Ordine internazionale e diritti umani
International Legal Order and Human Rights
Ordenamiento Jurídico Internacional y Derechos Humanos
Ordre Juridique International et Droits de l'Homme
Diretta da Claudio Zanghì, Lina Panella, Carlo Curti Gialdino
EDITORIALE SCIENTIFICA

MARCO SILIGATO*

IL SETTORE DELLA GIUSTIZIA NELL'AI ACT: IL DOVERE DI TRASPARENZA COME "ANIDRIDE SOLFOROSA"?

SOMMARIO: 1. Introduzione. – 2. Profili strutturali dell'*AI Act* e applicazione al settore della giustizia. – 3. L'utilizzo dei sistemi IA nel settore della giustizia penale: fra attività vietate e ad «alto rischio». – 4. Il dovere di trasparenza: reale garanzia?. – 5. Conclusioni.

1. *Introduzione*

L'idea al centro del presente lavoro affonda le radici nella potente metafora dell'anidride solforosa: con tale immagine, Lucio Dalla, già nel 1975 riportava in versi i propri dubbi in merito al futuro uso dei sistemi informatici, ai rischi legati all'uso di uno strumento calcolatore alienante e disumanizzante, anticipando profeticamente l'impatto dirompente e il potenziale effetto degli strumenti informatici sulla società, ben prima che la nuova tecnologia entrasse nel dibattito sociale.

Traendo spunto da tale provocazione, seppur non ispirata all'intelligenza artificiale, si indagherà sull'attuale regolamentazione in merito alla tutela dei diritti fondamentali riconosciuta agli individui, per provare a comprendere se l'accesso al nuovo strumento informatico sia ancora "fumoso".

Le tecnologie più moderne basate su sistemi di intelligenza artificiale (IA) hanno ormai fatto ingresso in ogni ambito della vita quotidiana nonostante tale strumento sia talmente nuovo che non se ne conoscono ancora la reale portata e dimensione.

La nuova strumentazione informatica, sulla base di una sempre maggiore disponibilità di dati legati a nuovi canali digitali e spazi comunicativi *online*, ha infatti dato vita a quello che è stato definito un "ecosistema digitale"[1], offrendo supporto ai cittadini tramite riferimenti e approfondimenti, anche in ambito giudiziario.

* Dottorando di ricerca in Studi europei, Dottorato di Interesse Nazionale (DIN) dell'Università di Genova, *Curriculum Governance multilivello e diritti fondamentali,* sede di afferenza Università degli Studi di Messina, Dipartimento di Scienze Politiche e Giuridiche.

[1] A. CORRERA, *Il ruolo dell'intelligenza artificiale nel paradigma europeo dell'*E-justice. *Prime riflessioni alla luce dell'*AI Act, in *QuaderniAISDUE*, n. 2/2024, p. 2.

In tale ottica, il presente lavoro intende analizzare il «dovere di trasparenza» così come delineato nel Regolamento (UE) n. 1689/2024 (d'ora innanzi *"AI Act*[2]*"*), sottolineandone la specificità e le implicazioni pratiche, nonché i limiti che possono emergere da definizioni non ancora applicate e interpretate.

Dopo aver presentato, seppur sinteticamente, i contenuti essenziali della regolamentazione del settore della giustizia penale dell'*AI Act*, evidenziando le peculiarità applicative degli strumenti di IA in tale ambito, si procederà all'analisi del dovere di trasparenza così come enucleato nel Regolamento, esaminando la coerenza e l'effettività di tali previsioni, nonché i profili di spiegabilità e responsabilità.

In questo senso, l'analisi dei riferimenti normativi dell'*AI Act* verterà sull'imposizione ai vari soggetti coinvolti (utilizzatori, fornitori, ecc.), di divieti o obblighi tecnico-procedurali a tutela della trasparenza, graduati in base al settore di applicazione e allo *standard* di tutela richiesto, mediante una "procedimentalizzazione" del rischio[3], con particolare riguardo al settore della giustizia penale.

Nelle conclusioni, dunque, volendo sfuggire dalla mera ritualità, partendo dal citato ambito applicativo dei sistemi di IA nel settore della giustizia penale, si tenterà di fornire un punto di vista circa la valutazione dottrinale del livello di garanzia della disciplina relativa a tale settore, nell'ambito della recente regolamentazione europea.

2. *Profili strutturali dell'AI Act e applicazione al settore della giustizia*

La rapida diffusione dei sistemi basati sull'IA ha generato un dibattito globale sulla necessità di un quadro normativo capace di bilanciare innovazione tecnologica e tutela dei diritti fondamentali.

In questo quadro si inserisce il solido approccio europeo adottato dalla Commissione europea già nel 2020 con il Libro bianco sull'intelligenza artificiale[4] e sviluppatosi negli anni attraverso numerosi atti legislativi di varia forma[5], di recente cristallizzatosi nell'*AI Act*, con

---

[2] 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 Serie L del 12/7/2024.

[3] A. FORMISANO, *L'impatto dell'intelligenza artificiale in ambito giudiziario sui diritti fondamentali*, in *Federalismi.it*, n. 22/2024, p. 136.

[4] Commissione europea, *Libro bianco sull'intelligenza artificiale - Un approccio europeo all'eccellenza e alla fiducia*, del 19 febbraio 2020, COM(2020) 65 final.

