

INTELLIGENZA ARTIFICIALE E GIUSTIZIA PENALE PREDITTIVA NEL DIRITTO EUROPEO E INTERNAZIONALE: OPPORTUNITÀ, RISCHI E TUTELA DEI DIRITTI FONDAMENTALI

di **Maria Giulia SCOGNAMIGLIO***

ABSTRACT

Il presente contributo analizza il fenomeno della giustizia predittiva attraverso la prospettiva del diritto europeo e internazionale, esaminando il rapporto tra innovazione tecnologica e principi dello Stato di diritto. L'analisi considera il progressivo sviluppo di un quadro normativo multilivello, composto sul piano del Consiglio d'Europa, dalla Convenzione europea dei diritti dell'uomo e dalla Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale e, sul piano dell'Unione europea, dalla Carta dei diritti fondamentali dell'Unione europea, dall'Artificial Intelligence Act, dal regolamento generale sulla protezione dei dati personali.

*Particolare attenzione sarà dedicata alle criticità derivanti dall'utilizzo di sistemi algoritmici in ambito penale, con riferimento ai rischi di discriminazione, alla trasparenza dei modelli decisionali, alla responsabilità degli operatori e alla compatibilità della logica predittiva con principi fondamentali quali la presunzione di innocenza e la responsabilità per fatto proprio. Attraverso l'analisi del caso *State v. Loomis* e del paradigma della pre-criminalità rappresentato dal celebre romanzo di Philip Dick *Minority Report*, il contributo evidenzia come l'utilizzo dell'intelligenza artificiale nella giustizia penale richieda un equilibrio costante tra esigenze di sicurezza e tutela della persona.*

La sfida futura non consiste pertanto nella sostituzione del giudizio umano con sistemi automatizzati, ma nella costruzione di un modello di governance tecnologica capace di garantire che l'innovazione rimanga compatibile con la dignità della persona e con lo Stato di diritto.

*The article takes as its point of departure the recent judgment no. 7898/2025 of the Council of State to examine the primary jurisprudential uncertainties regarding the application of Minimum Environmental Criteria (CAM). Specifically, it analyzes the mandatory nature of CAM as a parameter for the legality of the *lex specialis*, as well as the legal grounds exempting economic operators from the burden of immediate challenge against non-compliant tender notices.*

Through a reconstruction of the regulatory evolution, the article explores the function of CAM within Green Public Procurement (GPP), highlighting how environmental protection has become a structural and binding requirement for administrative action.

SOMMARIO

- 1. L'intelligenza artificiale nella giustizia penale: opportunità e rischi nel quadro europeo e internazionale2**
 - 1.1. L'intelligenza artificiale: definizione, caratteristiche e rilevanza giuridica2**
 - 1.2. L'impiego dell'intelligenza artificiale nella giustizia penale e le implicazioni sui diritti fondamentali3**
 - 1.3. La costruzione di un quadro normativo europeo e internazionale per l'intelligenza artificiale.....5**
- 2. Giustizia predittiva e strumenti algoritmici nella giustizia penale: applicazioni e criticità9**
 - 2.1. Origini ed evoluzione della giustizia predittiva.....9**
 - 2.2. Origini ed evoluzione della giustizia predittiva.....9**
 - 2.2.1. Predictive policing9**
 - 2.2.2. Risk assessment tools 10**
 - 2.2.3. Identificazione biometrica e sorveglianza algoritmica 10**
 - 2.3. Intelligenza artificiale criminale? La responsabilità nell'era degli algoritmi... 11**
 - 2.4. Il caso *State v. Loomis*: trasparenza algoritmica e diritto a un equo processo 11**
- 3. Giustizia predittiva e tutela dei diritti fondamentali: il difficile equilibrio tra innovazione e Stato di diritto 12**

* Dottoressa in giurisprudenza presso l'Università Suor Orsola Benincasa.

3.1. Dalla previsione del rischio alla pre-criminalità: il paradigma di Minority Report	13
3.2. Trasparenza, spiegabilità e controllo umano come garanzie dello Stato di diritto.	13
3.3. Sicurezza pubblica e tutela dei diritti fondamentali: verso un modello europeo di governance dell'IA	14
4. Conclusioni.....	15

1. L'INTELLIGENZA ARTIFICIALE NELLA GIUSTIZIA PENALE: OPPORTUNITÀ E RISCHI NEL QUADRO EUROPEO E INTERNAZIONALE

L'intelligenza artificiale costituisce una delle trasformazioni tecnologiche più rilevanti dell'epoca contemporanea e rappresenta una sfida crescente per gli ordinamenti giuridici, chiamati a confrontarsi con sistemi capaci di elaborare informazioni, individuare correlazioni e produrre risultati idonei a incidere sui processi decisionali.¹ Tale fenomeno assume particolare rilevanza nell'ambito della giustizia penale, settore nel quale l'impiego di strumenti algoritmici può incidere direttamente sull'esercizio del potere pubblico e, conseguentemente, sulla sfera dei diritti fondamentali dell'individuo.

L'utilizzo dell'intelligenza artificiale nel settore penale si inserisce infatti in una tensione costante tra due esigenze potenzialmente contrapposte: da un lato, la ricerca di una maggiore efficienza nell'amministrazione della giustizia e il

rafforzamento degli strumenti di prevenzione e contrasto della criminalità; dall'altro, la necessità di preservare le garanzie fondamentali proprie dello Stato di diritto. Proprio per tale ragione, la regolazione dell'intelligenza artificiale non può essere considerata una questione esclusivamente tecnica, ma assume una dimensione propriamente giuridica, nella quale vengono in rilievo principi quali dignità umana, uguaglianza, tutela della vita privata, diritto a un equo processo e divieto di discriminazione.²

1.1. L'INTELLIGENZA ARTIFICIALE: DEFINIZIONE, CARATTERISTICHE E RILEVANZA GIURIDICA

Definire l'intelligenza artificiale rappresenta il primo passaggio necessario per comprenderne l'impatto sul diritto. Sebbene non esista una definizione universalmente condivisa, con tale espressione si fa generalmente riferimento a sistemi tecnologici capaci di elaborare dati, apprendere dall'esperienza e produrre risultati assimilabili, sotto alcuni profili, a processi cognitivi umani.³

A differenza dei tradizionali strumenti informatici, fondati sull'esecuzione automatica di istruzioni previamente determinate, i sistemi di intelligenza artificiale più avanzati operano attraverso tecniche di apprendimento automatico (*machine learning*)⁴ e apprendimento profondo (*deep learning*), individuando schemi ricorrenti all'interno di grandi quantità di dati e formulando previsioni o classificazioni sulla base delle informazioni elaborate.

Tale caratteristica assume particolare rilievo nel settore della giustizia penale. Gli algoritmi non si limitano infatti a svolgere una funzione meramente strumentale, ma possono contribuire

¹ L. FLORIDI, *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*, Oxford, Oxford University Press, 2023, parte I, capp. 1-3; S. RUSSELL - P. NORVIG, *Artificial Intelligence: A Modern Approach*, 4^a ed., Harlow, Pearson, 2021, pp. 1-35.

² M. HILDEBRANDT, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology*, Cheltenham, Edward Elgar Publishing, 2015, cap. 5, *Threats to Fundamental Rights in the Onlife World*, pp. 77-103; F. PASQUALE, *The Black Box Society: The Secret Algorithms*

That Control Money and Information, Cambridge (MA), Harvard University Press, 2015, cap. 2, pp. 19-58.

³ H. SURDEN, *Artificial Intelligence and Law: An Overview*, in *Georgia State University Law Review*, vol. 35, 2019, pp. 1305-1337; M. HILDEBRANDT, *Law for Computer Scientists and Other Folk*, Oxford, Oxford University Press, 2020.

⁴ S. RUSSELL - P. NORVIG, *Artificial Intelligence: A Modern Approach*, cit.; L. FLORIDI, *The Ethics of Artificial Intelligence*, cit.



alla valutazione di situazioni complesse, influenzando decisioni relative alla prevenzione del crimine, alla valutazione del rischio e, in alcuni ordinamenti, persino alla determinazione della risposta sanzionatoria.

Proprio la capacità predittiva di tali sistemi costituisce il principale elemento di interesse e, al contempo, di criticità. L'algoritmo non accerta necessariamente un fatto già verificatosi, ma elabora probabilità sulla base di dati storici e correlazioni statistiche.⁵ Questa caratteristica pone interrogativi fondamentali circa la possibilità di utilizzare valutazioni probabilistiche nell'ambito di decisioni che possono incidere sulla libertà personale e sugli altri diritti garantiti dagli strumenti internazionali ed europei di tutela.

