Image

CAMMINO DIRITTO
Rivista di informazione giuridica
https://rivista.camminodiritto.it

Articolo Scientifico

GLI SPAZI APPLICATIVI DELL'INTELLIGENZA ARTIFICIALE NELLA GIUSTIZIA PENALE: RIFLESSIONI ALLA LUCE DELL'AI ACT DELL'UE E DEL BLUEPRINT STATUNITENSE

Con l'entrata in vigore del Regolamento UE 2024/1689 (AI Act), l'intelligenza artificiale ha ottenuto una prima regolamentazione formale anche nel settore della giustizia penale. Attraverso il confronto con il Blueprint for an AI Bill of Rights 2022 degli Stati Uniti, l'articolo si propone di evidenziare l'importanza di una lettura integrata dell'AI Act con il Regolamento UE 2016/679 (GDPR), al fine di garantire un giusto equilibrio tra diritto alla privacy ed esigenze securitarie. In questa prospettiva, viene esaminata l'ammissibilità dell'adozione di risk assessment tools, di strumenti di identificazione biometrica e di algoritmi a supporto dell'attività giurisdizionale dell'organo giudicante, evidenziando luci e ombre della disciplina comunitaria.

di **Mattia Giangreco**
IUS/16 - DIRITTO PROCESSUALE PENALE
Estratto dal n. 3/2025 - ISSN 2532-9871

Direttore responsabile
Alessio Giaquinto

Pubblicato, Lunedì 10 Marzo 2025

CamminoDiritto.it
ISSN 2532-9871

Abstract ENG

With the EU Regulation 2024/1689 (AI Act), artificial intelligence has been formally regulated for the first time, including in the field of criminal justice. Through a comparison with the United States' Blueprint for an AI Bill of Rights 2022, this article aims to highlight the importance of an integrated interpretation of the AI Act and the EU Regulation 2016/679 (GDPR) to ensure an appropriate balance between the right to privacy and security needs. In this context, it is examined the admissibility of adopting risk assessment tools, biometric identification systems, and algorithms to support the judicial activity of the judge, highlighting both the strengths and weaknesses of the EU framework.

Sommario: 1. Introduzione; 2. I risk assessment tools: un crocevia per la giustizia penale tra AI Blueprint e AI Act; 2.1. La regolamentazione dell'intelligenza artificiale e i principi dell'AI Blueprint negli Stati Uniti; 2.2. Il risk-based approach e le attività vietate dall'AI Act: tra problemi ermeneutici e possibili letture conformi; 3. L'uso di strumenti di riconoscimento facciale per fini investigativi; 4. Il considerando (61) dell'AI Act e l'apertura alla figura del "giudice robot"; 5. Gli algoritmi di predictive policing basati su big data e gli argini delle normative sulla privacy; 6. Conclusioni.

1. Introduzione

La storia insegna che l'adattamento dell'uomo al progresso tecnologico non è mai stato indolore. Accogliere una nuova tecnologia, infatti, significa rimettere in discussione solidi equilibri per ricercarne di nuovi, assumendosi il rischio di perdere, nel percorso di transizione, prassi consolidate e garanzie. Nell'epoca contemporanea, la frontiera tecnologica da oltrepassare è rappresentata dall'intelligenza artificiale [1], che inevitabilmente sta investendo anche la giustizia penale. L'esigenza di regolamentare le potenzialità e confinare i rischi dell'intelligenza artificiale ha spinto l'Unione Europea ad adottare il 13 giugno 2024 il Regolamento (UE) 2024/1689 (AI Act) [2], con lo scopo di introdurre regole normative armonizzate. Nonostante i propositi del legislatore comunitario, tuttavia, il testo normativo – particolarmente ampio ed articolato – soffre di eccessive generalizzazioni, dovute alla sua destinazione multidisciplinare. Nel campo della giustizia penale, così, si rischia di limitare o indulgere troppo sulle concrete applicazioni dell'intelligenza artificiale. Al fine di comprendere la concreta portata delle disposizioni dell'AI Act, appare utile lo studio del Blueprint for an AI Bill of Rights 2022 [3] statunitense, che enuclea in un documento di soft law importanti principi regolatori [4]. Gli Stati Uniti, infatti, hanno già da tempo adottato algoritmi predittivi sia per la prevenzione dei reati (predictive policing), sia per la valutazione del rischio di recidiva

CamminoDiritto.it
ISSN 2532-9871

(risk assessment tools) [5]. Il confronto tra i due testi normativi, inoltre, offre molteplici spunti di riflessione, perché nelle scelte sulla gestione dell'intelligenza artificiale si manifesta una differente visione sul futuro della giustizia penale: più prudente, nell'Unione Europea, e più efficiente, negli Stati Uniti.

2. I risk assessment tools: un crocevia per la giustizia penale tra AI Blueprint e AI Act

2.1. La regolamentazione dell'intelligenza artificiale e i principi dell'AI Blueprint negli Stati Uniti

Nell'era analogica, i dati personali erano prevalentemente contenuti in documenti e atti cartacei, la cui circolazione era fortemente limitata. Con l'avvento dell'era digitale e dell'intelligenza artificiale, i dati personali hanno assunto una maggiore capacità di diffusione, divenendo una «nuova ricchezza» [6] anche per la giustizia penale, perché potenzialmente in grado di costruire e aggiornare costantemente il profilo identitario di indagati, imputati e detenuti [7]. A beneficiare delle potenzialità della raccolta e profilazione dei dati personali nella giustizia penale sono, anzitutto, le innovative tecniche di predictive policing [8], che, attraverso l'incrocio di dati, elaborano previsioni statistiche sui luoghi in cui potrebbero essere commessi dei reati e sui potenziali autori o vittime, «orientando le attività di polizia alla prevenzione, più che alla sola repressione del crimine» [9]. I risk assessment tools, in particolare, hanno trovato ampio spazio applicativo negli Stati Uniti d'America, sia in ambito cautelare [10], sia nella semplice azione di prevenzione dei reati [11]. Le modalità di "costruzione" di questi algoritmi predittivi si differenzia in due tipologie: strutturata o attuariale [12]. Il primo si basa sulla valutazione delle risposte ad un questionario precedentemente somministrato; il secondo consiste sulla rielaborazione di dati già in possesso dall'amministrazione [13]. Tendenzialmente, si potrebbe affermare che il modello strutturato permette di valutare il rischio del singolo, mentre il modello attuariale ha una rilevanza più generalizzata, sul pericolo della collettività. A questo proposito, occorre notare, in primo luogo, che il modello attuariale può rappresentare una derivazione di quello strutturato, nel senso che, tra i dati in possesso, vi possono anche essere questionari precedentemente compilati da altri soggetti sospettati di avere commesso un reato. Ciò può determinare distorsioni anche nell'ottica della predizione, perché fondata su elementi soggettivi e personologici, quali sono i questionari, e non su dati oggettivi e fattuali, come sono gli eventi delittuosi registrati in un determinato arco temporale e in una specifica area territoriale. In secondo luogo, al di là delle modalità di acquisizione dei dati, la combinazione dei fattori di rischio presi in considerazione può risultare inattendibile, a seconda se sbilanciata sui rischi di tipo statico