[5] Commissione europea, *Piano coordinato sull'intelligenza artificiale*, del 7 dicembre 2018, COM(2018) 795 final (modificato il 21 aprile 2021); Id., *Creare fiducia nell'intelligenza artificiale antropocentrica*, dell'8 aprile 2019, COM/2019/168 final; Id., *Libro bianco sull'intelligenza artificiale Un approccio europeo all'eccellenza e alla fiducia*, del 19 febbraio 2020, COM(2020) 65 final; Id., *Relazione sulle implicazioni dell'intelligenza artificiale, dell'Internet delle cose e della robotica in materia di sicurezza e di responsabilità*, 19 febbraio 2020, COM (2020) 64 final; Parlamento europeo, *Risoluzione sui diritti di proprietà intellettuale per lo sviluppo di tecnologie di intelligenza artificiale*, del 20 ottobre 2020, (2020/2015(INI)); Id., *Risoluzione sui processi decisionali automatizzati: garantire la tutela dei consumatori e la libera circolazione di beni e servizi*, del 12 febbraio 2020, (2019/2915(RSP)); Id., *Risoluzione sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale*; *Dichiarazione europea sui diritti e i principi digitali per il decennio digitale,* 2023/C 23/01, del 6 ottobre 2021, P9_TA(2021)0405; Corte dei Conti UE, *Relazione speciale 08/2024*: *Le ambizioni dell'UE in materia di intelligenza artificiale – Per il futuro, una governance più forte e investimenti più consistenti e mirati sono essenziali*, del 29 maggio 2024.

cui l'Unione, perseguendo l'ambizioso obiettivo di guidare lo sviluppo dell'IA a livello mondiale[6], ha istituito vincoli uniformi e direttamente applicabili all'interno degli Stati membri in materia di utilizzo di sistemi di intelligenza artificiale[7].

Ai fini del presente lavoro pare opportuno, seppur brevemente, riferire in merito alla duplice base giuridica del Regolamento[8]: l'articolo 114 TFUE, in virtù del quale l'Unione può adottare misure volte al ravvicinamento delle normative nazionali in tema di funzionamento del mercato interno, e l'articolo 16 TFUE, sulla base del quale l'Unione può adottare atti legislativi in materia di protezione dei dati personali.

In tale ottica, infatti, si inserisce la finalità di promozione dell'innovazione e di tutela dei diritti fondamentali dell'*AI Act*[9], coerente con la scelta di legare la pervasività degli interventi in base al grado di rischio di ciascun sistema secondo il c.d. *risk-based approach*[10].

Entrando nel dettaglio della struttura dell'*AI Act*, il Regolamento distingue quattro tipi di sistemi IA in base al livello di rischio riconosciuto, definiti quali: inaccettabile, alto, moderato e basso[11].

Come già accennato, per ciascun livello il Regolamento ha graduato la portata della disciplina: per i sistemi a rischio inaccettabile è previsto il divieto di utilizzo dei sistemi IA con specifiche eccezioni[12]; per i sistemi a rischio alto, aprioristicamente ammissibili, sono previsti i necessari requisiti per l'applicabilità e l'utilizzo[13]; per i sistemi a rischio limitato sono previsti puntuali obblighi di trasparenza[14]; per quelli a rischio basso, la semplice sottoscrizione di codici di condotta[15].

A tale quadro si aggiunge la figura dei cosiddetti modelli di IA per finalità generali (GPAI)[16], così definiti in quanto non aprioristicamente classificabili in virtù della loro capacità di sviluppare abilità diverse rispetto a quelle predeterminate al momento della programmazione[17].

Entrando nel merito dell'analisi di cui al presente paragrafo, deve sottolinearsi che l'Unione europea, dinanzi alla sfida posta dalla "nuova" intelligenza artificiale, ha rivolto

---

[6] O. POLLICINO, F. DONATI, G. FINOCCHIARO, F. PAOLUCCI, *La disciplina dell'intelligenza artificiale*, Milano, 2025.
[7] P. INTURRI, S. FICHERA, A. COSTA, *La disciplina dei sistemi di intelligenza artificiale per l'amministrazione della giustizia nel Regolamento (UE) 2024/1689,* in *LavoroDirittiEuropa*, n. 1/2025, p. 2.
[8] Per approfondimenti, si consiglia: E. CIRONE, *L'*AI Act *e l'obiettivo (mancato?) di promuovere uno* standard *globale per la tutela dei diritti fondamentali*, in *QuaderniAISDUE*, n. 2/2024, p. 4.
[9] Regolamento (UE) 2024/1689, *cit.*, p.1; al considerando n. 1 si legge: «[Lo scopo del presente regolamento è] migliorare il funzionamento del mercato interno istituendo un quadro giuridico uniforme in particolare per quanto riguarda lo sviluppo, l'immissione sul mercato, la messa in servizio e l'uso di sistemi di IA in conformità con i valori dell'Unione, promuovere la diffusione di una IA antropocentrica e affidabile, garantendo nel contempo un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'Unione europea, compresi la democrazia, lo Stato di diritto e la protezione dell'ambiente, proteggere contro gli effetti nocivi dei sistemi di IA nell'Unione, nonché promuovere l'innovazione».
[10] G. ZICCARDI, *Una lettura dell'*Artificial Intelligence Act: *norme, etica, adempimenti, attuazione,* in AA. VV. (a cura di), *Intelligenza Artificiale. Diritto, giustizia, economia ed etica, Pubblicazioni del Dipartimento di Scienze Giuridiche "Cesare Beccaria"*, Bari, 2025, p. 20.
[11] S. QUATTROCOLO, *Intelligenza artificiale e processo penale: le novità dell'*AI Act, in *Diritto di difesa – La rivista dell'Unione delle Camere Penali Italiane*, 2025, p. 4.
[12] Regolamento (UE)1689/2024, *cit.*, Capo II.
[13] Regolamento (UE)1689/2024, *cit.*, Capo III.
[14] Regolamento (UE)1689/2024, *cit.*, Capo IV.
[15] Regolamento (UE)1689/2024, *cit.*, Capo X.
[16] Regolamento (UE)1689/2024, *cit.*, Capo V.
[17] Regolamento (UE)1689/2024, *cit.*, art. 3, par. 1, n. 63.

l'attenzione anche ai sistemi giudiziari degli Stati membri[18], comprendendo la necessità di compiere uno sforzo di adeguamento ulteriore per garantire un accesso alla giustizia moderno, efficiente, non invasivo e senza barriere, che rispetti il dogma dello Stato di diritto e che rafforzi la fiducia nel sistema istituzionale[19].

