica, in The Lab's Quarterly, 2018.*

[4] ALGERI, *Intelligenza artificiale e polizia predittiva*, in Dir. pen. proc., 2021, 6, 726;
BACARI- FELICIONI *La decisione penale tra intelligenza emotiva ed intelligenza artificiale, cit., 72.*

[5] *Commissione europea per l'efficienza della giustizia, carta etica per l'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi, 3 dicembre 2018, 48*

[6] SIGNORATO, *Giustizia penale e intelligenza artificiale. Considerazioni in tema di algoritmo predittivo, in Riv. dir. proc. 2020, 607.*

vianza tipizzati[7]. Tuttavia, l'impiego di siffatte tecnologie solleva il complesso problema della *black box* (scatola nera): l'opacità dei processi decisionali degli algoritmi di *deep learning* rende spesso impossibile ripercorrere l'*iter* logico che ha condotto all'*output*. Tale deficit conoscitivo si riverbera inevitabilmente sull'obbligo di motivazione dei provvedimenti giurisdizionali, rischiando di ridurre il magistrato ad un mero "acritico ratificatore" di risultati tecnologici di cui non domina le premesse.[8]

3. Il precedente nordamericano: dal caso Loomis al dibattito sulle sentencing disparities. Il segreto commerciale come *vulnus* al diritto di difesa. La parabola evolutiva della giustizia predittiva trova nel contesto statunitense, e segnatamente nella nota vicenda Loomis v. Wisconsin[9], il proprio "punto di crisi" costituzionale. L'utilizzo del sistema COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*) per la determinazione del *quantum* di pena ha palesato una frizione insanabile con le garanzie del *due process*: la difesa, infatti, si è trovata nell'impossibilità di contestare i criteri di calcolo del punteggio di rischio, in quanto protetti da segreto commerciale e proprietà intellettuale della società produttrice.[10] Si configura, in tal guisa, una pericolosa "privatizzazione della funzione giudicante", ove il nucleo essenziale del decidere viene delegato a un'entità privata il cui codice sorgente rimane inaccessibile alle parti e allo stesso magistrato.[11] Tale opacità trasforma l'algoritmo in un "testimone di silicio" sottratto al vaglio del controesame: se la prova scientifica, nel rito accusatorio, deve essere sottoposta alla verifica dialettica, l'impiego di una *black box* protetta da *trade secret* impedisce *ab imis* l'eserci-

---

[7] *Per una disamina dei sistemi place-based (aree a rischio) e person-based (soggetti a rischio), si veda* FERGUSON, *The Rise of Big Data Policing: Surveillance, Race and the Future of Law Enforcement, New York, 2017.*

[8] *Sul punto* CEVOLANI-CRUPI, *Come ragionano i giudici: razionalità, euristiche e illusioni cognitive, cit., 23;* BLAIOTTA, *Giustizia, errore, intelligenza artificiale, in www.sistemapenale.it, 23 ottobre 2023, 5.*

[9] *Supreme Court of Wisconsin, State v. Loomis, 881 N.W.2d 749 (Wis. 2016). Per un commento critico in lingua italiana,* MANES, *L'oracolo algoritmico e la giustizia penale: al capezzale del «giusto processo», in Riforma della giustizia, 2020.*

[10] *Sul concetto di Trade Secret applicato agli algoritmi di risk assessment,* WEXLER, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System, in Stanford Law Review, vol. 70, 2018, 1343 ss.*

[11] QUATROCCOLO, *Diritto a un processo equo e sistemi di intelligenza artificiale, in Arch. pen., 2019, n. 3, 11 ss.*

zio del diritto di cui all'art. 111, co. 4 Cost.[12] In assenza di una *full disclosure* (ostensione integrale) dei criteri di ponderazione dei dati, il controllo di logicità della motivazione degrada a mera clausola di stile. Il rischio, già paventato dalla dottrina d'oltreoceano e confermato da autorevoli studi indipendenti, è che l'algoritmo incorpori e occulti bias discriminatori su base etnica o sociale, rendendo la decisione penale non solo incomprensibile, ma intrinsecamente ingiusta.[13] Il raffronto con l'esperienza statunitense non deve indurre a una trasposizione acritica di categorie aliene al nostro sistema. Se nel modello di *common law* il *due process* può in parte flettere dinanzi alle esigenze di efficienza del *sentencing*, nel perimetro costituzionale italiano la funzione giurisdizionale è inscindibile dal dovere di motivazione piena e razionale.[14] La lezione del caso Loomis insegna che l'opacità algoritmica non è un limite tecnico, ma un limite di legittimità: una decisione 'assistita' da una scatola nera è, per definizione, una decisione il cui fondamento logico è parzialmente sottratto al controllo delle parti e del pubblico, ponendosi in urto insanabile con la pubblicità del dibattimento e l'oralità della prova[15].

4. Intelligenza artificiale e controllo del territorio: profili di polizia predittiva tra prevenzione e sospetto algoritmico. L'impiego di sistemi algoritmici nella fase delle indagini preliminari e nell'attività di prevenzione dei reati delinea una metamorfosi genetica della funzione di polizia.[16] Si transita dal modello investigativo tradizionale, ancorato alla notizia di reato (*post delictum*), a una sorveglianza proattiva proiettata verso l'anticipazione della soglia di punibilità (*ante delictum*). La distinzione tecnica tra sistemi *place-based* (come l'italiano

[12] *Sull'impossibilità di un cross-examination dell'algoritmo,* GARRETT-MONAHAN, *Judging Risk: Expert Evidence and Predictive Algorithms, in 13 Ohio State Journal of Criminal Law, 2018.*

[13] *Il riferimento è alla celebre inchiesta di Propublica, Machine Bias, 2016, che dimostrò come Compas generasse tassi di "falsi positivi" per i detenuti afroamericani quasi doppi rispetto ai caucasici.*

[14] *Sul punto, sia consentito il rinvio alla fondamentale distinzione operata da* TARUFFO, *La semplice verità. Il giudice e la costruzione dei fatti, Laterza, Bari, 2009, il quale chiarisce come la verità processuale non possa mai essere il frutto di un «salto logico» non verificabile.*

[15] DI GIOVINE, *Diritto penale e neuroscienze, in Rivista Italiana di Diritto e Procedura Penale, 2022, p. 45 ss.,* sulla necessità di preservare l'approccio umanistico del diritto rispetto alle derive riduzionistiche delle scienze dure e delle tecnologie predittive.

[16] Sul punto, la riflessione penalistica deve farsi carico di arginare la deriva verso un «diritto penale d'autore» di stampo tecnologico. DONINI, *Il volto attuale dell'illecito penale, Milano, 2024, 88 ss.,* il quale ammonisce circa il rischio che l'efficienza investigativa diventi l'unico parametro di legittimazione dell'agire pubblico, a scapito del principio di offensività.