L'introduzione dell'intelligenza artificiale nella giustizia penale richiede pertanto un approccio fondato non soltanto sull'efficienza tecnologica, ma anche sulla compatibilità dei sistemi utilizzati con i principi fondamentali che governano l'esercizio del potere pubblico. In tale prospettiva, il problema centrale non consiste nel determinare se l'intelligenza artificiale possa essere utilizzata nella giustizia, ma nel definire quali condizioni giuridiche debbano essere rispettate affinché il suo impiego rimanga conforme agli standard europei e internazionali di protezione dei diritti fondamentali.

1.2. L'IMPIEGO DELL'INTELLIGENZA ARTIFICIALE NELLA GIUSTIZIA PENALE E LE IMPLICAZIONI SUI DIRITTI FONDAMENTALI

L'applicazione dell'intelligenza artificiale nell'ambito della giustizia penale rappresenta uno degli ambiti nei quali le potenzialità degli strumenti algoritmici risultano maggiormente evidenti, ma anche uno dei settori nei quali emergono le criticità

più significative. La capacità dei sistemi di IA di analizzare grandi quantità di dati, individuare schemi ricorrenti e formulare previsioni ha favorito il loro progressivo impiego in attività tradizionalmente affidate alla valutazione umana, quali l'analisi del rischio criminale, l'organizzazione delle attività investigative e il supporto ai processi decisionali.⁶

Tali strumenti vengono spesso presentati come mezzi idonei a migliorare l'efficienza del sistema penale, consentendo una più razionale distribuzione delle risorse e una maggiore capacità di individuare situazioni potenzialmente problematiche. In questa prospettiva, l'intelligenza artificiale potrebbe contribuire al perseguimento di finalità di prevenzione della criminalità e di tutela della sicurezza pubblica, fornendo alle autorità competenti informazioni elaborate sulla base di dati empirici.

Tuttavia, proprio la natura predittiva di tali sistemi determina le principali problematiche giuridiche. Gli algoritmi operano infatti attraverso l'elaborazione di dati storici e l'individuazione di correlazioni statistiche, senza poter stabilire con certezza il verificarsi di un determinato evento futuro.⁷ Ne deriva che l'utilizzo di tali strumenti nell'ambito penale comporta inevitabilmente un passaggio dalla valutazione di fatti concretamente accaduti alla formulazione di giudizi probabilistici relativi a comportamenti futuri.

Questa trasformazione assume particolare rilevanza alla luce dei principi fondamentali riconosciuti dal diritto europeo e internazionale dei diritti umani. Un sistema penale fondato esclusivamente su valutazioni predittive rischierebbe infatti di incidere sul principio di responsabilità individuale, sulla presunzione di innocenza e sul diritto a essere giudicati sulla base di fatti effettivamente accertati. Va precisato che nessun ordinamento ha finora introdotto un sistema di responsabilità penale sostitutiva fondato

⁵ F. PASQUALE, *The Black Box Society*, cit.; D.K. CITRON - F. PASQUALE, *The Scored Society: Due Process for Automated Predictions*, in *Washington Law Review*, vol. 89, 2014, pp. 1-33, spec. p. 4.

⁶ A.G. FERGUSON, *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*, New York, New York University Press, 2017, cap. 5, pp. 97 ss.

⁷ R. RICHARDSON - J. SCHULTZ - K. CRAWFORD, *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice*, in *New York University Law Review Online*, vol. 94, 2019, pp. 15-55, spec. p. 21.

sulla sola previsione statistica di condotte future, sul modello descritto nella letteratura fantascientifica⁸. Va precisato che nessun ordinamento ha finora introdotto un sistema di responsabilità penale sostitutiva fondato sulla sola previsione statistica di condotte future. Esistono tuttavia fenomeni normativi che si muovono in questa direzione e ne confermano l'attualità: la criminalizzazione di condotte preparatorie nella legislazione antiterrorismo (si pensi, nell'ordinamento italiano, all'art. 270-quinquies c.p.) sposta la soglia della responsabilità penale a monte del fatto lesivo, su basi essenzialmente prognostiche; le misure di prevenzione post-condanna fondate sul rischio di recidiva, quali la *Sicherungsverwahrung* tedesca⁹ incidono sulla libertà personale sulla base di una valutazione statistica di pericolosità futura anche dopo l'espiazione della pena per il fatto già accertato; e la stessa dottrina criminologica ha teorizzato un passaggio da una logica penale "*post-crime*", reattiva rispetto al fatto commesso, a una logica "*pre-crime*", anticipatoria rispetto a condotte non ancora realizzate.¹⁰ Nessuno di questi fenomeni sostituisce l'accertamento del fatto con la mera previsione, ma tutti dimostrano che l'ordinamento non è concettualmente estraneo a fondare conseguenze restrittive della libertà su giudizi prognostici, il che rende ancora più urgente circoscrivere i limiti entro cui tali giudizi possono legittimamente operare. Tuttavia, gli strumenti predittivi già operativi in diverse giurisdizioni producono un paradosso strutturalmente analogo, pur restando formalmente distinti da tale scenario. Sistemi di *risk assessment* come COMPAS incidono sulla severità della pena o sulla concessione della libertà su cauzione sulla base di un punteggio di

rischio futuro, pur intervenendo solo successivamente all'accertamento di un fatto di reato.¹¹ Strumenti di *predictive policing place-based*, adottati da numerosi dipartimenti di polizia negli Stati Uniti e da almeno quattordici forze di polizia nel Regno Unito, orientano le risorse investigative verso soggetti o aree individuate come "a rischio" sulla base di correlazioni statistiche, determinando un incremento dei controlli che può tradursi, di fatto, in una forma di sorveglianza anticipata.¹² Il caso del *Gangs Matrix* britannico, un database che classifica individui come potenzialmente legati ad attività di gang sulla base di indicatori predittivi, è stato criticato proprio per l'ambiguità del confine tra strumento di prevenzione e anticipazione sostanziale del giudizio di pericolosità.¹³ La distanza formale tra questi strumenti e un sistema di responsabilità predittiva pura non elimina dunque il rischio che la logica preventiva, pur non sostituendo l'accertamento del fatto, ne condizioni progressivamente l'applicazione concreta.

Un ulteriore profilo problematico riguarda la possibile presenza di discriminazioni algoritmiche. Poiché i sistemi di intelligenza artificiale vengono addestrati attraverso dati raccolti nel passato, essi possono incorporare e riprodurre le distorsioni già presenti nei contesti sociali nei quali tali dati sono stati prodotti. In ambito penale, tale rischio assume una particolare gravità, poiché classificazioni algoritmiche apparentemente neutrali potrebbero determinare un trattamento differenziato di determinati gruppi sociali, incidendo sul principio di uguaglianza e sul divieto di discriminazione sanciti, tra gli altri, dall'art. 14 della Convenzione europea dei diritti dell'uomo e dall'art. 21 della Carta dei diritti fondamentali dell'Unione europea.

⁸ Cfr. P.K. DICK, The Minority Report, in *Fantastic Universe*, gennaio 1956; il riferimento è qui meramente descrittivo di uno scenario letterario, privo di corrispondenza normativa attuale.

⁹ Corte EDU, 17 dicembre 2009, M. c. Germania, ric. n. 19359/04; F. DÜNKEL - D. VAN ZYL SMIT, Preventive Detention of Dangerous Offenders Re-Examined: A Comment on Two Decisions of the German Federal Constitutional Court and the European Court of Human Rights, in *German Law Journal*, vol. 5, n. 6, 2004, pp. 619-637.

¹⁰ L. ZEDNER, Pre-Crime and Post-Criminology?, in *Theoretical Criminology*, vol. 11, n. 2, 2007, pp. 261-281.

¹¹ Supreme Court of Wisconsin, *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016), cert. denied, 137 S. Ct. 2290 (2017).

¹² LIBERTY, *Policing by Machine: Predictive Policing and the Threat to Our Rights*, London, gennaio 2019, pp. 1-48, report basato su novanta richieste di accesso civico rivolte alle forze di polizia del Regno Unito; ha rilevato l'uso di strumenti predittivi da almeno 14 forze, incluso il Harm Assessment Risk Tool della Durham Constabulary).