Invero, dal considerando n. 61 del Regolamento si ricava la collocazione dei sistemi di intelligenza artificiale applicati alla giustizia tra le attività ad «alto rischio», precisandosi che «al fine di far fronte ai rischi di potenziali distorsioni, errori e opacità» sono da considerarsi tali 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».

Tale locuzione è infatti ripresa nell'Allegato III del Regolamento[20], quale unico caso di utilizzo delle applicazioni IA con impatto potenziale sulla sicurezza dei cittadini e sui diritti fondamentali nel settore della giustizia[21].

In questo senso, circa la programmazione dei modelli di IA nei settori ad «alto rischio», l'*AI Act* precisa i criteri progettuali che devono garantire l'accessibilità, la qualità e l'adeguatezza dei *dataset* necessari e, allo stesso tempo, individuare, prevenire e colmare possibili distorsioni o gravi inadeguatezze nelle informazioni offerte agli utenti[22].

Diventa dunque importante facilitare l'accesso ai dati che vengono utilizzati per "addestrare" i sistemi di IA al fine di poterne verificare la genuinità e la qualità, nonché per garantire la trasparenza della decisione assunta in base all'algoritmo.

Il principio, infatti, così come delineato nel Regolamento, sembra richiamare la più classica declinazione del principio di trasparenza in materia di diritto dei consumatori, da intendersi quale accesso continuativo ad informazioni intelligibili, assenza di ambiguità e arbitrarietà nei comportamenti assunti dal professionista, il quale è soggetto a determinati obblighi professionali[23].

Inoltre, nell'ambito del dibattito dottrinale sull'utilizzo dei sistemi di IA nel settore della giustizia, si è comunque concordi sull'utilizzo della strumentazione informatica quale strumento complementare all'intelligenza e alla sensibilità umana, soprattutto per la necessità di garantire il rispetto dei valori democratici su cui si fonda l'Unione europea[24].

3. *L'utilizzo dei sistemi IA nel settore della giustizia penale: fra attività vietate e ad «alto rischio»*

Preliminare all'analisi della relazione tra intelligenza artificiale (IA) e il settore della giustizia penale è una breve ma essenziale ricognizione del concetto di intelligenza artificiale come definito nel contesto della normativa dell'Unione Europea.

---

[18] P. P. PAULESU, *Intelligenza artificiale e giustizia penale. Una lettura attraverso i principi*, in *Archivio penale*, n. 1/2022, p. 13.
[19] A. CORRERA, *op. cit.*, p. 2.
[20] Regolamento (UE) 2024/1689, *cit.*, Allegato III, p. 127 e s.
[21] Per un approfondimento, si veda O. CARAMASCHI, *Il costituzionalismo al cospetto dell'intelligenza artificiale: nuove sfide, quali soluzioni?*, in *Rivista italiana di informatica e diritto*, n. 1/2025.
[22] *Ivi*, art. 10, par. da 2 a 5.
[23] F. DEANA, *La mano invisibile dell'Intelligenza Artificiale e il principio di trasparenza nei rapporti B2C: la tutela del consumatore nel mercato unico digitale"*, in *Freedom, Security & Justice*, n. 1/2025, pp. 66-85.
[24] V. MANES, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in AA.VV., *Intelligenza Artificiale – Il diritto, i diritti, l'etica*, U. RUFFOLO (a cura di), Milano, 2020.

Nello specifico, il concetto di «sistema IA» di cui all'articolo 3 dell'*AI Act*[25] è legato al riferimento a «livelli di autonomia variabile», dovendo pertanto espungersi dal nucleo concettuale i sistemi «automatizzati» che eseguono comandi senza alcuna forma di rielaborazione[26].

La "rivoluzione algoritmica" paventata in dottrina rappresenta dunque la nuova frontiera tecnologica in cui l'umano è chiamato a "misurarsi" con la macchina[27].

Merita certamente sottolineatura il mancato riferimento espresso alla natura "collaborativa" dell'intelligenza artificiale, potendo la stessa, secondo la lettera della norma, generare «decisioni».

Come osservato[28], tale riferimento non permette di escludere aprioristicamente la previsione di un modello di intelligenza artificiale «forte» volta «all'automazione del processo decisionale», in luogo di quella «debole», che collabora con la figura preposta in assenza di qualsivoglia autorità decisionale.

In questo senso, si è sostenuto che in un'ottica mitigatoria di tale accezione devono intendersi i dettami dei considerando n. 2 e n. 27, oltre che lo stesso articolo 1 del Regolamento, i quali precisano che l'intelligenza artificiale deve rispettare i criteri di antropocentrismo e affidabilità, garantendo così il necessario apporto umano[29].

A parere di chi scrive, debbono ritenersi coerenti anche le specifiche disposizioni relative ai settori ad «alto rischio» relative all'obbligo di garantire la trasparenza[30], la supervisione e la sorveglianza umana[31], oltre a una valutazione sui possibili pregiudizi sui diritti fondamentali[32].

Innanzitutto, deve segnalarsi che il Regolamento include fra le attività vietate nel settore della giustizia penale due specifici impieghi dei sistemi di intelligenza artificiale[33].

La prima attività vietata[34] riguarda l'utilizzo degli algoritmi di *risk assessment*[35], ad esclusione di quelli volti a fornire valutazioni basate su fatti oggettivi e verificabili, direttamente connessi a un'attività criminosa.