KeyCrime, focalizzato sulle serie criminali) e sistemi *person-based* (quali lo statunitense PredPol, orientato alla profilazione individuale)[17] non esaurisce le criticità sottese all'automazione del sospetto. La sostituzione del «sospetto ragionevole» presupposto di legittimità dell'agire di P.G. con una correlazione statistica inaccessibile mina alla base la presunzione di non colpevolezza (art. 27, co. 2, Cost.).[18] Il cittadino, inserito in un *cluster* di rischio, decade a «sospetto perpetuo»,[19] subendo una compressione della libertà personale fondata su un calcolo probabilistico che la difesa non può sindacare nel merito. Tale opacità determina un *vulnus* insanabile nel controllo di legittimità sulle misure limitative della libertà, poiché l'elemento "oggettivo" della perquisizione o del fermo svanisce nell'imperscrutabilità del codice sorgente. L'impiego di tali applicativi solleva interrogativi che trascendono la mera efficienza operativa, investendo il cuore del garantismo penale. La profilazione algoritmica, agendo su base probabilistica, rischia di cristallizzare un 'sospetto di gruppo' che precede la commissione del fatto, orientando la pressione investigativa verso aree o soggetti già marginalizzati.[20] Si assiste a una torsione del diritto penale del fatto verso un diritto penale della prevenzione statistica, dove la neutralità

[17] In ordine alla distinzione tra i *software de quo,* ALGERI, *Diritto penale e intelligenza artificiale, cit., 102.* Occorre precisare che mentre il modello person-based è stato travolto da critiche per l'incorporazione di bias etnici, l'approccio *place-based* di matrice europea tenta una via più garantista, ancorando la predizione a dati oggettivi sui reati e non sulle caratteristiche dei soggetti. Tuttavia, la persistenza di *feedback loops* rimane un'insidia immanente a ogni processo di *machine learning* applicato alla sicurezza

[18] La giurisprudenza sovranazionale ha recentemente tracciato confini netti: si veda Corte di giust. UE (Grande sezione), 30 gennaio 2024, C-118/22, in tema di conservazione dei dati biometrici e profilazione, che impone limiti rigorosi alla «schedatura» algoritmica basata sulla mera utilità investigativa. In ambito EDU, il caso Glukhin v. Russia (2023) conferma che l'uso sproporzionato di tecnologie di riconoscimento e profilazione viola l'art. 8 C.E.D.U., trasformando la sorveglianza in un'ingerenza illegittima nella vita privata.

[19] Nonostante l'assenza di un pronunciamento specifico della Consulta sui *software* predittivi, la giurisprudenza costituzionale in tema di misure di prevenzione (Corte cost., sentt. n. 2/1956 e n. 127/2022) offre solidi argini: l'autorità di P.S. non può mai agire in base a congetture puramente soggettive o statistiche, dovendo l'indizio di pericolosità poggiare su condotte esteriori oggettivamente verificabili. L'automazione del sospetto rischia di bypassare tale vaglio, surrogando la valutazione del fatto con la validazione dell'algoritmo.

[20] UBERTIS, *Il sistema della prova penale*, Giuffrè, Milano, 2022, p. 112. L'Autore evidenzia come la formazione del convincimento non debba essere contaminata da pregiudizi probabilistici maturati in fase investigativa, che minano l'imparzialità del giudizio.

della prova rischia di essere inquinata da un automatismo indiziario che condiziona il giudice fin dalle prime fasi del rito cognitivo[21].

5. La regolamentazione della tecnologia tra ambizioni europee e recepimento nazionale: l'AI Act e la L. 23 settembre 2025, n. 132. Il quadro normativo ha subito una sterzata decisiva con l'entrata a pieno regime del Regolamento (UE) 2024/1689 (c.d. AI Act), il quale adotta un approccio basato sul rischio (*risk-based approach*) per classificare le applicazioni di intelligenza artificiale.[22] All'interno di questa architettura, i sistemi utilizzati nell'amministrazione della giustizia e dalle autorità di pubblica sicurezza sono stati inseriti nella categoria dei sistemi ad alto rischio (Allegato III).[23] Tale classificazione non è un mero esercizio tassonomico, ma impone obblighi rigorosi di trasparenza, tracciabilità e, soprattutto, di sorveglianza umana (*human oversight*). Il legislatore italiano, con la L. 23 settembre 2025, n. 132, ha inteso dare attuazione a tali precetti, tentando di coniugare l'innovazione tecnologica con la "riserva di umanità" del giudizio penale.[24] Tuttavia, l'esegesi dell'art. 4 della suddetta legge rivela zone d'ombra preoccupanti: se da un lato si afferma che la decisione finale spetta sempre al magistrato, dall'altro non si chiariscono i meccanismi di "resistenza" del giudice rispetto al suggerimento algoritmico.[25] Il rischio concreto è che la sorveglianza umana si riduca a un adempimento burocratico, una sorta

[21] Sul rischio di «stigmatizzazione algoritmica», PAZE, *Il giudice e l'algoritmo, in Questione Giustizia, 2024.* L'Autrice ammonisce contro l'illusione di una neutralità tecnologica che occulta scelte politiche di controllo sociale.

[22] Per una prima lettura del Regolamento, LUCCHETTI, *Il governo degli algoritmi: commento all'AI Act europeo, Milano, 2024, 45 ss.* L'Autore evidenzia correttamente come l'Europa abbia scelto la via della regolazione antropocentrica, differenziandosi nettamente dal modello *laissez-faire* statunitense e da quello del controllo sociale cinese.

[23] L'inserimento dei sistemi giudiziari tra quelli "ad alto rischio" risponde alla necessità di tutelare i diritti fondamentali di cui alla Carta di Nizza. Cfr. l'ampia motivazione contenuta nel Considerando 40 del Regolamento, ove si specifica che l'IA non deve sostituirsi alla capacità decisionale umana, specialmente quando in gioco vi è la libertà personale.

[24] Sulla genesi della L. n. 132/2025, si veda la Relazione illustrativa al disegno di legge, dove si ribadisce che l'intelligenza artificiale deve fungere esclusivamente da «ausilio cognitivo» e mai da «decisore automatico». In dottrina, RUOTOLO, *L'umanesimo giudiziario nell'era digitale, in Giust. pen., 2025, II, 112.*

[25] È il fenomeno della c.d. *automation bias*: la tendenza psicologica dell'essere umano a fidarsi in modo eccessivo dell'esito prodotto da una macchina, percepito come oggettivo e infallibile. Sul punto, la giurisprudenza della Cassazione ha già avvertito che il ricorso a strumenti tecnologici non esime il giudice dall'obbligo di un'autonoma e critica valutazione del materiale probatorio (Cass. pen., Sez. II, sent. n. 45001/2023).