¹³ AMNESTY INTERNATIONAL, *Trapped in the Matrix: Secrecy, Stigma, and Bias in the Met's Gangs Database*, maggio 2018, p. 2 (40% degli individui schedati privi di precedenti per reati violenti negli ultimi due anni; database dismesso dalla Metropolitan Police nel febbraio 2024).

A ciò si aggiunge la questione della trasparenza degli algoritmi. Molti sistemi di intelligenza artificiale, soprattutto quelli basati su tecniche avanzate di apprendimento automatico, presentano un elevato grado di complessità tecnica che rende difficile comprendere il percorso attraverso il quale viene raggiunto un determinato risultato. Tale caratteristica, definita comunemente come problema della "scatola nera" (*black box*), assume particolare rilevanza quando l'output algoritmico incide su decisioni suscettibili di limitare diritti fondamentali.¹⁴ In tali circostanze, la possibilità di comprendere e contestare il funzionamento del sistema diventa una componente essenziale del diritto a un processo equo, tutelato dall'art. 6 della Convenzione europea dei diritti dell'uomo e dall'art. 47 della Carta dei diritti fondamentali dell'Unione europea.

L'impiego dell'intelligenza artificiale nella giustizia penale richiede quindi un approccio regolatorio capace di bilanciare le esigenze di innovazione e sicurezza con la necessità di preservare le garanzie fondamentali dell'individuo. In questa prospettiva, il ricorso a strumenti algoritmici non può tradursi in una sostituzione della valutazione umana, ma deve essere accompagnato da adeguati meccanismi di controllo, trasparenza e responsabilizzazione.

La sfida posta dall'intelligenza artificiale non consiste pertanto nel determinare se la tecnologia possa essere utilizzata nel settore penale, bensì nel definire un modello giuridico nel quale l'innovazione rimanga subordinata ai principi fondamentali che caratterizzano gli ordinamenti europei e internazionali.

1.3. LA COSTRUZIONE DI UN QUADRO NORMATIVO EUROPEO E INTERNAZIONALE PER L'INTELLIGENZA ARTIFICIALE

La crescente diffusione dei sistemi di intelligenza artificiale ha evidenziato l'insufficienza degli strumenti normativi tradizionali nel disciplinare fenomeni caratterizzati da un elevato grado di innovazione tecnologica e da un impatto potenzialmente significativo sui diritti fondamentali. La natura transnazionale delle tecnologie algoritmiche rende infatti inadeguata una regolazione esclusivamente nazionale, imponendo l'elaborazione di un quadro giuridico multilivello fondato sulla cooperazione tra ordinamenti e sulla definizione di standard comuni.

In tale contesto, l'ordinamento dell'Unione europea muove anzitutto da due coordinate di diritto primario: l'art. 2 TUE, che annovera lo Stato di diritto tra i valori fondanti dell'Unione, e l'art. 16 TFUE, che eleva la protezione dei dati personali a diritto autonomo. Su tali basi, l'Unione europea ha assunto un ruolo di primo piano nella costruzione di una disciplina organica dell'intelligenza artificiale, perseguendo l'obiettivo di coniugare innovazione tecnologica e tutela dei diritti fondamentali. Il principale risultato di tale percorso è rappresentato dal regolamento (UE) 2024/1689 (*Artificial Intelligence Act*), primo strumento normativo al mondo volto a disciplinare in maniera organica lo sviluppo, l'immissione sul mercato e l'utilizzo dei sistemi di intelligenza artificiale.¹⁵

L'AI Act adotta un approccio basato sul rischio (*risk-based approach*), distinguendo i sistemi di IA in funzione del livello di rischio che essi presentano per la sicurezza e per i diritti fondamentali delle persone. In particolare, il regolamento individua una categoria di sistemi considerati a rischio inaccettabile, il cui utilizzo è, in linea di principio, vietato, accanto ai sistemi ad alto rischio, ammessi soltanto nel rispetto di rigorosi requisiti di trasparenza, documentazione, supervisione umana, gestione del rischio e qualità dei dati utilizzati per il loro addestramento.

¹⁴ F. PASQUALE, *The Black Box Society*, cit.; J.A. KROLL et al., *Accountable Algorithms*, in *University of Pennsylvania Law Review*, vol. 165, 2017, pp. 633-705.

¹⁵ Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024 che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n.

300/2008, (UE) n. 167/2013, (UE) n. 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull'intelligenza artificiale), in GUUE L, 2024/1689, 12 luglio 2024.

Va peraltro segnalato che il pacchetto di semplificazione noto come Digital Omnibus on AI,¹⁶ sul quale Parlamento europeo e Consiglio hanno raggiunto un'intesa politica il 7 maggio 2026 e che è stato successivamente adottato in via definitiva dal Parlamento europeo il 16 giugno 2026 e dal Consiglio dell'Unione europea il 29 giugno 2026, differisce l'applicazione degli obblighi previsti per i sistemi ad alto rischio dell'allegato III – categoria che comprende diversi strumenti rilevanti ai fini della giustizia penale, quali gli strumenti di valutazione del rischio e i sistemi impiegati dalle autorità di contrasto, dal 2 agosto 2026 al 2 dicembre 2027, mentre restano fermi, sin dalla data originaria, i divieti previsti dall'art. 5. Un simile differimento, per quanto motivato dall'esigenza di attendere la disponibilità delle norme tecniche armonizzate, espone al rischio che la semplificazione normativa si traduca, nei fatti, in un rinvio delle tutele proprio nei settori in cui l'incidenza dell'intelligenza artificiale sui diritti fondamentali risulta più immediata.¹⁷

Tra le applicazioni che assumono particolare rilievo nel settore della giustizia penale rientrano i sistemi di identificazione biometrica remota, gli strumenti di valutazione del rischio e alcune forme di analisi predittiva, rispetto ai quali il legislatore europeo ha previsto specifiche limitazioni, nella consapevolezza della loro potenziale incidenza sui diritti fondamentali e sul corretto funzionamento dello Stato di diritto. In particolare, l'art. 5, par. 1, lett. d), dell'AI Act vieta l'uso di sistemi che valutino o predicano il rischio che una persona fisica commetta un reato sulla base della sola profilazione o della valutazione di tratti e caratteristiche di personalità, salvo che la valutazione supporti un accertamento umano fondato su fatti oggettivi e

verificabili direttamente collegati a un'attività criminosa; l'art. 5, par. 1, lett. h), vieta inoltre, salvo tassative eccezioni, l'identificazione biometrica remota «in tempo reale» in spazi accessibili al pubblico a fini di contrasto. Gli strumenti di valutazione del rischio di recidiva e i sistemi di analisi predittiva impiegati dalle autorità di contrasto rientrano invece, in linea generale, tra i sistemi ad alto rischio elencati nell'allegato III, punto 6, dell'AI Act, e sono pertanto assoggettati ai relativi obblighi di gestione del rischio, documentazione tecnica e sorveglianza umana.¹⁸

L'AI Act non costituisce, tuttavia, un intervento normativo isolato, ma si inserisce in un sistema più ampio di tutela multilivello dei diritti fondamentali.

Un ruolo centrale continua a essere svolto dalla Convenzione europea dei diritti dell'uomo,¹⁹ la quale, pur essendo stata adottata in un contesto storico precedente allo sviluppo dell'intelligenza artificiale, offre un insieme di garanzie suscettibili di trovare piena applicazione anche rispetto ai sistemi algoritmici. In particolare, il diritto al rispetto della vita privata (art. 8 CEDU), il diritto a un equo processo (art. 6 CEDU) e il divieto di discriminazione (art. 14 CEDU) costituiscono parametri essenziali per valutare la compatibilità dell'impiego dell'intelligenza artificiale con gli obblighi convenzionali gravanti sugli Stati.

La centralità della Convenzione europea dei diritti dell'uomo è stata progressivamente confermata anche dalla giurisprudenza della Corte di Strasburgo, che, pur non essendosi ancora pronunciata specificamente sulla giustizia predittiva, ha elaborato principi destinati a incidere sull'utilizzo dell'intelligenza artificiale da parte delle

¹⁶ Commissione europea, Proposta di Regolamento del Parlamento europeo e del Consiglio che modifica i regolamenti (UE) 2024/1689 e (UE) 2018/1139 per quanto riguarda la semplificazione dell'attuazione di regole armonizzate sull'intelligenza artificiale (Omnibus digitale sull'IA), COM(2025) 836 definitivo, 19 novembre 2025.