Si intende qui condividere il punto di vista dottrinale per cui la formula relativa ai «fatti oggettivi e verificabili» sia di scarsa efficacia e rischi di ricondurre l'utilizzo degli strumenti di

---

[25] Letteralmente: «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».

[26] M. GIANGRECO, *Gli spazi applicativi dell'intelligenza artificiale nella giustizia penale: riflessioni alla luce dell'*AI Act *dell'UE e del* Blueprint *statunitense,* in *Cammino Diritto, Rivista di Informazione Giuridica*, n. 3/2025, p. 5.

[27] S. LORUSSO, *La sfida dell'intelligenza artificiale al processo penale nell'era digitale*, in *Sistema penale*, 2024, p. 1 ss.

[28] G. CANZIO, AI Act *e processo penale: sfide e opportunità*, in *Sistema penale*, 2024, p. 2.

[29] In *Ibidem,* p. 4.

[30] Regolamento (UE)1689/2024, *cit.*, art. 13.

[31] Regolamento (UE)1689/2024, *cit.*, art. 14.

[32] Regolamento (UE)1689/2024, *cit.*, art. 27.

[33] I divieti di cui all'art. 5 del Regolamento hanno trovato applicazione a decorrere dal 2 febbraio 2025.

[34] Regolamento (UE) 1689/2024, art. 5, lett. d): «l'immissione sul mercato, la messa in servizio per tale finalità specifica o l'uso di un sistema di IA per effettuare valutazioni del rischio relative a persone fisiche al fine di valutare o prevedere il rischio che una persona fisica commetta un reato, unicamente sulla base della profilazione di una persona fisica o della valutazione dei tratti e delle caratteristiche della personalità; tale divieto non si applica ai sistemi di IA utilizzati a sostegno della valutazione umana del coinvolgimento di una persona in un'attività criminosa, che si basa già su fatti oggettivi e verificabili direttamente connessi a un'attività criminosa».

[35] E. GUIDO, *Intelligenza artificiale e procedimento penale: ragionando di valutazione del rischio* de libertate, in *Archivio penale*, 2023, n. 1.

predictive policing vietati a valutazioni soggettive nonostante la preclusione dalla prima parte della disposizione in esame[36].

In questo senso, nella valutazione del rischio che un soggetto possa commettere un reato, l'algoritmo predittivo dovrebbe escludere ogni apprezzamento di dati soggettivi - fra cui i precedenti penali - altrimenti sfociando nel diritto penale d'autore[37], oltre a entrare in contrasto con il principio di presunzione di innocenza.

Inoltre, risulta immediatamente evidente che la norma sia indirizzata al divieto di utilizzo degli algoritmi di *risk assessment* nella valutazione di una singola persona, escludendo ogni tipo di riferimento al divieto di applicazione su di un gruppo di persone o su di un'intera comunità[38].

L'altra attività vietata riguarda l'utilizzo dei sistemi di identificazione biometrica remota «in tempo reale»[39] in luogo pubblico a fini di attività di contrasto[40], anche qui con specifiche eccezioni, a vario titolo riguardanti il contrasto a specifici crimini[41].

Tale decisione è coerente con quanto già sostenuto dall'Agenzia europea per i diritti fondamentali (FRA), la quale caldeggiava un utilizzo «strictly determined» e limitato dell'utilizzo dello strumento di riconoscimento biometrico[42].

Ulteriore limite è posto con riferimento all'obbligo di ottenere specifica autorizzazione da un'autorità giudiziaria indipendente dello Stato membro per poter svolgere attività di

---

[36] In M. GIANGRECO, *op. cit.*, p. 5; E. SACCHETTO, *Spunti per una riflessione sul rapporto tra biometria e processo penale*, in *Diritto penale contemporaneo*, n. 2/2019, p. 465 ss.

[37] In tal senso dispone il considerando n. 42, secondo cui 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 dall'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».

[38] Di questo avviso M. GIANGRECO, *op. cit.*, pag. 6, il quale ritiene che «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».

[39] Sull'utilizzo di strumenti di identificazione biometrica, si rimanda a: E. SACCHETTO, *La prova biometrica*, in A. MARANDOLA, C. CONTI (a cura di), *La prova scientifica*, Milano, 2023, p. 243 ss.; G. MOBILIO, *Tecnologie di riconoscimento facciale. Rischi per i diritti fondamentali e sfide regolative*, Napoli, 2021; J. DELLA TORRE, *Novità dal Regno Unito: il riconoscimento facciale supera il vaglio della High Court of Justice (ma raccoglie le critiche del Garante Privacy D'Oltremanica)*, in *Sistema penale*, 2020, p. 3.

[40] Per approfondimenti, si consiglia E. SACCHETTO, Face to face*: il complesso rapporto tra* automated facial recognition technology *e processo penale*, in *La legislazione penale*, 2020; B. PIETROCARLO, Predictive policing*: criticità e prospettive dei sistemi di identificazione dei potenziali criminali*, in *Sistema penale*, 2023.

[41] Regolamento (UE) 1689/2024, *cit.*, art. 5, lett. h): «l'uso di sistemi di identificazione biometrica remota «in tempo reale» in spazi accessibili al pubblico a fini di attività di contrasto a meno che, e nella misura in cui, tale uso sia strettamente necessario per uno degli obiettivi seguenti: i) la ricerca mirata di specifiche vittime di sottrazione, tratta di esseri umani o sfruttamento sessuale di esseri umani, nonché la ricerca di persone scomparse; ii) la prevenzione di una minaccia specifica, sostanziale e imminente per la vita o l'incolumità fisica delle persone fisiche o di una minaccia reale e attuale o reale e prevedibile di un attacco terroristico; iii) la localizzazione o l'identificazione di una persona sospettata di aver commesso un reato, ai fini dello svolgimento di un'indagine penale, o dell'esercizio di un'azione penale o dell'esecuzione di una sanzione penale per i reati di cui all'allegato II, punibile nello Stato membro interessato con una pena o una misura di sicurezza privativa della libertà della durata massima di almeno quattro anni.»