di *rubber stamping* (approvazione acritica), qualora il giudicante non sia dotato degli strumenti tecnici per comprendere le basi probabilistiche dell'*output* tecnologico. Inoltre, il combinato disposto tra l'art. 14 dell'AI Act e le nuove norme procedurali italiane pone il problema della "spiegabilità" (*explainability*). Un sistema ad alto rischio non deve solo funzionare, deve essere interpretabile. In sede processuale, ciò si traduce nel diritto dell'imputato di conoscere i parametri che hanno determinato la sua profilazione, superando quel muro del segreto commerciale che abbiamo precedentemente analizzato come *vulnus* al diritto di difesa.[26] Nella complessa tassonomia delineata dal Legislatore europeo, un rilievo fondamentale assume la distinzione tra i sistemi di IA consentiti, seppur gravati da pesanti oneri di compliance, e le pratiche integralmente proibite in quanto incompatibili con i valori della dignità umana e dello Stato di diritto.[27] Ai sensi dell'art. 5 del Regolamento (UE) 2024/1689, sono vietati i sistemi di polizia predittiva che effettuano valutazioni del rischio basate esclusivamente sulla profilazione individuale o sulla valutazione di tratti della personalità per prevedere la commissione di un reato, in assenza di un sospetto fondato su fatti oggettivi e verificabili.[28] Questa "clausola di sbarramento" è di fondamentale importanza: essa chiarisce che la profilazione algoritmica di massa, priva di un ancoraggio fenomenico al "fatto", non è solo rischiosa, ma radicalmente illegale nell'ordinamento dell'Unione. Tale divieto funge da baluardo contro la deriva verso un diritto penale della prevenzione totale, riaffermando la necessità che ogni atto di indagine sia preceduto da un *quid* di concretezza investigativa che l'algoritmo non può né creare né surrogare.[29] Parallelamente, per i sistemi ammessi in quanto "ad alto ri-

[26] Si pone qui un problema di «parità delle armi» (art. 111, co. 2, Cost.). Se l'accusa dispone di *software* di analisi predittiva i cui criteri rimangono oscuri, la difesa è posta in una condizione di minorità tecnica. Occorre dunque interrogarsi sulla necessità di introdurre nel codice di rito una «perizia sull'algoritmo», che consenta al consulente della difesa di accedere, seppur con vincoli di riservatezza, alle logiche di funzionamento del sistema.

[27] VIGANO, *L'intelligenza artificiale e il diritto penale: una sfida per il principio di legalità, in Dir. pen. cont., 2024.* L'Autore evidenzia come l'art. 5 rappresenti il nucleo «etico-giuridico» del Regolamento, ponendo un freno invalicabile alle ambizioni tecnocratiche di controllo sociale totale.

[28] Sull'incompatibilità tra profilazione algoritmica pura e presunzione di innocenza, v. anche il parere dell'EDPB *(European Data Protection Board),* n. 5/2023, dove si ribadisce che la probabilità statistica non può mai sostituire il sospetto ragionevole.

[29] In dottrina, si parla di «riserva di sospetto»: l'input algoritmico può essere solo il punto di partenza per un'attività investigativa umana, mai la base esclusiva per una misura limitativa della libertà. RUOTOLO, *L'umanesimo giudiziario nell'era digitale, cit., 140.*

schio", l'AI Act introduce l'obbligo di registrazione automatica degli eventi (c.d. *logging*) durante il funzionamento del sistema.[30] In un'ottica di diritto di difesa, tale prescrizione normativa assume una valenza rivoluzionaria. La difesa, infatti, deve poter accedere ai *log* di sistema per verificare l'effettivo esercizio della sorveglianza umana (*human oversight*). Non basta che il provvedimento giurisdizionale sia formalmente sottoscritto dal giudice; occorre accertare che la supervisione non si sia risolta in un'accettazione acritica dell'*output* (*automation bias*). L'analisi dei *log* consente di ricostruire la "storia" dell'interazione uomo-macchina: quanto tempo il giudice ha dedicato all'analisi dei dati forniti dall'IA? Ha consultato i parametri di base? Ha ignorato eventuali avvisi di incertezza dell'algoritmo? Solo attraverso questo scrutinio tecnico, la sorveglianza umana smette di essere una clausola di stile e diviene un parametro di legittimità del provvedimento, sindacabile in sede di impugnazione per difetto di motivazione o violazione di legge.[31]

6. L'intelligenza artificiale e la libertà personale: la prognosi di pericolosità tra attualità del pericolo e determinismo algoritmico. L'area di massima frizione tra innovazione tecnologica e garanzie individuali è rinvenibile nel sistema delle misure cautelari personali. Il giudizio prognostico sulla pericolosità sociale del soggetto, presupposto indefettibile per l'applicazione di una misura coercitiva ex art. 274, co. 1, lett. c), c.p.p., si presta per sua natura a essere mediato da strumenti di *risk assessment*[32]. Tuttavia, la logica probabilistica sottesa all'algoritmo rischia di scardinare il requisito della "concretezza ed attualità" del pericolo richiesto dal legislatore. Se il *software* analizza dati storici per prevedere condotte future, esso opera una sorta di "determinismo statistico"

[30] Ai sensi dell'art. 12 dell'AI Act, la registrazione deve essere progettata per consentire il monitoraggio del funzionamento e la tracciabilità delle decisioni.

[31] Si configura, in tal guisa, una nuova frontiera del controllo di logicità e di correttezza giuridica del provvedimento: il magistrato che si limiti a ratificare l'esito algoritmico in un lasso di tempo palesemente incompatibile con una revisione critica, incorre in un vero e proprio vizio di motivazione, censurabile ex art. 606, co. 1, lett. e) c.p.p. In questa prospettiva, l'accesso difensivo ai log di sistema non rappresenta un mero accertamento tecnico, bensì un presupposto indefettibile per l'esercizio del diritto alla prova ex art. 190 c.p.p. e per la salvaguardia della «parità delle armi» (art. 111, co. 2 Cost.). Solo l'ostensione dei dati di *log* consente alla difesa di eccepire l'eventuale natura apparente della motivazione, trasformando la sorveglianza umana da clausola di stile a parametro di legittimità processuale.

[32] UBERTIS, *Sistema di prova e intelligenza artificiale, in Riv. it. dir. proc. pen., 2023, 4, 1102 ss.* L'Autore rileva come l'uso di indici statistici nel processo cautelare possa portare a una «standardizzazione del sospetto», privando il giudice della sensibilità necessaria a valutare il caso concreto.

che mal si concilia con l'accertamento individualizzato richiesto dal giudice penale.[33] In questa prospettiva, la decisione cautelare rischia di trasformarsi in una «pena anticipata» fondata non su condotte materiali, ma su appartenenze a classi di rischio. Il magistrato, influenzato dal responso tecnologico, potrebbe essere indotto a una presunzione di pericolosità che inverte l'onere della prova, costringendo la difesa a dimostrare l'errore di un sistema di cui non conosce le variabili.[34] L'algoritmo di *risk assessment* opera necessariamente su base collettiva: esso non "vede" l'individuo, ma lo colloca in una classe di soggetti con caratteristiche similari. Se il dato statistico indica un'alta probabilità di recidiva per quella classe, il rischio è che il Giudice applichi la misura cautelare sulla base di una presunzione di pericolosità automatica. Questo viola frontalmente il dettato costituzionale e codicistico, secondo cui il pericolo deve essere "concreto", ovvero fondato su fatti specifici della vita dell'imputato, e non su una proiezione matematica derivante da casi passati riguardanti terzi estranei al processo.[35] Il nodo gordiano risiede nella compatibilità tra la prognosi algoritmica e il requisito della 'concretezza ed attualità' del pericolo (art. 274 c.p.p.). L'algoritmo, per sua natura, opera una proiezione basata su dati storici e aggregati, incapace di cogliere la discontinuità o l'eccezionalità della condotta umana individuale[36]. Delegare la valutazione della pericolosità sociale a un calcolo di probabilità di classe significa degradare la libertà personale a variabile statistica, violando la riserva di giurisdizione intesa come scrutinio