¹⁷ Sull'iter del c.d. Digital Omnibus on AI, che differisce il calendario di applicazione del regolamento (UE) 2024/1689 per i sistemi ad alto rischio: Consiglio dell'Unione europea, comunicato stampa «*Artificial Intelligence: Council and Parliament agree to simplify and streamline rules*», 7 maggio 2026 (intesa politica provvisoria); Parlamento europeo, approvazione in

seduta plenaria, 16 giugno 2026; Consiglio dell'Unione europea, comunicato stampa «*Artificial Intelligence: Council gives final green light to simplify and streamline rules*», 29 giugno 2026 (adozione definitiva). Alla data del presente contributo l'atto risulta adottato dai co-legislatori ma non ancora pubblicato in Gazzetta ufficiale dell'Unione europea; l'entrata in vigore, tre giorni dopo la pubblicazione, è attesa entro il 2 agosto 2026.

¹⁸ Regolamento (UE) 2024/1689 (AI Act), art. 5, par. 1, lett. d) e h), e allegato III, punto 6.

¹⁹ Convenzione europea per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali, Roma, 4 novembre 1950.

autorità pubbliche.²⁰ In particolare, nelle sentenze *S. and Marper c. Regno Unito* (2008) e *Big Brother Watch e altri c. Regno Unito* (2021), la Corte ha ribadito che la raccolta e il trattamento di dati personali da parte dello Stato devono rispettare i principi di necessità, proporzionalità e prevedibilità, richiedendo garanzie effettive contro il rischio di arbitrarietà.²¹ Più recentemente, nella sentenza *Glukhin c. Russia* (2023), relativa all'utilizzo di sistemi di riconoscimento facciale nell'ambito di attività di polizia, la Corte ha evidenziato come l'impiego di tecnologie biometriche possa determinare un'ingerenza particolarmente intensa nella vita privata, imponendo agli Stati un rigoroso bilanciamento tra esigenze di sicurezza pubblica e tutela dei diritti fondamentali.

Analogamente, nell'ordinamento dell'Unione europea, la Carta dei diritti fondamentali²² assume una funzione di riferimento imprescindibile. Le garanzie riconosciute dagli artt. 7 e 8 in materia di vita privata e protezione dei dati personali, dall'art. 21 in tema di non discriminazione, dall'art. 47 sul diritto a un ricorso effettivo e a un giudice imparziale e dall'art. 48 relativo alla presunzione di innocenza delineano un quadro di tutela che deve orientare tanto l'elaborazione quanto l'utilizzo dei

sistemi di intelligenza artificiale da parte delle autorità pubbliche.

A tali strumenti si affianca il regolamento (UE) 2016/679 (General Data Protection Regulation – GDPR),²³ il quale disciplina il trattamento dei dati personali introducendo principi di liceità, correttezza, trasparenza, minimizzazione dei dati e limitazione delle finalità. Particolare rilievo assume inoltre l'art. 22 del GDPR, che riconosce all'interessato specifiche garanzie rispetto alle decisioni fondate unicamente su trattamenti automatizzati, evidenziando la centralità dell'intervento umano nei processi decisionali suscettibili di produrre effetti giuridici significativi. Nel settore penale, tale garanzia trova il proprio corrispondente nell'art. 11 della direttiva (UE) 2016/680, che disciplina in termini pressoché speculari le decisioni individuali automatizzate, inclusa la profilazione, adottate dalle autorità di polizia e giudiziarie, e costituisce pertanto la disposizione più direttamente rilevante ai fini della presente analisi.²⁴

Anche la Corte di giustizia dell'Unione europea ha progressivamente rafforzato la tutela dei diritti fondamentali nell'ambito del trattamento automatizzato dei dati personali.²⁵ Sebbene le principali pronunce non riguardino direttamente

²⁰ Corte Europea dei Diritti Umani, *S. and Marper c. Regno Unito* [GC], ricc. nn. 30562/04 e 30566/04, sentenza del 4 dicembre 2008, § 125, ove la Corte afferma che la conservazione indiscriminata dei profili genetici di soggetti non condannati «*could not be regarded as necessary in a democratic society*»; Corte Europea dei Diritti Umani, *Big Brother Watch e altri c. Regno Unito* [GC], ricc. nn. 58170/13, 62322/14 e 24960/15, sentenza del 25 maggio 2021, §§ 347-361 e 425, sulle otto garanzie «*end-to-end*» richieste per la compatibilità con l'art. 8 CEDU dei regimi di intercettazione massiva; Corte Europea dei Diritti Umani, *Glukhin c. Russia*, ric. n. 11519/20, sentenza del 4 luglio 2023, §§ 90-91, ove l'uso del riconoscimento facciale in un contesto di protesta pacifica è definito «*incompatible with the ideals and values of a democratic society*».

²¹ F. FABBRINI, *Fundamental Rights in Europe: Challenges and Transformations in Comparative Perspective*, Oxford, Oxford University Press, 2014; K. DZEHTSIAROU, *European Consensus and the Legitimacy of the European Court of Human Rights*, Cambridge, Cambridge University Press, 2015.

²² Carta dei diritti fondamentali dell'Unione europea, in GUUE C 202, 7 giugno 2016, pp. 389-405.

²³ Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016 (General Data Protection Regulation).

²⁴ Direttiva (UE) 2016/680 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativa alla protezione delle persone fisiche con riguardo al trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, nonché alla libera circolazione di tali dati, art. 11.

²⁵ Corte di giustizia dell'Unione europea, *Digital Rights Ireland Ltd*, cause riunite C-293/12 e C-594/12, sentenza dell'8 aprile 2014, §§ 65-69, ove la Corte rileva che la direttiva 2006/24/CE comporta un'ingerenza nei diritti sanciti dagli artt. 7 e 8 della Carta che «risulta essere di vasta portata e deve essere considerata particolarmente grave», priva di regole chiare e precise idonee a limitarla a quanto strettamente necessario; Corte di giustizia dell'Unione europea, *La Quadrature du Net e a.*, cause riunite C-511/18, C-512/18 e C-520/18, sentenza del 6 ottobre 2020, §§ 168-179, sui limiti alla conservazione generalizzata e indiscriminata dei dati di traffico e ubicazione ai fini di contrasto della criminalità grave; Corte di giustizia dell'Unione europea, *Data Protection Commissioner c. Facebook Ireland e Maximilian Schrems* (Schrems II), causa C-311/18, sentenza del 16 luglio 2020, §§ 174-185, sul livello di protezione sostanzialmente equivalente richiesto in caso di trasferimento di dati personali verso Stati terzi.

l'intelligenza artificiale applicata alla giustizia penale, decisioni quali *Digital Rights Ireland* (C-293/12 e C-594/12), *La Quadrature du Net* (C-511/18, C-512/18 e C-520/18) e *Schrems II* (C-311/18) hanno affermato che l'impiego di tecnologie idonee a incidere sulla vita privata deve essere sottoposto a rigorosi requisiti di proporzionalità, necessità e controllo giurisdizionale.²⁶ Tali principi costituiscono oggi un parametro interpretativo imprescindibile anche per l'utilizzo di sistemi algoritmici nel settore della giustizia penale.

A tale giurisprudenza generale si affianca un filone più prossimo al tema della giustizia predittiva. Nella sentenza *Ligue des droits humains* (causa C-817/19, 2022), la Corte si è pronunciata sulla direttiva PNR, chiarendo che lo screening algoritmico automatizzato dei dati dei passeggeri aerei, ove non accompagnato da adeguate garanzie, rischia di produrre un numero significativo di falsi positivi e di alimentare discriminazioni fondate su criteri predittivi non sufficientemente verificabili. Nelle sentenze *SCHUEA Holding* (causa C-634/21, 2023) e *Dun & Bradstreet Austria* (causa C-203/22, 2025), rese in materia di scoring creditizio, la Corte ha inoltre precisato la portata dell'art. 22 GDPR, qualificando come decisione automatizzata anche l'attività di chi elabora il punteggio predittivo destinato a orientare la decisione di un terzo, e ha imposto al titolare del trattamento di fornire una spiegazione concisa e comprensibile della logica sottostante, non superabile con il mero richiamo al segreto industriale. Sebbene rese in ambiti diversi dalla giustizia penale, tali pronunce offrono

coordinate interpretative direttamente trasponibili alla valutazione algoritmica del rischio in sede predittiva.²⁷

Sul piano internazionale, un ulteriore sviluppo è rappresentato dalla Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale, i diritti umani, la democrazia e lo Stato di diritto, aperta alla firma nel 2024.²⁸ Essa costituisce il primo trattato internazionale giuridicamente vincolante dedicato all'intelligenza artificiale e pone al centro della disciplina la protezione dei diritti umani, la salvaguardia delle istituzioni democratiche e il rispetto dello Stato di diritto.²⁹ Pur lasciando agli Stati un margine di discrezionalità nell'attuazione delle relative disposizioni, la Convenzione conferma la progressiva affermazione di un modello regolatorio fondato sulla centralità della persona e sulla necessità che ogni impiego dell'intelligenza artificiale sia accompagnato da adeguate garanzie di trasparenza, responsabilità e controllo umano.