[42] European Union Agency for Fundamental Rights (FRA), Facial recognition technology: fundamental rights considerations in the context of law enforcement, p. 25.

contrasto in luogo pubblico ma, anche in questo caso, la lettera della norma riferisce le eccezioni relative ai casi una situazione di urgenza debitamente giustificata[43].

Va in ogni caso sottolineato che la formulazione letterale delle eccezioni al generale divieto di utilizzo può risultare generica e, secondo la dottrina, tale difetto rischia di determinare disomogeneità applicativa nei diversi Stati membri, in particolare, circa la ricerca e l'individuazione degli autori di determinati reati[44].

Passando alle attività ritenute ad «alto rischio»[45] con riferimento al settore della giustizia penale, l'Allegato III[46] qualifica come tali i «sistemi usati per assistere un'autorità giudiziaria nella ricerca e nell'interpretazione dei fatti e del diritto e nell'applicazione della legge a un insieme concreto di fatti».

È opportuno soffermarsi sulla portata di questa formula che, se letta da sola, non è scevra di potenziali fraintendimenti.

Il riferimento è ovviamente all'uso dei sistemi di IA per scopi ermeneutici che minano la genuinità della decisione[47] e limitano il principio fondamentale del libero convincimento del giudice[48], a causa dell'opacità che caratterizza i sistemi di giustizia predittiva[49].

Tale caratteristica, fra l'altro, a detta della Corte di giustizia, è legata al fatto che potrebbe «risultare impossibile comprendere la ragione per la quale un dato programma sia arrivato ad un riscontro positivo» e, dunque, «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]»[50].

Con riferimento a questo tipo di sistemi, la disciplina regolamentare prevede una serie di obblighi specifici, fra cui quello di predisporre e mantenere una documentazione tecnica dettagliata, la trasparenza per gli utenti – di cui si dirà –, la necessaria sottoposizione al controllo umano e l'adozione di misure *ad hoc* per i relativi rischi legati all'utilizzo del sistema.

Va detto che dal corpo della norma di cui all'articolo 9 del Regolamento si ricava quello che, come riportato in rubrica, è il «sistema di gestione dei rischi», consistente in un'articolata

---

43 Regolamento (UE) 1689/2024, *cit.*, art. 5 par. 3: "Ai fini del paragrafo 1, primo comma, lettera h), e del paragrafo 2, ogni uso di un sistema di identificazione biometrica remota «in tempo reale» in spazi accessibili al pubblico a fini di attività di contrasto è subordinato a un'autorizzazione preventiva rilasciata da un'autorità giudiziaria o da un'autorità amministrativa indipendente, la cui decisione è vincolante, dello Stato membro in cui deve avvenire l'uso, rilasciata su richiesta motivata e in conformità delle regole dettagliate del diritto nazionale di cui al paragrafo 5. Tuttavia, in una situazione di urgenza debitamente giustificata, è possibile iniziare a usare il sistema senza autorizzazione a condizione che tale autorizzazione sia richiesta senza indebito ritardo, al più tardi entro 24 ore. Se tale autorizzazione è respinta, l'uso è interrotto con effetto immediato e tutti i dati nonché i risultati e gli *output* di tale uso sono immediatamente eliminati e cancellati.»

44 In V. VASTA, *Diritto dell'Unione europea e intelligenza artificiale. Riflessi sul procedimento penale*, in AA. VV. (a cura di), *Intelligenza Artificiale. Diritto, giustizia, economia ed etica*, *Pubblicazioni del Dipartimento di Scienze Giuridiche "Cesare Beccaria"*, Bari, 2025, p. 127 ss., l'A ritiene che difficoltà applicative possano in particolare rinvenirsi nella definizione di un utilizzo relativo ad un nucleo di reati individuati sulla base del limite massimo edittale della pena, potendo questo mutare nelle varie legislazioni nazionali.

45 La disciplina relativa ai sistemi ad «alto rischio» di cui all'art. 6, par. 1, dell'*AI Act* si applicano a decorrere dal 2 agosto 2027.

46 Reg. (UE) 1689/2024, *cit.*, Allegato III, punto 8, lett. a).

47 A. MACERATINI, *La giustizia predittiva: potenzialità e incognite,* in *MediaLAWS*, n. 2/2024, p. 239.

48 In merito a tali riflessioni, si consiglia: R. E. KOSTORIS, *Intelligenza artificiale, strumenti predittivi e processo penale*, in *Cassazione penale*, n. 5/2024, p. 1642 ss.; G. UBERTIS, *Necessaria compatibilità dell'intelligenza artificiale con il giusto processo*, in *Archivio penale*, 2024; A. BALSAMO, *L'impatto dell'intelligenza artificiale nel settore della giustizia*, in *Sistema penale*, 2024; S. LORUSSO, *La sfida dell'intelligenza artificiale al processo penale nell'era digitale*, in *Sistema penale*, 2024.

49 G. BARONE, *Giustizia predittiva e certezza del diritto*, Pisa, 2024.

50 Corte di Giustizia UE (Grande Camera), sentenza del 21 giugno 2022, C-817/19, ECLI:EU:C:2022:491.

struttura giuridica composta da cautele, informazioni, garanzie, prove, sviluppi e progettazioni specifiche da adottare «al fine di eliminare o ridurre i rischi connessi all'uso del sistema di IA ad alto rischio».

A chiosa, lo stesso considerando n. 61 esclude dal novero dei sistemi di intelligenza artificiale ad «alto rischio», quelli «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»[51].