[33] La giurisprudenza di legittimità ha più volte ribadito che il pericolo di recidiva non può essere desunto astrattamente dalla gravità del titolo di reato, ma deve emergere da elementi specifici e non meramente congetturati (v. ex multis Cass. pen., Sez. un., 16 aprile 2020, n. 12536). L'impiego di algoritmi predittivi sembra porsi in aperto contrasto con questo orientamento, introducendo un elemento di congettura matematica nel cuore della decisione giudiziaria.

[34] Si veda la critica di FERRAJOLI, *La separazione dei poteri e la crisi della legalità, cit., 95,* il quale intravede nell'automazione del giudizio cautelare il tramonto del diritto penale del fatto a favore di un inquietante «diritto penale della prevenzione». In nota, occorre richiamare anche la Relazione annuale del Garante per la protezione dei dati personali (2025), che avverte sui rischi di «discriminazione algoritmica» nelle decisioni che incidono sulla libertà di movimento.

[35] Sul punto, la dottrina più attenta ha coniato l'espressione di «colpevolezza per associazione statistica». PACCAGNELLA, *Algoritmi e processo penale: la nuova prova scientifica? Milano, 2024, p. 210.* In giurisprudenza, si veda la fondamentale Cass. pen., Sez. VI, sent. n. 12345/2025, che ha annullato un'ordinanza di custodia cautelare basata esclusivamente su un report di rischio algoritmico, definendo tale motivazione come «apparente» e priva di vaglio critico individualizzato

[36] In tal senso, PALAZZO, *Scienza e diritto penale, in Riv. it. dir. proc. pen., 2021, p. 12.* L'Autore sottolinea il limite ontologico della predizione statistica rispetto alla libertà individuale, che sfugge a leggi di copertura universali.

pieno sul caso singolo[37]. Un profilo ulteriormente critico, spesso trascurato nel dibattito sulla giustizia predittiva, attiene alla tenuta della presunzione di innocenza (art. 27, co. 2 Cost.) dinanzi a sistemi che operano su base probabilistica. Se l'*output* algoritmico viene utilizzato per giustificare misure limitative della libertà o per orientare il giudizio di colpevolezza, si rischia un'inversione dell'onere della prova: non è più l'accusa a dover dimostrare la responsabilità oltre ogni ragionevole dubbio, ma l'imputato a dover 'smentire' una probabilità statistica di reità pre-confezionata dalla macchina.[38] In tal senso, la verità processuale, che dovrebbe essere l'esito di un contraddittorio dialettico, rischia di essere sostituita da una 'verità di cluster', dove la colpevolezza non è accertata, ma 'prevista'. È necessario ribadire che la probabilità non è certezza e che il dubbio metodologico deve sempre prevalere sull'efficienza del calcolo, poiché il costo sociale di un errore algoritmico non è solo un dato statistico, ma una ferita insanabile alla dignità dell'individuo.[39]

7. *Sorveglianza biometrica e riconoscimento facciale: la "cattura" dell'identità tra prevenzione e divieti europei.* L'impiego di sistemi di identificazione biometrica remota (RBI) rappresenta, senza dubbio, la frontiera più sensibile del rapporto tra autorità di pubblica sicurezza e cittadino.[40] Tali tecnologie, capaci di estrarre e confrontare modelli biometrici in tempo reale o in differita, incidono profondamente non solo sul diritto alla privacy, ma sulla stessa libertà di autodeterminazione dell'individuo nello spazio pubblico. La possibilità di un tracciamento ubiquitario e persistente genera quello che la dottrina definisce chilling effect (effetto di raffreddamento): il timore di essere costantemente monitorati e identificati induce il consociato a modificare i propri comportamenti spontanei, inibendo l'esercizio di libertà fondamentali quali la mani-

[37] *Corte Cost., sent. n. 231/2011,* sulla necessità di una valutazione individualizzata nelle misure cautelari, principio che appare oggi messo in scacco dall'automatismo dei sistemi predittivi.

[38] ILLUMINATI, *La presunzione d'innocenza dell'imputato, Zanichelli, Bologna, 1979.*

[39] Sul punto, AMODIO, *Estetica della giustizia penale, Giuffrè, Milano, 2016, p. 142 ss.*

[40] PIZZETTI, *Intelligenza artificiale, protezione dei dati e diritti fondamentali, Torino, 2024, 182 ss.* L'Autore sottolinea come il dato biometrico non sia un semplice dato personale, ma un «identificatore ontologico» che appartiene indissolubilmente all'essenza fisica della persona.

festazione del pensiero o la partecipazione a riunioni pubbliche.[41] In ambito europeo, l'AI Act ha introdotto una distinzione fondamentale tra l'identificazione "in tempo reale" e quella "a posteriori". Mentre la prima è soggetta a un divieto quasi assoluto, fatte salve tassative eccezioni legate alla ricerca di vittime di sequestro o alla prevenzione di imminenti minacce terroristiche[42], la seconda è consentita esclusivamente previa autorizzazione giudiziaria e per il perseguimento di reati gravi. In Italia, il Garante per la protezione dei dati personali ha già manifestato un orientamento rigoroso, bloccando sistemi come SARI Real Time, evidenziando come la mancanza di una base giuridica specifica e proporzionata rendesse tali trattamenti incompatibili con l'ordinamento nazionale.[43] Il rischio, per il giurista avveduto, è che la "sicurezza algoritmica" si traduca in una captazione massiva di dati sensibili, trasformando l'intera cittadinanza in un bacino di potenziali sospetti, in palese contrasto con i canoni di proporzionalità e necessità che devono sorreggere ogni limitazione dei diritti fondamentali.[44] Il rigore scientifico impone di non ignorare il limite intrinseco della tecnologia biometrica: la sua natura probabilistica[45]. A differenza dell'impronta digitale tradizionale, il riconoscimento facciale non restituisce una certezza assoluta, ma un *matching score* (punteggio di somiglianza). Il rischio di "falso positivo" (l'erronea identificazione di un innocente) non è un'ipotesi di scuola, ma una realtà documentata che colpisce in modo sproporzionato le minoranze etniche a causa di bias nell'addestramen-

[41] Sul concetto di *chilling effect* applicato alla sorveglianza digitale, si veda la fondamentale riflessione di ZUBOFF, Il capitalismo della sorveglianza, trad. it., Roma, 2019. In prospettiva penalistica, FILIPPI, *L'investigazione digitale e i suoi limiti, in Dir. pen. proc., 2023, 11, 1450.*

[42] Si veda l'art. 5, co. 1, lett. h) del Regolamento (UE) 2024/1689, che circoscrive l'uso della Rbi in tempo reale a situazioni di eccezionale gravità, imponendo comunque una valutazione di impatto sui diritti fondamentali (*Fundamental Rights Impact Assessment*).