CamminoDiritto.it
ISSN 2532-9871

(come il genere), o dinamico (come i precedenti penali). Così, ad esempio, il profilo di rischio fondato prevalentemente su fattori dinamici, potrebbe essere scarsamente affidabile [14], mentre quello basato su fattori statici potrebbe incorporare bias cognitivi potenzialmente discriminatori, come nel noto leading case Loomis, in cui l'appartenenza all'etnia afro-americana di un individuo era stata rilevata dall'algoritmo COMPAS come un fattore di pericolosità sociale [15]. Si è giunti così all'elaborazione di risk assessment tools più attenti a queste problematiche, tra cui: il Privacy Impact Assessment (PIA), il Social Impact Assessment (SIA) e l'Ethical Impact Assessment (EtIA) [16]. Inoltre, nel fare emergere il rischio di bias all'interno degli algoritmi, il precedente Loomis ha affermato che non è possibile adottare una decisione solo sulla base delle valutazioni di un algoritmo predittivo. Viene, quindi, accolta una visione «debole» di intelligenza artificiale, in cui le decisioni devono essere corroborate da «ulteriori e diversi elementi di prova» [17]. In realtà, non si tratta di un principio di diritto del tutto inedito, perché in tema di prova scientifica [18] la Corte Suprema statunitense aveva chiarito, nel caso Daubert, [19] che il giudice deve prima verificare l'affidabilità della teoria e del metodo scientifico proposto dal perito, per potere utilizzare una prova ai fini del proprio convincimento. Anche l'intelligenza artificiale dovrà confrontarsi, prima o poi, con i criteri Daubert, per potere assumere una concreta rilevanza in ambito processuale. Nella consapevolezza di questi limiti, l'AI Blueprint mira a circoscrivere l'applicazione dei sistemi automatizzati a rigorosi principi. Il primo principio riguarda la sicurezza e l'efficacia (safe and effective systems), sia in fase di progettazione sia durante l'utilizzo degli algoritmi, per prevenire anomalie che possono avere impatti negativi sulla società. Un esempio è il problema dei software di polizia predittiva, che possono generare feedback loops, ossia l'innalzamento delle attività di controllo per la prevenzione dei reati solo in aree specifiche, causando episodi di sorveglianza di massa difficilmente accettabili [20]. Per mitigare tali rischi, quindi, non sarebbe sufficiente evitare l'acquisizione illegale di dati, ma sarebbe necessaria l'istituzione di organi indipendenti di controllo sulla sicurezza e l'efficacia di tali algoritmi [21], pur nella consapevolezza che non è semplice monitorarne il funzionamento, a causa del c.d. effetto black box [22]. Il secondo principio riguarda la non discriminazione nei sistemi automatizzati (algorithmic discrimination protections). In particolare, durante la fase di apprendimento dell'algoritmo, è fondamentale evitare l'incorporazione di bias che possano influenzare in modo discriminatorio l'assunzione di decisioni. Un altro principio attiene alla tutela della privacy attraverso il riconoscimento di poteri di controllo sui dati personali che sono funzionali all'apprendimento degli algoritmi (data privacy). A questo proposito, occorre rilevare che, differentemente dal Regolamento generale per la protezione dei dati personali n. 2016/679 (GDPR) dell'Unione Europea [23], gli Stati Uniti non prevedono facoltà e diritti sul titolare del trattamento dei dati, pertanto, la concretizzazione di questo principio avrebbe ripercussioni positive, in termini garantistici, anche esterni al campo dell'intelligenza artificiale [24]. L'AI Blueprint riconosce, inoltre, una sorta di diritto ad essere giudicati da un uomo nel caso in cui dovesse riscontrarsi un errore nell'attività decisionale dei sistemi automatizzati (consideration and fallback) [25]. A rendere effettivo questo principio sembra potersene richiamare un ultimo: quello di

CamminoDiritto.it
ISSN 2532-9871

trasparenza e informazione (notice and explanation), inteso come necessità di informare ogni soggetto interessato circa l'uso dell'intelligenza artificiale per l'assunzione di decisioni giudiziarie.