L'evoluzione del quadro normativo europeo e internazionale dimostra, dunque, come la regolazione dell'intelligenza artificiale non persegua l'obiettivo di ostacolare l'innovazione tecnologica, bensì quello di assicurare che il suo sviluppo e il suo utilizzo avvengano nel rispetto dei valori fondamentali che caratterizzano gli ordinamenti democratici. In tale prospettiva, il settore della giustizia penale rappresenta uno degli ambiti nei quali il corretto bilanciamento tra esigenze di sicurezza, efficienza amministrativa e tutela dei diritti fondamentali assume la massima rilevanza.

²⁶ P. CRAIG - G. DE BÚRCA, *EU Law: Text, Cases, and Materials*, 8ª ed., Oxford, Oxford University Press, 2024; S. PEERS et al. (a cura di), *The EU Charter of Fundamental Rights: A Commentary*, 2ª ed., Oxford, Hart Publishing, 2021.

²⁷ Corte di giustizia dell'Unione europea, *Ligue des droits humains*, causa C-817/19, sentenza del 21 giugno 2022, §§ 197 e 203-206, ove la Corte precisa che il riesame individuale non automatizzato deve tenere al minimo il numero di persone identificate erroneamente, posto che il trattamento algoritmico dei dati PNR produce un «*fairly substantial number of false positives*»; Corte di giustizia dell'Unione europea, *SCHUEA Holding e a. (Scoring)*, causa C-634/21, sentenza del 7 dicembre 2023, § 57, secondo cui costituisce decisione automatizzata ai sensi dell'art. 22 GDPR anche l'elaborazione, da parte di un terzo, di un valore di probabilità sul quale il

destinatario si fondi in maniera determinante; Corte di giustizia dell'Unione europea, *CK c. Magistrat der Stadt Wien e Dun & Bradstreet Austria GmbH*, causa C-203/22, sentenza del 27 febbraio 2025, §§ 55-58 e 66, che impone al titolare del trattamento di spiegare «*in a concise, transparent, comprehensible and easily accessible manner*» la logica dell'elaborazione automatizzata, non superabile con il mero richiamo al segreto industriale.

²⁸ Consiglio d'Europa, *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, CETS No. 225, Vilnius, 2024.

²⁹ Consiglio d'Europa, *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, Vilnius, 17 maggio 2024.

2. GIUSTIZIA PREDITTIVA E STRUMENTI ALGORITMICI NELLA GIUSTIZIA PENALE: APPLICAZIONI E CRITICITÀ

La progressiva integrazione dell'intelligenza artificiale nei sistemi di giustizia penale ha favorito lo sviluppo della cosiddetta giustizia predittiva, ossia un modello decisionale fondato sull'impiego di strumenti algoritmici capaci di elaborare dati e formulare valutazioni probabilistiche relative alla commissione di reati, alla valutazione del rischio individuale o all'individuazione di aree e contesti considerati maggiormente esposti a fenomeni criminali.³⁰

La giustizia predittiva rappresenta una delle applicazioni più controverse dell'intelligenza artificiale nel settore pubblico, poiché modifica il tradizionale rapporto tra decisione giuridica e accertamento dei fatti.³¹ Mentre il modello classico della giustizia penale si fonda sulla ricostruzione di una condotta già realizzata e sulla successiva attribuzione di una responsabilità individuale, gli strumenti predittivi operano prevalentemente attraverso l'analisi di dati e la formulazione di ipotesi relative a eventi futuri.

Tale caratteristica determina una trasformazione significativa della funzione preventiva dello Stato, poiché sposta l'attenzione dall'accertamento della violazione alla previsione del rischio. Proprio questa anticipazione dell'intervento pubblico costituisce il principale elemento di interesse, ma anche la fonte delle maggiori criticità rispetto ai principi europei e internazionali di tutela dei diritti fondamentali.

2.1. ORIGINI ED EVOLUZIONE DELLA GIUSTIZIA PREDITTIVA

Le origini della giustizia predittiva risalgono allo sviluppo dei primi sistemi di analisi

quantitativa applicati al fenomeno criminale e alla progressiva diffusione delle tecnologie informatiche nell'ambito delle attività investigative. L'idea alla base di tali strumenti consiste nella possibilità di individuare schemi ricorrenti all'interno di grandi quantità di dati, al fine di prevedere la probabilità che determinati eventi possano verificarsi.

Con il progressivo sviluppo dell'apprendimento automatico (*machine learning*),³² tali strumenti hanno acquisito una capacità analitica sempre maggiore, consentendo di elaborare informazioni provenienti da fonti eterogenee e di individuare correlazioni non immediatamente percepibili attraverso una valutazione esclusivamente umana.

L'interesse verso la giustizia predittiva deriva principalmente dalla promessa di una maggiore efficacia preventiva. Se correttamente utilizzati, gli strumenti algoritmici potrebbero infatti contribuire a orientare le risorse investigative, migliorare la gestione degli apparati giudiziari e favorire una risposta più tempestiva rispetto a determinati fenomeni criminali.

Tuttavia, la logica predittiva presenta una tensione strutturale con i principi fondamentali del diritto penale contemporaneo. Gli algoritmi non stabiliscono infatti una verità giuridica, ma elaborano probabilità sulla base di dati storici.³³ Ne consegue che il risultato prodotto dal sistema non rappresenta un accertamento, bensì una valutazione di rischio, la cui utilizzazione in ambito penale deve essere sottoposta a rigorose garanzie.

2.2. ORIGINI ED EVOLUZIONE DELLA GIUSTIZIA PREDITTIVA

2.2.1. PREDICTIVE POLICING

Una delle applicazioni più note della giustizia predittiva è rappresentata dal *predictive policing*,³⁴ ossia dall'utilizzo di sistemi algoritmici finalizzati a individuare aree geografiche, periodi temporali o

³⁰ A.G. FERGUSON, *The Rise of Big Data Policing*, cit.; H. SURDEN, *Artificial Intelligence and Law: An Overview*, cit., pp. 1305 ss.

³¹ A.G. FERGUSON, *The Rise of Big Data Policing*, cit.; U. BECK, *Risk Society: Towards a New Modernity*, London, Sage, 1992.

³² S. RUSSELL - P. NORVIG, *Artificial Intelligence: A Modern Approach*, cit.

³³ M. HILDEBRANDT, *Smart Technologies and the End(s) of Law*, cit.; F. PASQUALE, *The Black Box Society*, cit.

³⁴ A.G. FERGUSON, *The Rise of Big Data Policing*, cit.

contesti sociali considerati maggiormente esposti alla commissione di reati.

Attraverso l'analisi di dati relativi alla criminalità pregressa, tali strumenti cercano di individuare schemi ricorrenti e fornire indicazioni operative alle forze dell'ordine. L'obiettivo dichiarato è quello di consentire una distribuzione più efficiente delle risorse e una prevenzione più mirata degli episodi criminosi.

Tuttavia, il funzionamento di tali sistemi presenta rilevanti problematiche sotto il profilo dei diritti fondamentali. Poiché gli algoritmi si basano prevalentemente su dati storici,³⁵ essi rischiano di riprodurre e amplificare eventuali distorsioni già presenti nella società. Un territorio sottoposto in passato a maggiori controlli potrebbe infatti produrre un numero maggiore di dati relativi alla criminalità, determinando un rafforzamento artificiale della previsione iniziale.

Tale fenomeno evidenzia il rischio di una discriminazione algoritmica indiretta,³⁶ in contrasto con il principio di uguaglianza e con il divieto di discriminazione sancito dalle principali fonti europee e internazionali di tutela dei diritti umani.