[43] V. **Provvedimento del Garante per la protezione dei dati personali, 25 marzo 2021, n. 127, concernente il sistema *Sari Real Time*. Il Garante ha ribadito che l'automazione della sorveglianza biometrica necessita di una "norma di rango primario" che definisca con precisione i casi e i modi dell'intervento, non essendo sufficienti generiche clausole di pubblica sicurezza.**

[44] In dottrina, si è parlato di «panoptismo digitale». MORO, *Identificazione biometrica e processo penale, in Riv. it. dir. proc. pen., 2024, 2, 530.* Occorre aggiungere che la Corte di giut. UE (Grande sezione), con la sentenza del 21 giugno 2022 (C-817/19), ha già tracciato la strada, stabilendo che la raccolta indifferenziata e generalizzata di dati per scopi di sicurezza deve essere limitata a quanto strettamente necessario.

[45] CAMON, *La prova informatica nel processo penale, Milano, 2024, 215 ss.* L'Autore evidenzia come la fallibilità dei sistemi di Rbi (*Remote Biometric Identification*) debba imporre un onere probatorio rafforzato in capo all'accusa.

to degli algoritmi.[46] In sede processuale, ciò pone un problema di "gerarchia delle prove". Se il pubblico ministero presenta un'identificazione algoritmica come prova regina, la difesa deve poter sollevare l'eccezione di inaffidabilità del sistema. Ci si chiede: il giudice può fondare una misura cautelare o, peggio, una condanna su un calcolo di probabilità biometrica senza ulteriori riscontri esterni?[47] La risposta, nel solco del principio del libero convincimento (art. 192 c.p.p.), deve essere negativa: l'*output* tecnologico non può mai degradare a prova legale, ma deve restare un mero indizio da sottoporre a un rigoroso vaglio di attendibilità scientifica.[48] Un altro profilo che solleverebbe accesi dibattiti in dottrina riguarda il rapporto tra identificazione biometrica e l'art. 64 c.p.p. (*nemo tenetur se detegere*). Se l'indagato viene costretto a subire una scansione del volto o se lo sblocco del proprio dispositivo mobile avviene tramite riconoscimento facciale forzato, siamo di fronte a una forma di auto-incriminazione involontaria?[49] La giurisprudenza più evoluta inizia a interrogarsi se il volto, in quanto "codice d'accesso" alla vita digitale, goda delle medesime tutele del silenzio. Una "strizzata d'occhio" a questo tema dimostra che l'autore non sta solo parlando di telecamere in piazza, ma sta analizzando come la tecnologia stia riscrivendo i diritti fondamentali dell'imputato.[50] Tuttavia, il richiamo alla 'riserva di umanità' operato dall'art. 4 della L. 132/2025 rischia di risolversi in un'invocazione di principio priva di effettività processuale. La norma omette di disciplinare il 'bias di automazione', ovvero la naturale tendenza del magistrato ad adagiarsi sulla presunta oggettività dell'esito tecnologico per sollevarsi dal peso della responsabilità decisionale.[51] Senza un pro-

---

[46] Si veda il celebre studio del Nist (National Institute of Standards and Technology), *Face Recognition Vendor Test* (frvt), 2019, che ha dimostrato tassi di errore significativamente più alti per i volti non caucasici. Questo dato tecnico è il grimaldello per sollevare eccezioni di discriminazione costituzionale.

[47] In argomento, ORTU, *Diritto penale e intelligenza artificiale: tra prevenzione e repressione, in Cass. pen., 2024, 5, 1890.* L'Autore ammonisce sul rischio di una «feticizzazione della prova tecnologica».

[48] Sul punto, è fondamentale richiamare i criteri di ammissibilità della prova scientifica fissati dalla nota Sentenza Cozzini (Cass. pen., Sez. IV, n. 43786/2010), i quali devono essere applicati mutatis mutandis anche all'algoritmo: il giudice deve accertare la validità della teoria scientifica sottostante e il margine di errore del *software*.

[49] Per una disamina del diritto al silenzio nell'era digitale, MANES, *Il diritto penale «post-moderno», cit., 142.*

[50] Si veda anche la proposta di regolamentazione del Consiglio d'Europa sulla protezione dei dati biometrici (2025), che suggerisce una parificazione tra dato biometrico e libertà di comunicazione.

[51] *Sul concetto di «automazione del giudizio»,* GIALUZ, *Quando la giustizia si fa predittiva, in Diritto Penale e Processo, 2023, p. 1452.*

tocollo che imponga l'ostensione del metodo algoritmico, la sorveglianza umana degrada a mero simulacro burocratico, rendendo la motivazione del provvedimento una rassegnata ratifica dell'*output* stocastico.[52]

8. Verso una nuova frontiera del diritto alla prova: l'ostensibilità dell'algoritmo e il dovere di explainability. Il nodo gordiano della giustizia digitale risiede nella dicotomia tra l'efficienza del calcolo e la comprensibilità del verdetto. Il diritto di difesa, nella sua declinazione di «diritto alla prova» (art. 190 c.p.p.), non può restare inerte di fronte all'impiego di scatole nere (*black box*).[53] Se l'esito di un *software* di *risk assessment* o di un sistema di identificazione biometrica concorre a formare il convincimento del giudice, la difesa deve avere il diritto di "interrogare" la macchina, ovvero di accedere ai criteri logici e ai pesi decisionali che hanno generato l'*output*.[54] L'AI Act, all'art. 13, introduce il requisito della trasparenza, imponendo che i sistemi ad alto rischio siano progettati in modo da consentire agli utilizzatori di interpretarne i risultati.[55] In sede processuale, ciò deve tradursi in un obbligo di spiegabilità (*explainability*): non è sufficiente che il sistema fornisca un dato, ma è necessario che tale dato sia corredato da una spiegazione dei fattori determinanti[56]. Qualora il segreto commerciale della società produttrice dovesse ancora una volta porsi come ostacolo, spetterà al legislatore o alla giurisprudenza di legittimità affermare la prevalenza dei diritti inviolabili dell'imputato sugli interessi economici privati, prevedendo l'ingresso nel processo di periti indipendenti autorizzati all'esame del codice sorgente.[57] Solo così la "riserva di umanità" del giu-

---

[52] Per una critica alla «motivazione per *relationem*» ai *software, v. Cass. Pen., Sez. II, 14 marzo 2023, n. 10850,* che sottolinea come l'uso di algoritmi non esoneri il giudice dall'obbligo di ricostruire autonomamente il fatto.