Come si vedrà di seguito, il sistema predisposto dall'*AI Act* è volto a promuovere un intervento di tipo "proattivo" più che "reattivo"[52].

Tale tipo di approccio, infatti, è legato all'imposizione nei confronti dei soggetti coinvolti di una serie di obblighi e attività che danno vita ad un articolato sistema di responsabilità[53] che, esulando in ogni caso dalla stringente disciplina penale, permette di delimitare una zona bianca in cui considerare comunque lecito l'utilizzo di strumenti IA.

4. *Il dovere di trasparenza: reale garanzia?*

Fra gli obblighi più stringenti e potenzialmente più influenti nel settore della giustizia vi è l'obbligo di trasparenza, più volte inserito nel corpo del testo del Regolamento, quale componente necessaria per la conformità dei sistemi ad «alto rischio», da intendersi quale dovere di informare in modo chiaro, dettagliato e comprensibile gli utenti sul funzionamento dell'IA, specialmente in ambiti come la sorveglianza biometrica o la creazione di contenuti generati artificialmente[54].

Tale impegno è stato di recente ribadito anche nella Strategia europea in materia di giustizia elettronica, ove si legge che «spostare l'attenzione sulla prospettiva delle persone e rendere i sistemi giudiziari più accessibili, efficaci e trasparenti sarà […] essenziale per rafforzare la fiducia tra le persone e le istituzioni pubbliche»[55].

In senso più pratico, deve essere assolutamente chiaro che l'utente sta interagendo con un sistema IA e lo stesso deve avere facile accesso alle informazioni relative al funzionamento del modello e ai criteri che guidano le decisioni automatizzate[56].

Dal considerando n. 27, invero, si ricava l'interpretazione del concetto di «trasparenza» nel senso che «i sistemi di IA sono sviluppati e utilizzati in modo da consentire un'adeguata tracciabilità e spiegabilità, rendendo gli esseri umani consapevoli del fatto di comunicare o

---

[51] M. TORRE, *Il Regolamento europeo sull'intelligenza artificiale: profili processuali*, in *Processo penale e giustizia*, 2024, n. 6, p. 1534 ss.

[52] J. DELLA TORRE, *Quale spazio per i tools di riconoscimento facciale nella giustizia penale?*, in G. DI PAOLO, L. PRESSACCO (a cura di), *Intelligenza artificiale e processo penale. Indagini, prove, giudizio*, Napoli, 2022, p. 20 ss.

[53] A tal proposito, per un quadro generale si veda M. E. FLORIO, *Il dibattito sulla responsabilità penale diretta delle I.A.: "molto rumore per nulla"?*, in *Sistema penale*, n. 2/2024.

[54] G. ZICCARDI, *op. cit.*, p. 27 ss.

[55] Strategia europea in materia di giustizia elettronica 2024-2028 (C/2025/437), in *G.U.U.E.*, 16 gennaio 2025, C/1, Capo II-Principi, Par. A Principi sostanziali, lettera c) centralità delle persone.

[56] S. QUATTROCOLO, *Quesiti nuovi e soluzioni antiche?, Consolidati paradigmi normativi vs rischi e paure della giustizia digitale*, in *Cassazione penale*, 2019, pp. 1748-1765.

interagire con un sistema di IA e informando debitamente i *deployer*[57] delle capacità e dei limiti di tale sistema di IA e le persone interessate dei loro diritti».

Con riferimento alla «tracciabilità», si è già detto che per i sistemi ad «alto rischio» è fatto obbligo di predisporre documentazione tecnica che garantisca la conformità del sistema e che sia dettagliata in merito al funzionamento dell'algoritmo, sui dati utilizzati per l'addestramento e sui meccanismi di mitigazione del rischio.

Circa il profilo della «spiegabilità» del sistema pare utile soffermarsi, ben comprendendosi che l'uso dell'intelligenza artificiale nel settore della giustizia potrebbe avere un ruolo decisivo per i cittadini nella chiarificazione di concetti giuridici non di immediata comprensione.

In questo senso, la dottrina è giunta a riconoscere quattro caratteri chiave del concetto di «spiegabilità»: l'interpretabilità, la trasparenza, la giustificabilità e la riproducibilità[58].

Si ricava, pertanto, che per i sistemi di IA ad «alto rischio» è prevista la necessità di fornire informazioni appropriate agli utilizzatori, mediante istruzioni per l'uso che coprano le caratteristiche, le prestazioni, i limiti del sistema, i rischi possibili e le misure di controllo umano pertinenti, anche attraverso esempi esplicativi[59].

Non a caso, in ossequio ai dettami di cui all'articolo 96 del Regolamento, rubricato "Orientamenti della Commissione sull'attuazione del regolamento", nel luglio del 2025 la Commissione europea ha pubblicato il Codice di buone pratiche sull'IA per finalità generali, cui si affiancano orientamenti sui concetti chiave relativi ai modelli di intelligenza artificiale per finalità generali (GPAI).

Limitandoci qui a riferire del carattere volontario di tale Codice, la Commissione ha allegato, nel capo sulla trasparenza, un "modulo di documentazione tipo" che potrà consentire ai fornitori di modelli GPAI di documentare facilmente le informazioni necessarie in conformità al Regolamento sull'IA[60].

In effetti, dal considerando 72 non era chiaro in che modo sia in concreto possibile descrivere e spiegare parametri tecnici nel linguaggio naturale, quale livello di chiarezza e precisione nella spiegazione sia necessario e come sia possibile valutarne il raggiungimento[61].

---

[57] Tale termine, non tradotto nella versione italiana del Regolamento, intende, secondo l'art. 3, par. 4: «una persona fisica o giuridica, un'autorità pubblica, un'agenzia o un altro organismo che utilizza un sistema di IA sotto la propria autorità, tranne nel caso in cui il sistema di IA sia utilizzato nel corso di un'attività personale non professionale».