2.2.2. RISK ASSESSMENT TOOLS

Un'altra applicazione particolarmente significativa riguarda gli strumenti di valutazione del rischio (*risk assessment tools*),³⁷ utilizzati in alcuni ordinamenti per stimare la probabilità di recidiva di un individuo o supportare decisioni relative alla fase cautelare, alla determinazione della pena o all'accesso a misure alternative.³⁸

Tali strumenti sollevano questioni particolarmente delicate poiché possono incidere direttamente sulla posizione giuridica della persona, influenzando decisioni suscettibili di limitare la libertà individuale.

Il principale problema riguarda il valore attribuito alla previsione algoritmica. Un sistema fondato su variabili statistiche non valuta infatti la responsabilità per un fatto commesso, ma elabora una probabilità relativa al comportamento futuro dell'individuo. L'utilizzo di tale informazione deve pertanto essere attentamente bilanciato con il principio della responsabilità personale e con il diritto a una decisione giudiziaria adeguatamente motivata.

2.2.3. IDENTIFICAZIONE

BIOMETRICA

E

SORVEGLIANZA

ALGORITMICA

Un ulteriore settore di applicazione dell'intelligenza artificiale nella giustizia penale riguarda i sistemi di identificazione biometrica, in particolare quelli basati sul riconoscimento facciale.³⁹

Tali tecnologie possono offrire strumenti estremamente efficaci per le attività investigative, consentendo l'identificazione rapida di persone coinvolte in determinati eventi. Al contempo, però, esse comportano un significativo impatto sulla protezione della vita privata⁴⁰ e dei dati personali, soprattutto quando utilizzate in spazi pubblici o attraverso sistemi di sorveglianza su larga scala.

La particolare sensibilità dei dati biometrici ha determinato una crescente attenzione da parte del legislatore europeo, che ha introdotto specifiche limitazioni all'utilizzo di tali strumenti proprio in ragione del loro potenziale impatto sui diritti fondamentali.

³⁵ R. RICHARDSON - J. SCHULTZ - K. CRAWFORD, *Dirty Data, Bad Predictions*, cit.

³⁶ S. BAROCAS - A.D. SELBST, *Big Data's Disparate Impact*, in *California Law Review*, vol. 104, 2016, pp. 671-732.

³⁷ R. BERK, *Artificial Intelligence, Predictive Policing, and Risk Assessment for Law Enforcement*, in *Annual Review of Criminology*, vol. 4, 2021, pp. 209-237.

³⁸ J. DRESSSEL - H. FARID, *The Accuracy, Fairness, and Limits of Predicting Recidivism*, in *Science Advances*, vol. 4, n. 1, 2018, e.aao5580.

³⁹ Regolamento (UE) 2024/1689 (AI Act), artt. 5 e ss.; Commissione Europea, *Ethics Guidelines for Trustworthy AI*, 2019.

⁴⁰ GDPR and artt. 7 e 8 Carta di Nizza.

2.3. INTELLIGENZA ARTIFICIALE CRIMINALE? LA RESPONSABILITÀ NELL'ERA DEGLI ALGORITMI

L'impiego dell'intelligenza artificiale nell'ambito della giustizia penale non pone soltanto il problema dell'utilizzo degli algoritmi da parte delle autorità pubbliche, ma solleva anche un interrogativo ulteriore: può l'intelligenza artificiale assumere un ruolo nella produzione o nella determinazione del fenomeno criminale?⁴¹

L'espressione "intelligenza artificiale criminale" non deve essere intesa nel senso di attribuire una soggettività penale autonoma ai sistemi algoritmici, bensì come riferimento alla possibilità che tali tecnologie possano essere utilizzate per finalità illecite oppure possano contribuire indirettamente alla produzione di conseguenze dannose. Il problema riguarda quindi il rapporto tra tecnologia, responsabilità umana e capacità degli ordinamenti giuridici di individuare soggetti responsabili in contesti caratterizzati da processi decisionali complessi e distribuiti.

Nel settore penale tradizionale, l'attribuzione della responsabilità presuppone l'individuazione di una condotta umana riconducibile a un soggetto determinato. L'impiego di sistemi di intelligenza artificiale introduce invece una pluralità di attori coinvolti nel processo tecnologico: sviluppatori, fornitori, utilizzatori istituzionali e soggetti incaricati dell'elaborazione dei dati.⁴² Tale frammentazione rende più complessa l'individuazione della responsabilità nel caso in cui l'utilizzo dell'algoritmo produca conseguenze lesive dei diritti individuali. Sul piano dell'Unione, tale incertezza non trova oggi un pieno rimedio sul versante della responsabilità civile: la proposta di direttiva sulla responsabilità da intelligenza artificiale (AI Liability Directive), pensata proprio per alleggerire l'onere probatorio del danneggiato

in presenza di sistemi opachi, è stata formalmente ritirata dalla Commissione europea nell'ottobre 2025, lasciando la materia priva di un regime uniforme e rimessa, allo stato, ai diritti nazionali.⁴³

La questione assume particolare rilevanza nel contesto della giustizia predittiva. Un sistema che suggerisca un determinato livello di rischio o orienti una decisione pubblica non costituisce un soggetto autonomo di decisione, ma uno strumento utilizzato da un'autorità umana. Tuttavia, il grado di affidamento attribuito al risultato algoritmico può determinare una progressiva delegazione sostanziale del potere decisionale alla macchina.

Proprio per evitare tale rischio, il quadro europeo sull'intelligenza artificiale insiste sul principio della supervisione umana (*human oversight*),⁴⁴ inteso non come un controllo meramente formale, ma come una capacità effettiva dell'operatore di comprendere, valutare e, se necessario, discostarsi dal risultato prodotto dal sistema.

La centralità del controllo umano rappresenta quindi una condizione essenziale affinché l'intelligenza artificiale rimanga uno strumento di supporto e non si trasformi in un meccanismo autonomo di determinazione delle decisioni pubbliche. In assenza di tale garanzia, il rischio sarebbe quello di trasferire progressivamente la responsabilità decisionale verso sistemi algoritmici la cui logica di funzionamento potrebbe risultare difficilmente verificabile.

2.4. IL CASO STATE V. LOOMIS: TRASPARENZA ALGORITMICA E DIRITTO A UN EQUO PROCESSO

Uno dei casi più significativi nell'ambito della giustizia predittiva è rappresentato dalla decisione della Corte Suprema del Wisconsin nel caso *State v. Loomis* del 2016.⁴⁵ La controversia riguardava

⁴¹ M. HILDEBRANDT, *Law for Computer Scientists and Other Folk*, cit.

⁴² OECD, *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449, adottata il 22 maggio 2019 e aggiornata il 3 maggio 2024.

⁴³ Commissione europea, ritiro della proposta di direttiva del Parlamento europeo e del Consiglio relativa all'adeguamento delle norme in materia di responsabilità civile

extracontrattuale all'intelligenza artificiale (AI Liability Directive), GUUE C/2025/5423 del 6 ottobre 2025.

⁴⁴ Artificial Intelligence Act, art. 14; Consiglio d'Europa, *Framework Convention on AI, Human Rights, Democracy and Rule of Law*, 2024.

⁴⁵ Wisconsin Supreme Court, *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016), cert. denied, 137 S. Ct. 2290 (2017), spec. §§ 66, 71, 88 e 98, ove la Corte Suprema del Wisconsin ammette l'uso

l'utilizzo del software COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*),⁴⁶ un sistema algoritmico sviluppato per valutare il rischio di recidiva degli imputati sulla base di una serie di variabili personali e comportamentali.

Nel caso concreto, il giudice aveva utilizzato la valutazione prodotta dal sistema COMPAS nell'ambito della determinazione della pena, contribuendo alla scelta della sanzione da applicare. Il ricorrente contestava tale utilizzo sostenendo, tra gli altri profili, che l'impiego di un algoritmo proprietario e non pienamente conoscibile violasse il diritto a un processo equo, impedendo una contestazione effettiva dei criteri utilizzati nella valutazione della propria posizione.⁴⁷

La Corte Suprema del Wisconsin respinse il ricorso, ritenendo che il sistema potesse essere utilizzato purché considerato esclusivamente come uno strumento di supporto e non come elemento determinante della decisione giudiziaria. La Corte sottolineò inoltre la necessità che il giudice fosse consapevole dei limiti dello strumento e dei possibili margini di errore.⁴⁸

Sebbene la decisione abbia ammesso l'utilizzo di sistemi algoritmici nel processo penale, essa ha evidenziato alcune delle principali criticità che caratterizzano la giustizia predittiva. In particolare, il caso ha posto in luce il problema della trasparenza degli algoritmi e della possibilità per l'imputato di conoscere e contestare gli elementi che incidono sulla valutazione della propria posizione giuridica.