2.2. Il risk-based approach e le attività vietate dall'AI Act: tra problemi ermeneutici e possibili letture conformi

Con l'AI Act, l'Unione Europea ha scelto di regolamentare in senso "negativo" l'uso dell'intelligenza artificiale, secondo un risk-based approach che individua quattro differenti classi di rischio: inaccettabile, alto rischio, medio rischio, basso rischio [26]. L'intelligenza artificiale è definita dall'art. 3 dell'AI Act come «un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali» [27]. A tal riguardo, occorre notare che la norma offre una definizione ampia di intelligenza artificiale che, tuttavia, è ancorata all'esistenza di «livelli di autonomia variabile», escludendo così i sistemi automatizzati che eseguono comandi senza alcuna forma di rielaborazione. Ad essere assente è, invece, un riferimento espresso alla natura collaborativa dell'intelligenza artificiale, che nella definizione dell'art. 3 dell'AI Act può «generare [...] decisioni». In astratto, dunque, non sembrerebbe possibile escludere l'ammissibilità di un modello di intelligenza artificiale «forte», ossia volta «all'automazione del processo decisionale», in luogo di quella «debole», che collabora senza alcuna autorità decisionale [28]. A mitigare questa possibilità, tuttavia, vi sono i considerando (2) e (27), oltre che l'art. 1 dell'AI Act, i quali precisano che l'intelligenza artificiale deve rispettare i criteri di antropocentrismo e affidabilità [29]. Il modello scelto dal legislatore comunitario sembra volere evitare, anche in prospettiva futura, qualsiasi forma di surrogazione dell'apporto umano, non solo nel settore della giustizia, ma anche in qualsiasi altro ambito. Inoltre, nelle forme ad alto rischio, l'AI Act dispone che l'intelligenza artificiale debba garantire la trasparenza (art. 13 AI Act) la supervisione e la sorveglianza umana (art. 14 AI Act), oltre a una valutazione sui possibili pregiudizi sui diritti fondamentali (art. 27 AI Act) [30]. Si tratta di garanzie inevitabili, la cui effettività è subordinata allo sviluppo di metodologie operative in grado di superare l'effetto black box tipico dei meccanismi di intelligenza predittiva [31]. Nel classificare le possibili applicazioni dell'intelligenza artificiale, l'AI Act ha, anzitutto, individuato le attività vietate. L'art. 5 dell'AI Act, in particolare, prevede nelle lettere d) e h) disposizioni che riguardano espressamente la giustizia penale. La lettera d) riguarda sostanzialmente l'adozione di algoritmi di risk assessment, in quanto vieta l'uso dell'intelligenza artificiale «per effettuare valutazioni del rischio relative a persone fisiche al fine di valutare o prevedere il rischio che una persona fisica commetta un reato,

CamminoDiritto.it
ISSN 2532-9871

unicamente sulla base della profilazione di una persona fisica o della valutazione dei tratti e delle caratteristiche della personalità». La norma, comunque, prevede che non ne è vietato l'uso se «si basa già su fatti oggettivi e verificabili direttamente connessi a un'attività criminosa». Su questo punto, autorevole dottrina ha fatto notare come la norma miri a vietare in senso generalizzato l'uso di strumenti di predictive policing, perché anche i «fatti oggettivi e verificabili», cui fa riferimento il testo normativo, non possono non ricadere su valutazioni soggettive, vietate dalla prima parte della disposizione in esame. Viene, inoltre, rilevato come tale disposizione sia in contraddizione con quanto previsto nell'allegato III, collegato all'art. 6 dell'AI Act, che non vieta l'uso di sistemi automatizzati «per valutare i tratti e le caratteristiche della personalità», considerando queste attività semplicemente come ad alto rischio [32]. A questo proposito, pur nell'apparente contraddittorietà normativa del testo dell'AI Act, è possibile ritrovare una coerenza sistematica. Occorre notare, in particolare, che la norma vieta l'uso di strumenti di risk assessment per valutare o prevedere il rischio che «una persona fisica» commetta un reato, non anche per valutare o prevedere il rischio che un gruppo o una comunità di soggetti commetta un reato. La norma sembra vietare che il risk assessment possa avere una diretta incidenza nell'ottica di prevenzione e valutazione su di una singola persona, ma non ne preclude l'applicazione nell'ottica esclusiva della prevenzione su di un gruppo di persone o su di un'intera comunità. Questa possibile lettura ermeneutica permetterebbe di superare l'apparente contraddizione con l'allegato III, che, nel legittimare i risk assessment tools «per valutare i tratti e le caratteristiche della personalità», sembra riferirsi alla possibile applicazione nei confronti di più soggetti, ai sensi della lett. d) dell'art. 5 AI Act. L'adozione di risk assessment tools, inoltre, potrebbe anche porsi in contrasto con un altro principio fondamentale dell'Unione Europea, che è quello di presunzione di innocenza. Nel valutare il rischio che una persona fisica possa commettere un reato, infatti, l'algoritmo predittivo non dovrebbe mai basarsi, neanche parzialmente, su elementi soggettivi (come i precedenti penali), perché ricadrebbe nel diritto penale d'autore [33]. Si comprende, allora, che l'art. 5 lett. d) dell'AI Act vieta il ricorso a risk assessment tools basati su elementi meramente soggettivi, perché, in assenza, violerebbe il principio di presunzione di innocenza. Coerentemente, l'AI Act dispone nel considerando (42) che le persone fisiche nell'Unione Europea «dovrebbero sempre essere giudicate in base al loro comportamento effettivo [e] non dovrebbero mai essere giudicate sulla base di un comportamento previsto dell'IA basato unicamente sulla profilazione, sui tratti della personalità o su caratteristiche quali la cittadinanza, il luogo di nascita, il luogo di residenza, il numero di figli, il livello di indebitamento o il tipo di automobile, senza che vi sia un ragionevole sospetto che la persona sia coinvolta in un'attività criminosa sulla base di fatti oggettivi verificabili e senza una valutazione umana al riguardo». La stessa disposizione, inoltre, vieta che l'adozione dell'intelligenza artificiale possa essere strumentalizzata per determinare la probabilità che le persone fisiche possano compiere un reato basandosi unicamente sulla «loro profilazione o [...] valutazione dei loro tratti della personalità e delle loro caratteristiche». La norma è molto dettagliata nel prevedere – quasi con una

CamminoDiritto.it
ISSN 2532-9871

tecnica-redazionale di tipo casistico – le condotte non ammissibili, ma lascia perplessi la scelta di adottare il condizionale («dovrebbero»), quasi come se vi fosse uno spazio normativo per l'ammissibilità di simili valutazioni.