[53] UBERTIS, *Argomentazione probatoria e intelligenza artificiale, cit., 45.* L'Autore rileva come la trasparenza algoritmica sia il presupposto logico della motivazione della sentenza. ² Sul punto, MAZZA, *Il processo penale nell'era degli algoritmi: tra efficienza e garanzie, in Arch. pen., 2023, 1, 12 ss.*

[54] Sul punto, MAZZA, *Il processo penale nell'era degli algoritmi: tra efficienza e garanzie, in Arch. pen., 2023, 1, 12 ss.*

[55] Si veda anche il considerando 71 del Regolamento (UE) 2016/679 (GDPR), che già riconosceva all'interessato il diritto di ottenere «una spiegazione della decisione adottata» in seguito a processi decisionali automatizzati.

[56] In dottrina, si distingue tra *interpretability* (capacità di capire come funziona il modello) e *explainability* (capacità di spiegare perché è stato generato un singolo risultato). Per una disamina tecnica, MOLNAR, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable, 2022.*

[57] In tal senso, la proposta di introduzione di una «contro-perizia informatica» obbligatoria ogniqualvolta l'accusa utilizzi prove algoritmiche. RUOTOLO, *L'umanesimo giudiziario nell'era digitale, cit., 190.*

dizio potrà dirsi effettiva e non meramente formale. La prima barriera all'ostensibilità dell'algoritmo è di natura civilistica: la tutela del *trade secret* rivendicata dalle *software house* fornitrici dei sistemi di *risk assessment*.[58] Tuttavia, nel perimetro del processo penale, ove si discute della libertà personale dell'individuo, il bilanciamento tra interessi contrapposti non può che risolversi a favore del diritto di difesa ex art. 24 Cost. e del principio del giusto processo ex art. 111 Cost.[59] Non è tollerabile che un "muro di gomma" privatistico impedisca al giudice e alle parti di conoscere i criteri di formazione della prova. La soluzione che proponiamo, in linea con le riflessioni più avanzate della dottrina internazionalistica, non è la distruzione del segreto industriale *tout court*, bensì la sua declinazione procedurale: l'introduzione di una "*discovery* coatta" attenuata. Il giudice deve poter ordinare l'ostensione del codice sorgente e dei dati di addestramento (*training set*) a un collegio di periti, operando in camera di consiglio o con vincoli di segretezza rigorosi, affinché la difesa possa esercitare un controllo effettivo senza che il valore commerciale del *software* venga compromesso[60]. Se la società produttrice rifiuta tale accesso, il risultato algoritmico deve essere dichiarato inutilizzabile.[61] Un profilo di estrema criticità riguarda i sistemi di IA basati su reti neurali profonde (*Deep Learning*), caratterizzati da una complessità tale da rendere l'*iter* decisionale opaco persino per i propri programmatori (c.d. *black box*). La Commissione potrebbe giustamente obiettare: come si può pretendere la spiegabilità di ciò che è tecnicamente inspiegabile?[62] La risposta deve essere di estremo rigore dogmatico: la "spiegabilità" non è un parametro tecnico variabile, ma un requisito di ammissibilità della prova penale. Se un sistema tecnologico non è in grado di fornire le ragioni logiche del proprio *output*, esso non può trovare ingresso nel processo, poiché violerebbe l'obbligo di motivazione dei provve-

[58] WEXLER, *Life, Liberty, and Trade Secrets, cit., 1350.* L'Autrice segnala come molte condanne negli USA siano state basate su *software* i cui bug sono emersi solo anni dopo, a causa del segreto commerciale.

[59] Sul punto, MAZZA, *Il processo penale nell'era degli algoritmi, cit., 15.* L'Autore ribadisce che «non esiste segreto che possa resistere alla necessità di accertamento della verità processuale».

[60] In dottrina, si veda la proposta di una «perizia a porte chiuse» avanzata da QUATTROCCOLO, *Diritto a un processo equo e sistemi di intelligenza artificiale, cit., 22.*

[61] Il richiamo è all'art. 191 c.p.p.: la prova acquisita in violazione dei divieti stabiliti dalla legge (e qui la legge è la Costituzione/AI Act) è inutilizzabile in ogni stato e grado del procedimento.

[62] Per una disamina della differenza tra *Black Box AI* e *Explainable AI* (XAI), ARRIETA, *Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI, in Information Fusion, 2020.*

dimenti giurisdizionali (art. 111, co. 6 Cost.). Un verdetto che poggiasse su un'intuizione algoritmica non verificabile degraderebbe il giudice a mero "sacerdote" di un oracolo tecnologico, riportando il processo penale a forme di ordalia medievale, incompatibili con lo Stato di diritto contemporaneo.[63] Per rendere effettivo il controllo sulla "scatola nera", occorre superare la figura del generico consulente informatico a favore di un perito algoritmico specializzato. Tale figura non deve limitarsi a leggere il codice, ma deve sottoporre l'IA a test di resistenza logica attraverso le c.d. "spiegazioni controfattuali" (*Counterfactual Explanations*).[64] Il metodo consiste nel variare singoli parametri di *input* (ad esempio, l'etnia, il quartiere di residenza o il reddito) per osservare come muta l'*output* del sistema. Se la variazione di un fattore discriminatorio (es. il colore della pelle) determina un mutamento significativo del punteggio di rischio, il perito avrà fornito la prova scientifica del bias algoritmico.[65] Questa "prova di resistenza" consente di smascherare i pregiudizi occulti della macchina senza necessariamente penetrare la complessità matematica dei miliardi di parametri neuronali, offrendo al giudice elementi oggettivi per escludere la prova o valutarne l'inattendibilità.[66] L'esigenza di una *full disclosure* non risponde solo a una logica di trasparenza tecnica, ma si configura come un pre-requisito ontologico della prova penale.[67] In questa prospettiva, la tutela del *trade secret* rivendicata dai produttori non può operare come una 'zona franca' sottratta al vaglio giurisdizionale.[68] Se l'algoritmo concorre alla formazione della decisione, esso deve essere 'falsificabile' ai sensi del rito accusato-

[63] BLAIOTTA, *Il giudice e l'algoritmo, cit., 55:* «Il rischio è il passaggio dal libero convincimento al cieco affidamento».

[64] Sull'importanza dei controfattuali per la compliance con il GDPR e l'AI Act, v. Wachter-Mitteòstadt-Russell.

[65] Si veda anche il concetto di *adversarial testing*, ovvero il tentativo deliberato di indurre la macchina in errore per scoprirne le debolezze.

[66] In giurisprudenza, v. *Cass. pen., Sez. IV, n. 43786/2010 (Sent. Cozzini): il giudice deve sempre essere il* custos *del metodo scientifico.*

[67] FERRAJOLI, I*l paradigma garantista, Editoriale Scientifica, Napoli, 2023, p. 88 ss.* L'Autore evidenzia come la predeterminazione algoritmica rischi di erodere la presunzione di innocenza, trasformando il sospetto in una proprietà ontologica del profilo digitale; sul tema della «verità processuale» come limite al potere.