La questione assume particolare rilevanza alla luce del diritto europeo dei diritti fondamentali. Il diritto a un equo processo, riconosciuto dall'art. 6 della Convenzione europea dei diritti dell'uomo e dall'art. 47 della Carta dei diritti fondamentali dell'Unione europea, richiede infatti che le

decisioni giudiziarie siano fondate su elementi conoscibili e verificabili dalle parti interessate.

Il caso *Loomis* evidenzia inoltre il rischio che strumenti apparentemente neutrali possano produrre effetti discriminatori.⁴⁹ Poiché i sistemi di valutazione del rischio sono costruiti sulla base di dati storici, essi possono incorporare correlazioni legate a fattori sociali, economici o demografici, trasformando disuguaglianze preesistenti in criteri apparentemente oggettivi di valutazione.

La vicenda rappresenta quindi un importante punto di riferimento nel dibattito sulla compatibilità tra algoritmi e giustizia penale. Pur non escludendo l'utilizzo dell'intelligenza artificiale, essa conferma la necessità di subordinare ogni impiego di strumenti predittivi al rispetto di condizioni fondamentali: trasparenza, possibilità di controllo, supervisione umana e tutela effettiva dei diritti della persona.

In questa prospettiva, il caso *State v. Loomis* anticipa alcune delle questioni oggi affrontate dal legislatore europeo, dimostrando come la sfida principale non sia soltanto tecnologica, ma propriamente giuridica: assicurare che l'efficienza dell'algoritmo non prevalga sui principi che regolano l'esercizio del potere giudiziario.⁵⁰

3. GIUSTIZIA PREDITTIVA E TUTELA DEI DIRITTI FONDAMENTALI: IL DIFFICILE EQUILIBRIO TRA INNOVAZIONE E STATO DI DIRITTO

L'impiego dell'intelligenza artificiale nella giustizia penale pone una questione di fondo che supera il semplice problema dell'efficienza tecnologica: stabilire quali siano i limiti entro i quali lo Stato possa utilizzare strumenti predittivi senza

del sistema COMPAS a sostegno della decisione di sentencing a condizione che «*the risk scores may not be used as the determinative factor*» nella decisione relativa alla possibilità di supervisionare l'imputato in sicurezza nella comunità.

⁴⁶ A. CHOHLAS-WOOD, *Understanding Risk Assessment Instruments in Criminal Justice*, Washington (DC), Brookings Institution, 19 giugno 2018.

⁴⁷ D.K. CITRON - F. PASQUALE, *The Scored Society*, cit. Supreme Court of Wisconsin, *State v. Loomis*, cit., §§ 88 e 98

⁴⁸ *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016), §§ 88 e 98, ove la Corte precisa che il punteggio COMPAS può essere impiegato solo come uno dei fattori rilevanti ai fini del sentencing e non può costituire il fattore determinante, dovendo il giudice restare libero di discostarsene sulla base di una valutazione individualizzata del caso concreto.

⁴⁹ J. ANGWIN - J. LARSON - S. MATTU - L. KIRCHNER, *Machine Bias*, in *ProPublica*, 23 maggio 2016.

⁵⁰ J.A. KROLL et al., *Accountable Algorithms*, cit.

compromettere i principi fondamentali che governano l'esercizio del potere pubblico.⁵¹

La giustizia predittiva introduce infatti una trasformazione significativa del rapporto tra diritto e decisione. Il modello tradizionale della giustizia penale si fonda sull'accertamento di un fatto passato e sulla conseguente attribuzione di responsabilità a un soggetto determinato. I sistemi predittivi, invece, operano prevalentemente attraverso valutazioni probabilistiche relative a eventi futuri, modificando la prospettiva dell'intervento pubblico dalla repressione di una condotta realizzata alla prevenzione di un rischio ipotizzato.⁵²

Proprio tale trasformazione rende necessario interrogarsi sulla compatibilità della giustizia predittiva con i principi europei e internazionali di tutela dei diritti fondamentali.

3.1. DALLA PREVISIONE DEL RISCHIO ALLA PRE-CRIMINALITÀ: IL PARADIGMA DI MINORITY REPORT

La riflessione sui rischi di una giustizia fondata sulla previsione del comportamento umano trova una rappresentazione particolarmente efficace nell'opera di Philip K. Dick *The Minority Report* (1956), successivamente resa celebre dall'adattamento cinematografico diretto da Steven Spielberg.⁵³

Nella narrazione fantascientifica, il sistema denominato *Precrime* consente di arrestare individui prima della commissione di un omicidio, sulla base delle previsioni formulate da soggetti dotati di capacità precognitive. Il presupposto fondamentale del sistema è l'assoluta affidabilità della previsione: se il crimine può essere anticipato con certezza, l'intervento preventivo appare non soltanto legittimo, ma necessario.

Sebbene tale scenario appartenga alla dimensione della fantascienza, esso rappresenta una metafora particolarmente significativa dei rischi connessi a un uso incontrollato della giustizia predittiva. Il problema non riguarda infatti la possibilità tecnica di elaborare previsioni, bensì il valore giuridico che tali previsioni possono assumere nell'ambito di decisioni capaci di incidere sulla libertà individuale.⁵⁴

Un sistema penale fondato esclusivamente sulla previsione del rischio determinerebbe uno spostamento radicale del momento dell'intervento pubblico: il soggetto non verrebbe più valutato per un comportamento effettivamente realizzato, ma sulla base della probabilità che possa attuare una determinata condotta futura.

Tale prospettiva rischierebbe di entrare in tensione con principi fondamentali del diritto penale moderno e con gli standard internazionali di tutela dei diritti umani, in particolare con la presunzione di innocenza e con il principio secondo cui la responsabilità giuridica deve fondarsi su un fatto concretamente accertato.⁵⁵

Il pericolo maggiore sarebbe quello di trasformare la persona da soggetto titolare di diritti a oggetto di classificazione algoritmica, valutata non per le proprie azioni, ma per la sua appartenenza a una determinata categoria di rischio.⁵⁶

3.2. TRASPARENZA, SPIEGABILITÀ E CONTROLLO UMANO COME GARANZIE DELLO STATO DI DIRITTO

Alla luce di tali criticità, il controllo umano rappresenta una delle principali condizioni di legittimità dell'impiego dell'intelligenza artificiale nella giustizia penale.⁵⁷ Tuttavia, la supervisione

⁵¹ M. HILDEBRANDT, *Smart Technologies and the End(s) of Law*, cit.; L. FLORIDI, *The Ethics of Artificial Intelligence*, cit.

⁵² D.K. CITRON - F. PASQUALE, *The Scored Society*, cit.; H. SURDEN, *Artificial Intelligence and Law: An Overview*, cit.

⁵³ P.K. DICK, *The Minority Report*, in *Fantastic Universe*, gennaio 1956, poi in ID., *The Minority Report and Other*

Classic Stories, New York, Citadel Press, 2002; S. SPIELBERG (regia di), *Minority Report*, DreamWorks Pictures, 2002.

⁵⁴ F. PASQUALE, *The Black Box Society*, cit.

⁵⁵ Art. 6 CEDU; Art. 48 Carta dei diritti fondamentali dell'UE; Art. 49 Carta UE (principio di legalità).

⁵⁶ S. ZUBOFF, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York, PublicAffairs, 2019.

⁵⁷ Regolamento (UE) 2024/1689 (AI Act), art. 14.

umana non può essere intesa come una semplice approvazione formale del risultato prodotto dall'algoritmo.

Un effettivo *human oversight* richiede che il giudice e gli operatori del diritto siano posti nelle condizioni di comprendere il funzionamento generale del sistema, conoscere i dati utilizzati, valutare i possibili margini di errore e individuare eventuali distorsioni derivanti dal modello algoritmico.⁵⁸

Tale esigenza si collega direttamente ai principi di trasparenza e spiegabilità, oggi riconosciuti come elementi essenziali della governance europea dell'intelligenza artificiale.⁵⁹ Un algoritmo che contribuisce a determinare una decisione avente effetti giuridici rilevanti non può operare come una "scatola nera", sottratta a qualsiasi forma di controllo esterno.