3. L'uso di strumenti di riconoscimento facciale per fini investigativi

Una disciplina a parte è, invece, dedicata all'uso di strumenti di identificazione biometrica, che riguarda, in particolare, il riconoscimento facciale [34]. Le norme dell'AI Act distinguono tra identificazione biometrica in tempo reale e identificazione biometrica ex post. A norma dell'art. 5, lett. h) (1) dell'AI Act, emerge anzitutto che l'identificazione biometrica in tempo reale negli spazi pubblici è vietata, salvo il ricorrere di tre eccezioni consistenti nella: ricerca mirata di vittime di rapimento, tratta di esseri umani o sfruttamento sessuale, nonché per persone scomparse; prevenzione di una minaccia specifica, sostanziale e imminente per la vita o la sicurezza o un attacco terroristico; localizzazione o identificazione di una persona sospettata di aver commesso reati gravi elencati nell'Allegato II (ad esempio: terrorismo, crimine organizzato, sfruttamento sessuale di minori) [35]. Nonostante il divieto, l'AI Act prevede tante eccezioni, che consentono un ampio ricorso agli strumenti di identificazione biometrica, purché sussista una legislazione nazionale che ne regolamenti l'uso, a livello procedurale e temporale, e l'autorizzazione di un'autorità giudiziaria o amministrativa indipendente [36]. L'AI Act sembra volere uniformare la disciplina del riconoscimento facciale, ove consentita, a quella statuita in tema di acquisizione dei tabulati e data retention dal GDPR 2016 e dalla giurisprudenza della Corte di giustizia dell'Unione Europea. In particolare, la detenzione dei dati personali deve essere limitata ai parametri di necessità (art. 5 GDPR), e la loro acquisizione deve essere subordinata all'esistenza di un'esplicita previsione normativa da parte dei legislatori nazionali e alla previa autorizzazione da parte di un'autorità giudiziaria o indipendente [37]. Riguardo all'uso investigativo di strumenti di identificazione biometrica ex post, invece, l'AI Act non pone alcun divieto, ma prevede l'obbligo dell'autorizzazione giudiziaria, salvo sia funzionale al riconoscimento di un potenziale sospettato [38]. Il quadro normativo sollecita, anzitutto, la riflessione circa la ratio che anima questa differenza trattamentale. Se si considera, infatti, il valore intrinseco del dato biometrico, l'utilizzazione in tempo reale o ex post non rappresenta, rispettivamente, una minaccia maggiore o minore alla privacy. Sul punto, secondo alcuni, le cautele del legislatore comunitario sono riconducibili alla volontà di tutelare la libertà fondamentali di espressione e di riunione [39]. Purtuttavia, occorre notare che, la semplice attività di riconoscimento biometrico in tempo reale è priva di ripercussioni su questi diritti fondamentali. Ad avviso dello scrivente, invece, i vincoli normativi dell'AI Act sono funzionali a tracciare una linea di confine oltre la quale non è possibile andare, perché si rischia di scivolare verso forme di sorveglianza di massa, che verrebbero così

CamminoDiritto.it
ISSN 2532-9871

normalizzate, con potenziali ripercussioni sul diritto alla privacy [40]. Sulla scia delle attività vietate, infine, si pone l'art. 5 lett. e) dell'AI Act, che vieta l'adozione di sistemi di intelligenza artificiale in grado di realizzare attività di scraping sul web o su riprese di videosorveglianza, benché ex post [41]. Anche in questo caso, è difficile individuare una ratio differente da quella di volere tutelare il diritto di riservatezza. Infatti, nonostante sui dati personali presenti sul web sia possibile esercitare il diritto di cancellazione ex art. 17 GDPR [42], le difficoltà di attuazione concreta del diritto all'oblio non permettono di qualificare tutti i dati presenti sul web come frutto di una scelta consapevole e volontaria della persona interessata. Il divieto di cui all'art. 5 lett. e) dell'AI Act, pertanto, sembra seguire un principio di precauzione, a tutela del diritto alla riservatezza, per non aggravare ulteriormente la difficile attuazione dei principi del GDPR [43].

4. Il considerando (61) dell'AI Act e l'apertura alla figura del "giudice robot"

Nel considerando (61) dell'AI Act, il legislatore comunitario ha collocato l'intelligenza artificiale applicata alla giustizia tra la attività ad alto rischio. In particolare, ha previsto che «al fine di far fronte ai rischi di potenziali distorsioni, errori e opacità» sono collocate tra le attività ad alto rischio i sistemi di intelligenza artificiale destinati a «essere utilizzati da un'autorità giudiziaria o per suo conto per assistere le autorità [...] nelle attività di ricerca e interpretazione dei fatti e del diritto e nell'applicazione della legge a una serie concreta di fatti» [44]. La norma, in primo luogo, evidenzia i rischi legati all'uso dell'intelligenza artificiale per scopi ermeneutici, sottolineando il pericolo di compromettere principi fondamentali come il libero convincimento del giudice, dovuti ad errori o opacità dei sistemi di giustizia predittiva. Gli errori e le opacità dei sistemi di giustizia predittiva, infatti, rappresentano quella componente inestricabile degli algoritmi di intelligenza artificiale, che si può tentare di superare solo adottando algoritmi open-access, facilmente accessibili e trasparenti [45]. In secondo luogo, l'uso dell'intelligenza artificiale rischia di pregiudicare il concetto di oralità che caratterizza il processo penale [46], perché la giustizia predittiva, diversamente dai giudici, si fonda sugli atti del processo. Lo si comprende meglio dalla previsione del considerando (61) nella parte in cui prevede che «l'utilizzo di strumenti di IA può fornire sostegno al potere decisionale dei giudici o all'indipendenza del potere giudiziario, ma non dovrebbe sostituirlo: il processo decisionale finale deve rimanere un'attività a guida umana». L'uso del condizionale, pur se mitigato dalla figura umana del giudice – declassata al ruolo di guida del processo decisionale – sembra minare le fondamenta dei principi del giusto processo, i quali sottintendono che la decisione del giudice si formi in giudizio e non in una fase successiva, sulla base degli atti processuali. A questo proposito, nell'evidenziare l'importanza di un controllo umano che guidi le operazioni degli strumenti di intelligenza artificiale, la dottrina ha anche sottolineato come esistano condizioni «imprescindibili»