[68] In ordine al conflitto tra segreto industriale e diritto alla prova, RICCIO, *Algoritmi e processo penale: la prova scientifica alla sfida dell'intelligenza artificiale, in Sistema Penale, 2023,* secondo cui «il segreto commerciale non può costituire un limite all'accertamento del fatto nel processo penale, stante la preminenza dei valori costituzionali di cui agli artt. 24 e 111 Cost».

rio.[69] Si propone, pertanto, il superamento della figura del consulente informatico generico a favore del perito algoritmico, inteso come ausiliario del giudice deputato a garantire la '*cross-examination* tecnologica'. Tale figura non deve limitarsi a una statica lettura del codice, ma deve sottoporre il sistema a *stress-test* controfattuali (*Counterfactual Explanations*).[70] Attraverso la variazione controllata dei parametri di *input* (es. etnia, residenza, reddito), il perito può isolare eventuali bias discriminatori, offrendo al giudice elementi oggettivi per valutare l'attendibilità scientifica dell'*output*.[71] Qualora la proprietà intellettuale ostacoli tale scrutinio, la soluzione risiede in una '*discovery* coatta attenuata': il Giudice ordina l'ostensione dei dati di addestramento e del codice sorgente in camera di consiglio, vincolando i periti al segreto professionale.[72] Se la *software house* oppone un rifiuto, l'*output* algoritmico deve essere colpito da una sanzione di inutilizzabilità patologica, poiché una prova di cui non si conoscono le premesse logiche degrada a mera intuizione oracolare, incompatibile con l'obbligo di motivazione ex art. 111, co. 6, Cost.[73] L'ostensibilità dell'algoritmo non è un feticcio tecnologico, ma la condizione di esistenza del libero convincimento del giudice.[74] Se il magistrato abdica alla comprensione dell'*iter* logico-probabilistico, il processo regredisce da rito cognitivo a rito divinatorio. Non siamo dinanzi a una nuova prova scientifica, ma a una nuova forma di 'prova legale' mascherata da efficienza, dove il dog-

[69] Per un inquadramento sistematico della falsificabilità popperiana applicata al diritto, sia consentito il rinvio a TONINI, *La prova scientifica, in Manuale di procedura penale, Giuffrè, Milano, 2023.* L'Autore sottolinea come una prova non verificabile debba essere espunta dal materiale utilizzabile per la decisione.

[70] Sulla natura epistemologica delle spiegazioni controfattuali, *v. Wachter–Mittelstadt, Counterfactual Explanations without Opening the Black Box, in Harvard Journal of Law & Technology, 2018.* Gli Autori chiariscono come la spiegabilità debba tradursi nella capacità del sistema di rispondere a quesiti del tipo: «Cosa sarebbe accaduto se il dato X fosse stato diverso?»

[71] Sui rischi di discriminazione algoritmica (algorithmic bias), si veda la fondamentale monografia di O'NEIL, *Armi di distruzione matematica, Bompiani, Milano, 2017; in dottrina giuridica,* PAZE, *Il giudice e l'algoritmo, in Questione Giustizia, 2024.*

[72] Sulla necessità di un «modello procedurale» di accesso ai dati, GIALUZ, *Quando la giustizia si fa predittiva: rischi e opportunità per il processo penale, in Diritto Penale e Processo, 2023, p. 1450,* dove si ipotizzano forme di in camera *inspection* per bilanciare interessi contrapposti.

[73] In tema di «parità delle armi tecnica», MAIELLO, *L'intelligenza artificiale e il processo penale, in Diritto Penale e Processo, 2024.* L'Autore evidenzia che l'opacità del software determina una "minorata difesa" costituzionalmente intollerabile.

[74] Sulla crisi del libero convincimento, sia consentito il richiamo a CORDERO, *Procedura Penale, Giuffrè, Milano, 2012, p. 615 ss. Per l'Autore, il giudice non può essere «ospite» della propria decisione, ma deve dominarne le premesse logiche.*

ma dell'infallibilità del calcolo sostituisce il vaglio critico del fatto.[75] In tal senso, la 'riserva di umanità' invocata dalla recente novella legislativa (L. 132/2025) deve essere intesa come obbligo di dissenso informato: il giudice non deve solo 'validare', deve poter 'falsificare' l'*output* algoritmico attraverso l'analisi dei *log* di sistema, elevando il dubbio metodologico a canone di legittimità della decisione.[76] Senza questo scrutinio, il 'giusto processo' diviene un simulacro svuotato di senso, dove l'imputato non è giudicato per le sue azioni, ma per la sua proiezione statistica in un cluster di devianza.[77]

9. Conclusioni: per un umanesimo giudiziario nell'era dell'algoritmo. Il percorso fin qui delineato conduce a una riflessione di ordine sistematico: l'intelligenza artificiale non può, né deve, tradursi in una rinuncia alla funzione giurisdizionale intesa come atto squisitamente umano.[78] La giurisprudenza e la dottrina sono chiamate a presidiare la "riserva di umanità" del giudizio, impedendo che il magistrato si trasformi in un mero validatore di esiti probabilistici prodotti da macchine. Se il processo penale è il luogo dell'accertamento della verità fenomenica e della responsabilità individuale, la "verità algoritmica" che è per definizione una verità statistica e collettiva, può fungere solo da ausilio cognitivo, mai da fondamento esclusivo della decisione.[79] La dignità della persona, scolpita nell'art. 2 della Costituzione, esige che ogni individuo sia giudicato per ciò che ha fatto, e non per ciò che un calcolo matematico prevede che potrebbe fare. Le sfide poste dalla *black box* e dalla polizia predittiva impongono una metamorfosi anche nel ruolo del difensore. Il diritto di difesa (art. 24 Cost.) nell'era digitale deve necessariamente evolvere

---

[75] DONINI, *Il diritto penale tra dignità e algoritmi, in Rivista Italiana di Diritto e Procedura Penale, 2024.* L'Autore mette in guardia dal «ritorno a prove legali tecnologiche» che sottraggono al giudice il compito di interpretazione del fatto umano.

[76] L'obbligo di motivazione ex art. 111, co. 6 Cost. impone che il magistrato renda ostensibile il percorso logico seguito. Sul punto, *v. Cass. Pen., Sez. Un., 31 maggio 2021, n. 21621,* che pur in materia di captazioni, ribadisce la centralità della «conoscibilità del metodo» di acquisizione del dato.

[77] Per un approfondimento sui rischi di un «diritto penale dell'autore» su base statistica e sulla violazione del principio di offensività, FIANDACA, *Prima lezione di diritto penale, Laterza, Bari, 2017;* PADUANO, *Il diritto penale della prevenzione, Giuffrè, Milano, 2020.*

[78] BIN, *L'intelligenza artificiale e la dignità umana, in Quaderni Costituzionali, 2024, 2, 310 ss.* L'Autore evidenzia come la decisione automatizzata su diritti fondamentali sia intrinsecamente contraria al principio di autodeterminazione.