La possibilità di contestare il risultato prodotto dal sistema costituisce infatti una componente essenziale del diritto a un processo equo e del principio di effettività della tutela giurisdizionale. La decisione giudiziaria deve rimanere il risultato di una valutazione giuridica autonoma, nella quale l'intelligenza artificiale fornisca elementi di supporto senza determinare automaticamente l'esito del procedimento.⁶⁰

In questa prospettiva, la responsabilità finale deve rimanere attribuita all'autorità umana. L'utilizzo di strumenti algoritmici non può comportare una deresponsabilizzazione del decisore pubblico né un trasferimento della funzione giudiziaria a sistemi tecnologici privi di capacità valutativa autonoma.

3.3. SICUREZZA PUBBLICA E TUTELA DEI DIRITTI FONDAMENTALI: VERSO UN MODELLO EUROPEO DI GOVERNANCE DELL'IA

La principale sfida posta dalla giustizia predittiva consiste dunque nella ricerca di un equilibrio tra due esigenze entrambe meritevoli di tutela: da un lato, l'interesse pubblico alla prevenzione dei reati e alla sicurezza collettiva; dall'altro, la protezione dei diritti fondamentali dell'individuo.⁶¹

In tale prospettiva assume particolare rilievo la recente Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale, i diritti umani, la democrazia e lo Stato di diritto, aperta alla firma nel 2024. Il trattato rappresenta il primo strumento internazionale giuridicamente vincolante dedicato all'intelligenza artificiale e conferma l'emersione di un modello di governance fondato sul primato dei diritti fondamentali. Pur non disciplinando specificamente la giustizia predittiva, la Convenzione impone agli Stati l'obbligo di predisporre misure legislative e amministrative idonee a garantire che i sistemi di intelligenza artificiale siano sviluppati e utilizzati nel rispetto della dignità umana, del principio di uguaglianza, della non discriminazione, dell'accesso alla giustizia e del controllo umano effettivo. In questo senso, la Convenzione si pone in rapporto di complementarità con l'Artificial Intelligence Act, contribuendo alla progressiva costruzione di uno spazio giuridico europeo caratterizzato da standard comuni di tutela dei diritti fondamentali.

Il diritto europeo e internazionale sembrano orientarsi verso un modello nel quale l'innovazione tecnologica non viene rifiutata, ma sottoposta a condizioni rigorose. L'approccio basato sul rischio adottato dall'Unione europea e gli standard elaborati dal Consiglio d'Europa confermano infatti la necessità di subordinare lo sviluppo dell'intelligenza artificiale al rispetto della dignità

⁵⁸ Commissione Europea, Ethics Guidelines for Trustworthy AI, 2019.

⁵⁹ J.A. KROLL et al., Accountable Algorithms, cit.

⁶⁰ M. HILDEBRANDT, Law for Computer Scientists and Other Folk, cit., cap. 11, pp. 314-315.

⁶¹ Consiglio d'Europa, Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, 2024, cap II, pp. 15-16.

umana, della non discriminazione e dello Stato di diritto.⁶²

La questione centrale non è quindi stabilire se l'intelligenza artificiale possa essere utilizzata nella giustizia penale, ma individuare quali limiti debbano essere posti affinché essa non determini una trasformazione del modello di responsabilità giuridica.

L'eventuale passaggio da una giustizia fondata sul fatto a una giustizia fondata sul rischio rappresenterebbe infatti una modifica radicale dei presupposti dello Stato di diritto. Un ordinamento democratico non può punire o limitare la libertà di un individuo sulla base esclusiva di una previsione statistica relativa a ciò che potrebbe compiere in futuro.

La giustizia predittiva può dunque rappresentare un'opportunità, ma soltanto se inserita in un sistema di garanzie capace di preservare la centralità della persona. La tecnologia deve rimanere uno strumento al servizio della decisione giuridica e non trasformarsi nel criterio ultimo attraverso cui determinare la posizione dell'individuo all'interno dell'ordinamento.⁶³

In definitiva, la vera sfida della governance internazionale dell'intelligenza artificiale consiste nel garantire che il progresso tecnologico sia accompagnato da un corrispondente sviluppo delle garanzie giuridiche. Solo attraverso un equilibrio tra innovazione e tutela dei diritti fondamentali sarà possibile evitare che la promessa della prevenzione del crimine si trasformi nel rischio di una giustizia fondata sulla previsione e non sulla responsabilità.

4. CONCLUSIONI

L'analisi svolta ha evidenziato come l'intelligenza artificiale applicata alla giustizia penale rappresenti una delle più complesse sfide poste dalla trasformazione digitale al diritto contemporaneo. Gli strumenti di giustizia predittiva offrono indubbi vantaggi in termini di

capacità analitica, organizzazione delle risorse e prevenzione dei fenomeni criminali; tuttavia, il loro impiego solleva interrogativi fondamentali circa la compatibilità con i principi che costituiscono il fondamento dello Stato di diritto.

La questione centrale non riguarda la contrapposizione tra tecnologia e diritto, bensì la definizione delle condizioni giuridiche entro le quali l'innovazione tecnologica possa essere integrata nei sistemi di giustizia senza compromettere la tutela dei diritti fondamentali. L'esperienza degli strumenti predittivi dimostra infatti che l'efficienza algoritmica non può rappresentare l'unico criterio di valutazione, soprattutto quando tali sistemi incidono su libertà individuali e decisioni aventi rilevanza giuridica. L'argomento sostenuto è che l'intelligenza artificiale possa costituire uno strumento utile per l'amministrazione della giustizia solo a condizione che rimanga subordinata al controllo umano, alla trasparenza e ai principi fondamentali dell'ordinamento europeo e internazionale.

Il quadro normativo mostra tuttavia una duplice velocità. Mentre l'Unione europea ha optato per un modello regolatorio ex ante, fondato sul controllo umano, sulla trasparenza e sulla prevenzione delle discriminazioni, di cui l'Artificial Intelligence Act costituisce l'espressione più compiuta, il Consiglio d'Europa ha tentato di estendere questo approccio oltre i confini dell'Unione attraverso la Convenzione quadro sull'intelligenza artificiale, i diritti umani, la democrazia e lo Stato di diritto, primo trattato internazionale vincolante in materia. Il caso *State v. Loomis*⁶⁴ dimostra però che, al di fuori di questo perimetro, la tutela resta affidata quasi interamente alla giurisprudenza e al giudizio caso per caso, senza un corrispondente intervento del legislatore. La vera frattura internazionale non riguarda dunque la disponibilità di principi condivisi, quanto la loro trasposizione in obblighi giuridici vincolanti al di fuori dello spazio europeo.

⁶² Regolamento (UE) 2024/1689 (Artificial Intelligence Act).

⁶³ UNESCO, Recommendation on the Ethics of Artificial Intelligence, adottata dalla Conferenza generale il 23 novembre 2021.

⁶⁴ Wisconsin Supreme Court, *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016), cert. denied, 137 S. Ct. 2290 (2017).

La giustizia predittiva evidenzia inoltre una questione più profonda: il rischio che la logica della previsione sostituisca progressivamente quella della responsabilità. Un sistema fondato esclusivamente sulla valutazione del rischio potrebbe infatti condurre a una trasformazione del modello penale, spostando l'attenzione dal fatto commesso alla probabilità di una futura condotta illecita. Tale prospettiva, evocata simbolicamente dal paradigma di *Minority Report*, rappresenta il limite oltre il quale la prevenzione rischierebbe di entrare in conflitto con i principi della responsabilità individuale e della presunzione di innocenza.

Per tali ragioni, l'intelligenza artificiale nella giustizia penale deve essere concepita come uno strumento di supporto e non come un sostituto della valutazione giuridica. La decisione giudiziaria deve rimanere il risultato di un procedimento umano fondato sull'interpretazione della legge, sull'esame dei fatti e sulla considerazione delle circostanze concrete del caso.

La sfida futura per il diritto internazionale non consiste soltanto nel completare l'impianto di garanzie già delineato dall'esperienza europea, ma nell'esportarlo verso ordinamenti, come quello statunitense, nei quali la giustizia predittiva si sviluppa in assenza di una cornice normativa organica. Solo colmando questo divario tra modelli regolatori strutturalmente diversi la giustizia predittiva potrà rappresentare un'opportunità di miglioramento dell'amministrazione della giustizia, a livello non solo europeo ma globale, senza trasformarsi in un rischio per i diritti fondamentali che essa stessa è chiamata a proteggere.