CamminoDiritto.it
ISSN 2532-9871

per il loro corretto funzionamento, tra cui: la sottoposizione a processi di peer review; la trasparenza del tasso di errore; la spiegabilità delle formule tecniche in formule giuridiche; la salvaguardia del contraddittorio; l'accettazione da parte dell'organo giudicante in conformità al principio del libero convicimento (art. 192 c.p.p.) [47]. Il considerando (61) prevede, infine, che non devono essere considerati come sistemi di intelligenza artificiale ad alto rischio quei sistemi che sono «destinati ad attività amministrative puramente accessorie, che non incidono sull'effettiva amministrazione della giustizia nei singoli casi, quali l'anonimizzazione o la pseudonimizzazione di decisioni, documenti o dati giudiziari, la comunicazione tra il personale, i compiti amministrativi». È una precisazione che la norma adotta per evitare equivoci, benché già agevolmente deducibile dall'assenza di un contributo sostanziale nel «processo decisionale», che rappresenta l'unico chiaro elemento discretivo che separa le attività ad alto rischio dalle altre. L'attività di anonimizzazione o pseudonomizzazione dei documenti, infatti, rappresenta una mera attività ausiliatrice dei profili strumentali per l'amministrazione della giustizia, e non prevede alcun apporto decisionale o valutativo da parte dell'algoritmo preposto a svolgere questi compiti [48].

5. Gli algoritmi di predictive policing basati su big data e gli argini delle normative sulla privacy

Nel contesto normativo dell'Unione Europea, l'uso dei dati personali per programmare algoritmi applicati alla giustizia predittiva o per svolgere attività di indagine sollevano questioni di natura etica ma, soprattutto, giuridica e tecnica. In particolare, il GDPR rappresenta il principale riferimento normativo, stabilendo principi fondamentali come: la trasparenza (art. 5), il consenso esplicito (art. 7), il diritto all'oblio (art. 17) e l'autoresponsabilizzazione (art. 25) [49]. Sebbene il GDPR trovi applicazione all'ipotesi di trattamento dei dati personali per attività di machine learning, il rischio più significativo è rappresentato dal ricorso a tecniche di profiling [50], ad esempio per i risk assessment tools, in quanto i dati di riferimento mirano a costruire un quadro ben preciso di una persona fisica e non semplicemente di un contesto allargato. Il primo problema da affrontare riguarda la fonte dei big data, utilizzabile per addestrare un algoritmo di intelligenza artificiale. Nell'ambito della giustizia penale, oltre ai dati già in possesso dalle autorità giudiziarie, assumono rilievo le informazioni contenute nei metadati degli internet service providers, che rivelano informazioni sul traffico telefonico e telematico di ogni consumatore-cliente [51]. Queste informazioni, infatti, sono legittimamente oggetto di data retention solo se perseguono finalità di tutela della sicurezza pubblica, come chiarito nel precedente giurisprudenziale Tele2 Sverige AB e Watson [52]. Ci si domanda, se l'addestramento di algoritmi di intelligenza predittiva basata su metadati oggetto di data retention possa considerarsi legittima. Nel caso di algoritmi di predictive policing,

CamminoDiritto.it
ISSN 2532-9871

trattandosi di esigenze di prevenzione dei reati, la legittimità sarebbe teoricamente ammessa, purché, secondo l'interpretazione correttiva supra descritta [53], l'addestramento non sia poi finalizzato a valutare la pericolosità di un singolo individuo. Nel caso di valutazioni giurisdizionali, invece, gli algoritmi predittivi non potrebbero attingere a queste informazioni, non solo perché ricadrebbero nel divieto di cui all'art. 5, lett. d) dell'AI Act [54], ma anche perché le finalità di accertamento della responsabilità penale esulano da quella di prevenzione, ossia l'unica esigenza che legittima una limitazione del diritto di riservatezza nel contesto dell'Unione Europea. Una seconda questione attiene, infine, all'anonimizzazione dei dati personali da utilizzare per l'addestramento dell'algoritmo di intelligenza predittiva, su cui incide la quantità delle informazioni a disposizione. Infatti, tanto maggiori sono le informazioni anonimizzate a disposizione, quanto maggiori sono le possibilità di re-identificare i titolari dei dati personali, perché le reti neurali degli algoritmi hanno raggiunto un'elevata capacità di incrociare i dati [55], riuscendo a cogliere profili affini e analogie anche in elementi che sfuggono all'osservatore umano. Su questo punto, pur nell'estrema ampiezza dell'articolato normativo, l'AI Act non dispone alcunché, avvalorando quell'assenza di coordinamento con la disciplina del GDPR, che sarebbe invece stata indispensabile per evitare gravi violazioni del diritto di privacy. Guardando alla realtà giuridica statunitense, l'uso dei dati personali per finalità investigative non gode di una normativa unitaria, il che significa che non esiste una legge federale uniforme che tutela il diritto alla privacy. A differenza del GDPR, infatti, gli Stati Uniti si affidano a una combinazione di legislazione federale e statale, insieme a norme di rango secondario [56]. L'intervento dell'AI Blueprint rappresenta, dunque, un documento che indirizza unitariamente la disciplina dei sistemi automatizzati, ma costituisce allo stesso tempo l'occasione per rafforzare la tutela del diritto di riservatezza nell'ordinamento statunitense. A supporto di ciò può farsi riferimento al fatto che l'AI Blueprint definisce non semplicemente l'intelligenza artificiale, ma più ampiamente i sistemi automatizzati, come «qualsiasi sistema, software o processo che utilizza il calcolo come parte o intero di un sistema per determinare risultati, prendere o assistere decisioni, informare l'implementazione di politiche, raccogliere dati o osservazioni o interagire con individui e/o comunità» [57]. All'interno di questa definizione, vi si può fare rientrare anche la tecnologia volta alla registrazione e conservazione dei metadati, individuando così una cornice unitaria di soft law anche per la data retention.