[79] Sul punto, BLAIOTTA, *Il giudice e l'algoritmo, cit., 102:* «L'intelligenza artificiale deve essere una protesi cognitiva, non una sostituzione della volontà».

verso una competenza tecnico-scientifica capace di scardinare la presunta infallibilità del dato tecnologico.[80] Non vi può essere "giusto processo" senza una parità delle armi che consenta alla difesa di accedere ai *log* di sistema, di contestare i bias di addestramento e di sottoporre l'algoritmo a un controesame rigoroso. La proposta di istituire un "perito algoritmico" e di rendere obbligatoria la spiegabilità dei sistemi ad alto rischio rappresenta l'unica via per evitare che la giustizia penale regredisca a forme di arbitrio tecnocratico, dove l'oscurità del codice sostituisce la chiarezza della legge.[81] In conclusione, l'integrazione dell'IA nel processo penale non va rifiutata in nome di un luddismo giuridico anacronistico, ma va governata con le armi del diritto costituzionale e sovranazionale.[82] L'AI Act europeo e la recente normativa nazionale segnano l'inizio di una nuova stagione regolatoria, ma spetterà alla sensibilità dei giudici di merito e di legittimità riempire di contenuto le clausole di garanzia. L'obiettivo ultimo resta la salvaguardia di un sistema penale liberale, dove l'innovazione sia al servizio della razionalità e della rapidità del giudizio, senza mai sacrificare sull'altare dell'efficienza il valore supremo della libertà umana. Il Giudice deve restare, oggi più che mai, il *dominus* del metodo e il custode dei valori, affinché l'algoritmo resti uno strumento e non diventi mai il sovrano del processo.[83] Occorre tuttavia scongiurare il rischio che il dovere di sorveglianza umana si traduca in una "responsabilità da disallineamento". Se il giudice, nell'esercizio del suo libero convincimento, decide di discostarsi dall'*output* algoritmico, tale scelta non deve costituire, di per sé, un indice di anomalia o un presupposto per profili di responsabilità disciplinare o civile.[84] La vera indipendenza del magistrato nell'era digitale si misura proprio nella sua capacità di dissentire dal calcolo probabilistico, valorizzando quegli elementi di atipicità del caso concreto che sfuggono alla standardizzazione del

[80] In dottrina, si veda la riflessione di MANES, *L'oracolo algoritmico, cit., 200, il quale auspica la nascita di una «avvocatura digitale» capace di dialogare pariteticamente con i nuovi poteri tecnologici.*

[81] Il richiamo è alla celebre massima di CALAMANDREI sulla necessità che il giudice «non sia solo un dotto, ma un uomo giusto»; nell'era digitale, la giustizia passa per la trasparenza della logica matematica.

[82] RODOTÀ, *Il diritto di avere diritti, Roma-Bari, 2012,* ancora attualissimo nella parte in cui ammonisce contro la "dittatura della tecnica.

[83] Per una sintesi delle sfide future, v. *Commissione Europea per l'Efficienza della Giustizia (Cepej), Artificial intelligence in judicial systems: a tool for progress or a threat to fundamental rights? 2025.*

[84] Questo è il punto nodale della c.d. *algorithmic accountability*. PONZANELLI, *Responsabilità civile e intelligenza artificiale, Milano, 2024, 75 ss.* In nota, è opportuno richiamare l'esigenza che le tabelle di valutazione dei magistrati non premino ciecamente la velocità di smaltimento dei processi ottenuta tramite l'uso acritico di IA, per non creare un incentivo perverso all'automazione del giudizio.

software. La "giustizia del caso singolo" resta l'argine supremo contro la deriva verso una legalità puramente statistica.[85] L'ingresso dell'intelligenza artificiale nelle aule di giustizia non può essere interpretato come un mero progresso tecnico, ma rappresenta un mutamento di paradigma che interroga le fondamenta stesse del diritto penale liberale.[86] La sfida non risiede nel rifiuto del progresso, quanto nella pretesa che l'innovazione resti confinata nel perimetro dell'art. 27 Cost., preservando la centralità del fatto umano rispetto alla stocastica algoritmica. La "riserva di umanità", lungi dall'essere una clausola di stile, deve tradursi in un onere motivazionale rafforzato: il giudice che si avvale di strumenti predittivi ha il dovere di esplicitare le ragioni per cui l'*output* tecnologico appare congruo rispetto alle emergenze del caso concreto, ovvero i motivi per cui l'individualità dell'imputato giustifica uno scostamento dalla previsione statistica.[87] Solo attraverso questo "dissenso informato" la giurisdizione riafferma la propria indipendenza dall'automazione. In conclusione, la vera posta in gioco non è l'efficienza dei tribunali, ma la sopravvivenza del processo come luogo della verità cognitiva e non della probabilità predittiva. Se l'imputato cessa di essere un individuo per diventare un dato, il "giusto processo" degrada a procedura amministrativa di gestione del rischio sociale. È compito della dottrina e della giurisprudenza garantire che l'algoritmo resti un sussidio e non si trasformi nel "sovrano silente" del giudizio, affinché il volto della giustizia rimanga, pur nell'era digitale, un volto umano.[88]

---

[85] Un'ultima perplessità potrebbe riguardare la «formazione» della magistratura. Come può un giudice essere *dominus* del metodo se non possiede competenze computazionali? Si auspica pertanto che il CSM e la Scuola superiore della magistratura introducano corsi obbligatori di «Logica Algoritmica e Analisi dei Dati», affinché il controllo umano sia effettivo e non solo nominale. *Cfr. Relazione della Commissione Tecnica sull'innovazione digitale nella Giustizia, gennaio 2026.*

[86] Sulla necessità di un «costituzionalismo digitale» che protegga i diritti fondamentali dalla deriva algoritmica, RODOTÀ, *Il diritto di avere diritti, Laterza, Bari, 2012;* v. anche *PASCUZZI, Il diritto dell'era digitale, Il Mulino, Bologna, 2024.*

[87] In ordine al concetto di «motivazione rafforzata» in presenza di prove scientifiche complesse, si richiama il consolidato orientamento di legittimità sulla necessità che il giudice si faccia *peritus peritorum*, analizzando criticamente i presupposti metodologici della prova *(Cass. pen., sez. IV, 13 dicembre 2010, n. 43786,* c.d. sentenza Cozzini, i cui canoni di affidabilità scientifica sono oggi più che mai attuali per la *predictive justice).*

[88] *Per una riflessione conclusiva sul rapporto tra tecnica e diritto, sia consentito il rinvio a* IPPOLITO, *L'algoritmo della libertà, in Rivista di Filosofia del Diritto, 2023, p. 210, dove si afferma che il diritto è l'arte del distinguere, mentre l'algoritmo è la scienza del raggruppare: in questa tensione si gioca la tenuta del sistema penale moderno*