[58] In *ibidem*, p. 22, l'A. precisa i contenuti di ciascun carattere: « i) la interpretabilità (riguarda la possibilità di comprendere le relazioni tra *input* e *output* di un modello di IA, e alcuni modelli, come le reti neurali profonde, sono difficili da interpretare, mentre altri, come gli alberi decisionali, sono più trasparenti); ii) la trasparenza (si riferisce alla capacità di esaminare il funzionamento interno di un sistema di IA, come i pesi e i parametri di un modello di apprendimento automatico, e anche in questo caso alcuni algoritmi sono considerati "scatole nere" perché il loro processo decisionale è opaco e complesso); iii) la giustificabilità (implica la possibilità di fornire motivazioni coerenti e logiche che spieghino una decisione presa dal sistema, in modo comprensibile per gli esseri umani), e iv) la riproducibilità (significa che un sistema di IA dovrebbe produrre risultati simili quando vengono forniti dati e condizioni simili, permettendo di verificare e validare il suo funzionamento)».

[59] M. R. ALLEGRI, *L'ambiguo principio (anche costituzionale?) della trasparenza algoritmica fra tecnologia, diritto e linguaggio*, in *MediaLAWS,* n. 3/2024, pp. 19-27.

[60] Commissione europea, *Transparency Chapter,* in *Code of Practice for General-Purpose AI Models*, del 10 luglio 2025.

[61] Dal testo si ricava che: «i fornitori dovrebbero garantire che tutta la documentazione, comprese le istruzioni per l'uso, contenga informazioni significative, complete, accessibili e comprensibili, tenendo conto delle esigenze e delle conoscenze prevedibili dei *deployer* destinatari. Le istruzioni per l'uso dovrebbero essere messe a disposizione in una lingua che possa essere compresa facilmente dai *deployer* destinatari, secondo quanto stabilito dallo Stato membro interessato».

L'articolo 11 lascia analoghe zone di incertezza applicativa nel disciplinare che i sistemi di IA ad alto rischio devono essere corredati dalla documentazione tecnica prevista – in maniera molto specifica – dall'Allegato IV del Regolamento, da fornire alle autorità nazionali competenti e agli organismi notificati.

Tali perplessità sono emerse in particolare sulla possibilità di tradurre concretamente le regole tecniche dal linguaggio computazionale al linguaggio naturale, per renderle comprensibili a persone non necessariamente esperte del linguaggio matematico-informatico[62].

Come già accennato, infatti, l'Unione europea ha interesse a perseguire parallelamente sia lo sviluppo dello strumento di IA nel rispetto dei principi fondamentali, sia l'accessibilità da parte dei cittadini in termini di trasparenza e comprensibilità[63].

Tale volontà legislativa si è tradotta, nell'ambito della giustizia, nella scelta di sfruttare lo strumento informatico nella duplice ottica tutoria dell'autorità giudiziaria e degli addetti all'amministrazione della giustizia, e informativa e conoscitiva per i cittadini[64].

In questo senso, il dovere di trasparenza imposto dovrebbe garantire l'utilizzo di un linguaggio accessibile e l'estrapolazione di risultati giurisprudenziali, normativi e processuali corretti per i cittadini, ma soprattutto per giudici, avvocati e professionisti del diritto.

Detto che i benefici derivanti da un tale uso sarebbero di assoluto pregio sotto tanti profili, non ne vanno sottaciuti i rischi ancora attuali[65].

In questo senso, gli strumenti di "riassumi con IA" o "semplifica con IA" che molte banche dati giuridiche online hanno già adottato permette un'estrazione della conoscenza certamente più accessibile per il cittadino, ma potenzialmente meno efficace e, in alcuni casi, incompleta per gli esperti.

In effetti, la lunghezza dei testi legislativi, delle decisioni giudiziarie e dei procedimenti amministrativi è già notevolmente complessa per gli esperti del settore ma l'uso dello strumento IA può certamente contribuire a creare un ecosistema di fiducia e promuovere un avvicinamento fra tutte le parti.

D'altronde la legge entra ogni giorno nella vita di ciascuno di noi e, anche per evitare spiacevoli inconvenienti[66], l'algoritmo che propone una soluzione tecnica dovrebbe comunque, in un'ottica di trasparenza, suggerire il confronto con il testo giuridico integrale o la consultazione di un esperto[67].

In tal senso, potrebbe essere utile lo sviluppo nel settore giuridico della c.d. "intelligenza artificiale spiegabile" (XAI – *Explainable Artificial Intelligence*), strumento che ha

---

62 M. R. ALLEGRI, *op. cit.*, p. 24.

63 In questo senso, si veda: M. FRANCAVIGLIA, *IA e funzioni giurisdizionali: alcune questioni preliminari alla luce del quadro costituzionale*, in *Rivista italiana di informatica e diritto*, n. 1/2025, pp. 1-16.

64 S. QUATTROCOLO, *Intelligenza Artificiale e processo penale: le novità dell'AI Act*, in *Diritto di Difesa, la Rivista dell'Unione delle camere penali italiane*, pp. 1-13.

65 G. UBERTIS, *Processo penale telematico, intelligenza artificiale e Costituzione*, in *Cassazione penale*, 2024, n. 2.

66 Il riferimento è alla vicenda avvenuta presso il Tribunale di Firenze, in cui un Avvocato inseriva all'interno di un proprio atto una serie di riferimenti giurisprudenziali errati. Tale episodio rientra nel più ampio fenomeno delle c.d. "allucinazioni di intelligenza artificiale".
Per una lettura completa della vicenda, si veda: https://www.altalex.com/documents/2025/04/01/sentenza-viene-inventata-a-i-responsabilita-ex-art-96-c-p-c.