6. Conclusioni

Alla base dello sviluppo degli algoritmi di intelligenza artificiale vi è una fase di apprendimento fondamentale, realizzata attraverso i meccanismi di machine learning. Durante questo processo, la fase di identificazione ed elaborazione dei big data

CamminoDiritto.it
ISSN 2532-9871

rappresenta quella più delicata, perché da essa dipende la costruzione del "DNA" dell'algoritmo. Fin da questa prima fase, quindi, devono trovare applicazione i principi sanciti dall'AI Act e, in particolare, il principio di supervisione e sorveglianza umana di cui all'art. 14 [58]. Nel contesto della giustizia penale, l'AI Act prevede che l'intelligenza artificiale possa trovare applicazione sia come strumento di predictive policing, sia come strumento di ausilio per le decisioni giurisdizionali, benché con limiti e dubbi non marginali. Da un lato, la predictive policing non potrebbe affidarsi ai risk assessment tools per valutazioni soggettive, alla stregua dell'esperienza statunitense del caso Loomis [59]. Dall'altro, tuttavia, si muove in modo contraddittorio sull'ammissibilità degli strumenti di intelligenza predittiva per l'assunzione di provvedimenti giurisdizionali, in quanto all'antropocentrismo dell'art. 1 dell'AI Act, fa da contraltare il considerando (61), che ne ammette la possibilità, seppur sotto la guida umana. Quest'ultima, infatti, non sembra rispecchiare appieno quel modello di intelligenza artificiale «debole» [60], che sembra trasparire dalla lettura sistematica delle disposizioni dell'AI Act. Per garantire i diritti fondamentali, il principio di sorveglianza di cui all'art. 14 dell'AI Act deve trovare applicazione anche nella fase esecutiva, attraverso periodiche forme di monitoraggio e controlli, al pari di quanto previsto con l'AI Blueprint statunitense nel principio di «algorithmic discrimination protections», secondo cui la validazione di un algoritmo deve essere preceduta da una valutazione da parte di esperti indipendenti e seguita da reports continui che ne attestano il persistere dei presupposti di validazione [61]. L'adozione di strumenti di predictive policing, inoltre, solleva problematiche in seno alla possibilità di ricorrere a strumenti di monitoraggio dei dati biometrici (in tempo reale o ex post) per finalità securitarie. Su questo profilo, l'AI Act e il GDPR permettono di arginare l'avanzare di tecnologie potenzialmente contrastanti con il diritto di riservatezza, tramite le garanzie della riserva di legge e di giurisdizione [62]. L'AI Blueprint, invece, non prevedendo disposizioni di pari tenore, affida questa tutela ai suoi cinque principi generali: safe and effective systems; algorithmic discrimination protections; data privacy; notice and explanation e human alternatives, consideration and fallback [63]. Infine, l'AI Act non regolamenta espressamente il rischio di re-identificazione dei dati anonimizzati, nel caso in cui confluiscano in big data destinati ad addestrare algoritmi di predictive policing [64]. A questo proposito, sarebbe stato sufficiente integrare il considerando (69) dell'AI Act, che già riconosce «i principi della minimizzazione dei dati e della protezione dei dati fin dalla progettazione e per impostazione predefinita [e] l'anonimizzazione e la cifratura [dei dati]». Almeno per l'Unione Europea, dunque, questo aspetto dovrà essere regolamentato dalle legislazioni nazionali degli Stati membri. In conclusione, i principi dell'AI Act e dell'AI Blueprint, pur differendo per natura giuridica (hard law e soft law) e contenuti, convergono nella necessità di promuovere un uso trasparente e responsabile dell'intelligenza artificiale, soprattutto in settori cruciali come la giustizia penale. La consapevolezza condivisa è che, in assenza di tali garanzie, si rischia di compromettere ciò che l'intelligenza artificiale intende rafforzare: la fiducia nella giustizia e il rispetto dei diritti fondamentali.

CamminoDiritto.it
ISSN 2532-9871

CamminoDiritto.it
ISSN 2532-9871

Note e riferimenti bibliografici

[1] M. BOZZO, La grande storia del computer: dall'abaco all'intelligenza artificiale, Bari, 1996.

[2] Regolamento (UE) 2024/1689 del parlamento dell'Unione Europea e del Consiglio, disponibile all'indirizzo: www.eur-lex.europa.eu.

[3] V. Blueprint for an AI Bill of Rights. Making automated systems work for the American people (da ora in avanti AI Blueprint). Disponibile per la consultazione all'indirizzo: www.whitehouse.gov. Per un commento in rapporto ai diritti fondamentali v. A. CAHANE, D. DROR-SHPOLIANSKY, Human Rights and AI? On the The Blueprint for an AI Bill of Rights and its meaning, in RUNI Law Review Online, 18 maggio 2023.

[4] Occorre notare, tuttavia, che questi principi faticano a trovare concreta applicazione, a causa della non cogenza dell'AI Blueprint. Su questo punto si veda lo studio di D. LAGE, R. PRUITT, J.R. ARNOLD, Who Followed the Blueprint? Analyzing the Responses of US Federal Agencies to the Blueprint for an AI Bill of Rights, in Arxiv, 29 aprile 2024, 3 in cui si evidenzia che, a distanza di due anni dalla sua approvazione, solo il 20% dei documenti ufficiali contiene riferimenti all'AI Blueprint.

[5] Cfr. E. PIETRACOLO, Predictive policing: criticità e prospettive dei sistemi di identificazione dei potenziali criminali, in Sistema penale, 28 settembre 2023.

[6] Cfr. V. RICCIUTO, L'equivoco della privacy. Persona vs dato personale, 2022, Napoli.

[7] Cfr. D. ELLIOTT, E. SOIFER, AI technologies, privacy, and security, in Frontiers in Artificial Intelligence, Vol. 5, 2022, 1 ss..

[8] A. G. FERGUSON, The Rise of Big data Policing. Surveillance, Race, and The Future of Law Enforcement, New York, 2017; ID., Policing Predictive Policing, in Washington Law Review, 2017, vol. 94, no. 5, 1109.

[9] E. PIETRACOLO, Predictive policing: criticità e prospettive dei sistemi di identificazione dei potenziali criminali, cit., 1.