67 G. VACIAGO (a cura di), *Intelligenza artificiale generativa e professione forense*, Milano, 2024.

l'obiettivo di migliorare l'interpretabilità dei modelli AI e che punta alla fiducia e alla sicurezza degli utenti[68].

Con tale locuzione si intende un insieme di tecniche pensate per rendere interpretabili gli algoritmi, offrendo una spiegazione chiara e standardizzata del percorso logico che conduce a un determinato *output*, chiarendo quali variabili hanno influenzato la decisione e con quale peso[69].

Condividendosi l'indirizzo dottrinale secondo cui in settori sensibili deve pretendersi un grado di comprensibilità e un livello di spiegazione molto alti, pare opportuno specificare che l'implementazione del modello di IA spiegabile nel settore della giustizia potrebbe avere riflessi importanti in nuove declinazioni del diritto costituzionale alla difesa[70].

In tal senso, pare ci sia ancora da fare per eliminare incertezza teorica e pratica di tali disposizioni.

La sfida è aperta e spetta agli Stati e alle istituzioni europee mettere in pratica i principi etici contenuti nel Regolamento[71], i quali, nell'attuale contesto tecnologico e di mercato, rischiano di rimanere inapplicati o, ancor peggio, fumosi.

5. *Conclusioni*

Il quadro normativo europeo in materia di *AI Act* ha certamente segnato un importantissimo passo in avanti nel tentativo di incardinare lo sviluppo tecnologico all'interno di un perimetro giuridico chiaro e strutturato.

Allo stesso tempo però, come messo in evidenza, la disciplina in oggetto sconta ancora un margine di indeterminatezza che mette a rischio la tutela dei diritti fondamentali garantiti dal Regolamento.

In questo senso, è chiaro a tutti, istituzioni e *stakeholders*, che il percorso di introduzione dei sistemi di Intelligenza Artificiale nel settore della giustizia non può essere immediato e passivo.

Al contrario, dovrà essere accompagnato da un processo di adattamento graduale di carattere culturale *in primis*, ma anche formativo e giuridico, per garantire il principio base dell'antropocentrismo.

---

[68] Per un approfondimento sulle varie tecniche messe a punto da coloro i quali si occupano di *Explainable IA* si suggerisce: A. LOREGGIA, G. SARTOR, *L'Intelligenza Artificiale nella moderazione del digitale*, in *Sistemi Intelligenti*, pp. 53-57; P. BUONO, G. DESOLA, R. LANZILOTTI, *Interazione con sistemi intelligenti*, in V. V. CUOCCI, F. P. LOPS, C. MOTTI (a cura di), *La responsabilità civile nell'era digitale: atti della* Summer school *2021*, Bari, 2022, p. 115 ss.

[69] G. CONTISSA, G. LASAGNI, G. SARTOR, *Quando a decidere in materia penale sono (anche) algoritmi e IA: alla ricerca di un rimedio effettivo*, in *Dirittodiinternet.it*, n. 4/2019.

[70] C. CASONATO, *Costituzione e intelligenza artificiale: un'agenda per il prossimo futuro*, in *BioLaw Journal – Rivista di BioDiritto, Special Issue*, n. 2/2019, pp. 722-724. L'Autore inquadra tre importanti diritti di rango costituzionale, in particolare: «a) […] ottenere una spiegazione dei passaggi attraverso i quali la macchina ha generato il proprio risultato; b) […] essere resi consapevoli della natura, umana o artificiale, del proprio interlocutore; c) […] essere destinatari di decisioni che siano il risultato di un processo in cui sia presente una significativa componente umana».

[71] Strategia europea in materia di giustizia elettronica 2024-2028, in *GUUE*, C/2025/437, del 16 gennaio 2025.

Pare opportuno riferire che nei primi mesi del 2025, in ossequio all'articolo 96 del Regolamento, la Commissione ha pubblicato la bozza delle Linee Guida afferenti alle pratiche vietate di cui all'articolo 5[72].

Tale disposizione impone alla Commissione europea di procedere nelle medesime forme anche sulla disciplina relativa all'utilizzo dei sistemi identificati ad «alto rischio» dall'art. 6 del Regolamento.

Interessante dunque sarà comprendere come la Commissione suggerirà di procedere nell'ambito dell'amministrazione della giustizia per garantire i diritti fondamentali dinanzi al pericolo della disumanizzazione.

Infatti, il rischio potrebbe essere che l'introduzione dei modelli di IA a supporto del giudice, in assenza di linee guida chiare e adattabili a tutti gli Stati membri, determini un *vulnus* di tutela negli ordinamenti giuridici europei, con importanti conseguenze sull'accesso alla giustizia dei cittadini e sul rispetto dei principi fondamentali delle persone imputate[73].

In questo senso, si è voluto sottolineare il ruolo del principio di trasparenza, spesso evocato all'interno del Regolamento, in quanto ritenuto un ottimo punto di partenza per la chiarezza comunicativa, l'accessibilità tecnica e la garanzia del rispetto dell'eticità su cui è imperniata la disciplina.

Tuttavia, il rischio, come sottolineato in apertura e ripreso dalla provocazione "dallesca", è quello di trovarsi di fronte a un'infrastruttura giuridica tanto potente quanto opaca, una "anidride solforosa" normativa che illude con la promessa della modernità ma che, in mancanza di un'effettiva applicazione dei suoi principi, potrebbe accentuare le distorsioni anziché ridurle.

[72] Commissione europea, Annex to the Communication to the Commission, Approval of the content of the draft Communication from the Commission - Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (*AI Act*), del 4 febbraio 2025, C(2025) 884 final**.**

[73] L. PALMIERI, *Il deposito telematico tra esigenze efficientiste e garanzie difensive*, in *Diritto penale e processo*, n. 4/2024.