[10] Sia permesso il rinvio a M. GIANGRECO, Dal money bail alla predictive justice: il volto odierno della libertà su cauzione negli Stati Uniti d'America, in La legislazione penale, 2024, n. 3, 362 ss.

[11] Cfr. J.L. VILJOEN, D.M. COCHRANE, M.R. JONNSON, Do risk assessment tools help manage and reduce risk of violence and reoffending? A systematic review, in Law and human behavior, Vol. 42, 2018, n. 3, 181.

[12] Su ci v. S. QUATTROCOLO, Intelligenza artificiale e processo penale: le novità dell'AI Act, in Diritto di difesa, 16 gennaio 2025, 5.

[13] Ivi, 5.

[14] G. ZARA, D.P. FARRINGTON, Criminal Recidivism: explanation, prediction and prevention, Londra, 2016, 155.

[15] Supreme Court of Wisconsin, 31 luglio 2016, State v. Loomis, 881 NW 2d749 (WIS 2016). Cfr. S. QUATTROCOLO, Artificial Intelligence, Computational Modelling and Criminal Proceedings, cit., 156 e ss.

[16] Cfr. D. WIGHT, E. MORDINI, Privacy and Ethical Impact Assessment, in Privacy Impact Assessment, edited by Wright, De Hert, New York, 2012, 397–418.

[17] G. CANZIO, AI Act e processo penale: sfide e opportunità, in Sistema penale, 14 ottobre 2024, 2.

[18] Sulla definizione di prova scientifica v. M. CECCHI, Nozione e definizione di prova scientifica, in La prova scientifica, a cura di Conti, Marandola, Milano, 2023, 31.

[19] Corte Suprema degli Stati Uniti d'America, Daubert v. Merrel Dow Pharmaceuticals, 509 US 579 (1993).

CamminoDiritto.it
ISSN 2532-9871

[20] E.S. BRUNETTE, R.C. FLEMMER, C.L. FLEMMER, A review of artificial intelligence, in International Journal of Wireless Information Networks, 2009, n. 2, 386.

[21] Cfr. E. PIETRACOLO, Predictive policing: criticità e prospettive dei sistemi di identificazione dei potenziali criminali, cit., 30.

[22] Cfr. D. PEDRESCHI, F. GIANNOTTI, R. GUIDOTTI, A. MONREALE, S. RUGGIERI, F. TURINI, Meaningful explanations of black box AI decision systems, in Proceedings of the AAAI conference on artificial intelligence, Vol. 33, 2019, n. 1, 9780.

[23] Regolamento (UE) 2016/679 del parlamento dell'Unione Europea e del Consiglio, disponibile all'indirizzo: www.eur-lex.europa.eu.

[24] Su questo punto, occorre notare l'importante influenza della Corte costituzionale tedesca che, con la pronuncia Volkszählungsurteil del 15 dicembre 1983, sulla legge del censimento del 1983, ha introdotto il principio di autodeterminazione informativa. La pronuncia è disponibile integralmente all'indirizzo: www.bundesverfassungsgericht.de.

[25] Cfr. D. LAGE, R. PRUITT, J.R. ARNOLD, Who Followed the Blueprint? Analyzing the Responses of US Federal Agencies to the Blueprint for an AI Bill of Rights, cit., 3.

[26] Cfr. S. QUATTROCOLO, Intelligenza artificiale e processo penale: le novità dell'AI Act, in Diritto di difesa, 16 gennaio 2025, 4; M. KRETSCHMER, T. KRETSCHMER, A. PEUKERT, C. PEUKERT, The risks of risk-based AI regulation: taking liability seriously, in ArXiv, 2023.

[27] Art. 3 AI Act.

[28] Questa classificazione è proposta da G. CANZIO, AI Act e processo penale: sfide e opportunità, cit., 1.

[29] Considerando (2) dell'AI Act.

[30] Cfr. G. CANZIO, AI Act e processo penale: sfide e opportunità, cit., 5.

[31] Cfr. A.G. FERGUSON, Illuminating Black Data Policing, in Ohio State Journal of Criminal Law, 2018, 15, 503 ss.

[32] Ivi, 7 secondo cui «la contraddizione non potrebbe essere maggiore. Il testo, apparentemente, vieta prima e consente dopo il ricorso al risk assessment, aprendo uno spazio di incertezza assolutamente senza precedenti».

[33] Si veda E. STANIG, Il nuovo diritto penale d'autore, in Scuola Positiva e sistema penale: quale eredità?, a cura di Pittaro, Trieste, 45 ss.

[34] Su cui v. J. DELLA TORRE, Quale spazio per i tools di riconoscimento facciale nella giustizia penale? in Intelligenza artificiale e processo penale, a cura di Di Paolo, Pressacco, Napoli, 2022, 146 ss.; E. SACCHETTO, La prova biometrica, in La prova scientifica, cit., 243 ss.; A. MARANDOLA, Il riconoscimento facciale, in ibidem, 495 ss.

[35] Cfr. M. SIMMLER, G. CANOVA, Facial recognition technology in law enforcement: regulating data analysis of another kind, in Computer Law & Security Review, Vol. 56, 2025, 1 ss.

[36] In questo senso art. 5, lett. h (2) (3) dell'AI Act.

[37] Il riferimento è all'art. 5 GDPR e alle pronunce Corte di giustizia dell'U.E. H.K. v. Prokuratuur, Case C-746/18 e Corte di giustizia dell'U.E., Tele2 Sverige AB e Watson, C-203/15 e C-698/15. Cfr. Cfr. D. MARRANI, Cybersicurezza e tutela della riservatezza dei dati personali: le decisioni Breyer e Tele2 Sverige c. Watson della Corte di giustizia UE, in Il diritto dell'Unione Europea, 2017, 4, 791 ss.; Cfr. E. ANDOLINA, La sentenza della Corte di giustizia UE nel caso H.K. c. Prokuratuur: un punto di non ritorno nella lunga querelle in materia di data retention?, in Processo penale e Giustizia, 2021, n. 5. 1 ss.

[38] Si veda l'articolo 26, (10) dell'AI Act.

CamminoDiritto.it
ISSN 2532-9871

[39] Così S. QUATTROCOLO, Intelligenza artificiale e processo penale: le novità dell'AI Act, cit., 9.

[40] Cfr. T. GUIMARÃES MORAES, E. COSTA ALMEIDA, J.R. LARANJEIRA DE PEREIRA, Smile, you are being identifed! Risks and measures for the use of facial recognition in (semi-)public spaces, in AI and Ethics, 2021, n. 1, 159; J. FULTON, M. KIBBY, Millennials and the normalization of surveillance on Facebook, in Continuum Journal of Media & Cultural Studies, Vol. 31, 2017, n. 2, 189.

[41] Cfr. B. ZHAO, Web scraping, in Encyclopedia of big data, Cham, 2022, 951 ss. Un esempio è il software Clearview in grado di raccogliere immagini e informazioni dal web. Cfr. I.N. REZENDE, Facial recognition in police hands: Assessing the 'Clearview case'from a European perspective, in New Journal of European Criminal Law, Vol. 11, 2020, n. 3, 375.

[42] Sul diritto all'oblio v. R. PARDOLESI, S. BONAVITA, GDPR e diritto alla cancellazione (oblio), in Danno e responsabilità, 2018, n. 3, 269 ss.

[43] Occorre ricordare che, nel contesto statunitense, l'adozione della facial recognition applicata al web scraping ha condotto a scandali mediatici come quello del software Clearview Cfr. M. GIALUZ, Prove fondate sull'intelligenza artificiale e diritti fondamentali, in Diritto di difesa, 15 gennaio 2024, 4.

[44] Considerando 61 del Regolamento (UE) 2024/1689 del Parlamento Europeo e del Consiglio.

[45] Cfr. S. QUATTROCOLO, Prova e intelligenza artificiale, in La prova scientifica, cit., 467 ss. In giurisprudenza, si veda l'interessante pronuncia della Corte di Giustizia UE (Grande Camera), 21 giugno 2022, n. 817, in Dejure, §195 secondo cui «tenuto conto dell'opacità che caratterizza il funzionamento delle tecnologie di intelligenza artificiale, può risultare impossibile comprendere la ragione per la quale un dato programma sia arrivato ad un riscontro positivo. In tali circostanze, l'uso di siffatte tecnologie potrebbe privare gli interessati anche del loro diritto a un ricorso giurisdizionale effettivo sancito dall'articolo 47 della Carta [dei diritti fondamentali dell'Unione Europea]».

[46] Cfr. C. VALENTINI, I principi europei sul processo penale, a cura di Gaito, Roma, 2016, 455.

[47] Così G. UBERTIS, Intelligenza artificiale, giustizia penale, controllo umano significativo, in Diritto penale contemporaneo, 2020, n. 4, 75 ss. Cfr. ID., Necessaria compatibilità dell'intelligenza artificiale con il giusto processo, in Archivio penale, 19 dicembre 2024; M. GIALUZ, Prove fondate sull'intelligenza artificiale e diritti fondamentali, cit., 18.

[48] Si vedano, per esempio, i tools per l'anonimizzazione dei dati sensibili in ambito legale. Cfr. A. OKSANEN, E. HYVÖNEN, M. TAMPER, J. TUOMINEN, H. YLIMAA, K. LÖYTYNOJA, A. HIETANEN, An anonymization tool for open data publication of legal documents, in International Workshop on Artificial Intelligence Technologies for Legal Documents/International Workshop on Knowledge Graph Summarization, 2022, 12; M. EBERHARDINGER, P. TAKENAKA, D. GRIEßHABER, J. MAUCHER, Anonymization of Documents for Law Enforcement with Machine Learning, in ArXiv, 2025, 1 ss.

[49] Cfr. F. DE STEFANI, Le regole della privacy: guida pratica al nuovo GDPR, Milano, 2018; S. SANJAY; Data privacy and GDPR handbook, Hoboken, 2019.

[50] Cfr. S. QUATTROCOLO, Intelligenza artificiale e processo penale: le novità dell'AI Act, cit., 7.

[51] Tratta di sorveglianza attraverso i metadati P. BRANCH, Surveillance by metadata, Vol. 109, 2014, 10 ss.

[52] Corte di giustizia dell'U.E., Tele2 Sverige AB e Watson, cit.

[53] Cfr. supra, §2.2.

[54] Che, come nota S. QUATTROCOLO, Intelligenza artificiale e processo penale: le novità dell'AI Act, cit., 5, riflette il divieto di perizia personologica del nostro ordinamento di cui all'art 220, comma 2, c.p.p.

[55] Cfr. M. GORI, Introduzione alle reti neurali artificiali, in Mondo digitale, Vol. 8, 2003, n. 2, 4-20.

[56] M. B., SHAWN, Data Protection in the United States, in The American Journal of Comparative Law, Vol. 66, 2018, n. 1, 299–343; L. EBENIBO, J.O. ENYEJO, G.ADDO, M.O. TOYOSI, Evaluating the Sufficiency of the data protection act 2023 in the age of Artificial Intelligence (AI): A comparative case study of Nigeria and the USA, in International Journal of Scholarly, 2024, n. 1, 88-107.

CamminoDiritto.it
ISSN 2532-9871

[57] Cfr. A. CAHANE, D. DROR-SHPOLIANSKY, Human Rights and AI? On the The Blueprint for an AI Bill of Rights and its meaning, cit.

[58] Cfr. supra, §2.2.

[59] Cfr. supra, §2.1.

[60] G. CANZIO, AI Act e processo penale: sfide e opportunità, cit., 2.

[61] Cfr. supra, §2.1.

[62] Cfr. supra, §3.

[63] Cfr. D. BOGDANOV, P. ETTI, L. KAMM, F. STOMAKHIN, Artificial Intelligence System Risk Management Methodology Based on Generalized Blueprints, in 2024 16th International Conference on Cyber Conflict: Over the Horizon (CyCon), Berlino, 2024, 123 ss.

[64] Cfr. supra, §5.

---

* Il simbolo {https/URL} sostituisce i link visualizzabili sulla pagina:
https://rivista.camminodiritto.it/articolo.asp?id=10902