

## ORIENTAMENTI

---

**PIER PAOLO PAULESU**

### **Intelligenza artificiale e giustizia penale Una lettura attraverso i principi**

In un futuro che è già presente, l'intelligenza artificiale è destinata a costituire una risorsa preziosa in termini di efficacia e di efficienza della giustizia penale. Può contribuire a risolvere problemi investigativi, problemi cautelari, problemi probatori, problemi decisorii. Non c'è settore del processo penale che non ne possa ricavare una significativa utilità. A patto, naturalmente, di evitare striscianti e pericolose derive tecnocratiche che possono mettere in pericolo l'effettività di alcuni valori fondamentali. Di qui la necessità di assicurare il costante equilibrio tra la dimensione digitale e quella umana. Scopo del presente lavoro è analizzare l'impatto dei sistemi di *AI* su alcuni snodi peculiari del processo penale, verificando se e come alcuni principi fondamentali (presunzione di non colpevolezza, imparzialità del giudice, libertà personale, parità di armi, contraddittorio) su cui si regge il nostro sistema possano adattarsi alle nuove sfide tecnologiche senza perdere forza e incisività.

*Artificial intelligence and criminal trial. An overview through the principle*

*In a future that is already here, artificial intelligence is destined to be a valuable resource in terms of the effectiveness and efficiency of criminal justice. It can help to solve investigative problems, precautionary problems, evidentiary problems, decision-making problems. There is no area of the criminal trial that cannot benefit significantly from it. Provided, of course, that we avoid creeping and dangerous technocratic drifts that may endanger the effectiveness of certain fundamental values. Hence the need to ensure the constant balance between the digital and the human dimension. The purpose of this paper is to analyse the impact of AI systems on some specific aspects of the criminal trial, examining whether and how some fundamental principles (presumption of not guilty, impartiality of the judge, personal freedom, equality of arms, adversarial) must be able to adapt to new technological challenges without losing strength and effectiveness.*

**SOMMARIO:** 1. Premessa. - 2. *Segue.* Supervisione del giudice. Autonomia del giudice. Trasparenza dei metodi e dei risultati. Motivazioni rafforzate. - 3. Cautele. - 4. Prove. - 5. Decisioni.

1. *Premessa.* In un futuro che è già presente, gli strumenti di intelligenza artificiale sono destinati a costituire una preziosa risorsa tecnologica in termini di efficienza e di efficacia della giustizia penale. Possono infatti contribuire a risolvere problemi investigativi, problemi cautelari, problemi probatori, problemi decisorii. Non c'è settore del processo penale che non ne possa ricavare una significativa utilità. Con riserva di ritornare sul punto, preme qui subito segnalare come l'implementazione dell'intelligenza artificiale in ambito processuale tenda a realizzarsi principalmente su due piani: la valorizzazione delle cd. *automated evidence* di ultima generazione; il ricorso alla cd. "giustizia predittiva":<sup>1</sup> una endiadi, quest'ultima, che latamente allude all'impiego

---

<sup>1</sup> Cfr. nell'ambito di una letteratura che sta diventando via via sempre più vasta, CONTISSA, LASAGNI, SARTOR, *Quando a decidere in materia penale sono (anche) gli algoritmi e IA: alla ricerca di un rimedio effettivo*, in *Riv. trim. Diritto di Intenet*, n. 4 2009, 619 s. GIALUZ, *Quando il processo penale incon-*

dell'*AI* per ricavare, attraverso l'analisi di un numero significativo di provvedimenti giudiziari, previsioni attendibili sull'esito di alcune tipologie di controversie<sup>2</sup>. Con intuibili ricadute extraprocessuali che, a loro volta, possono condizionare le scelte degli operatori del diritto: la disponibilità di *open data* sugli esiti dei procedimenti penali potrebbe infatti spingere le cd. *Legaltech startup* ad immettere sul mercato *software* predittivi a vantaggio soprattutto degli studi legali, che potrebbero così orientare al meglio le strategie difensive<sup>3</sup>. Riassumendo, quando si affronta il tema dell'impatto dei sistemi di intelligenza artificiale sul processo penale<sup>4</sup> è importante distinguere tra *strumenti predittivi della IA ai fini decisori*, suscettibili di fornire dati che possono figurare come base per un provvedimento giurisdizionale (ordinanze cautelari, provvedimenti di archiviazione, decreti di citazione a giudizio, sentenze di rito, sentenze di merito) e *strumenti di predizione decisoria* basati sull'*AI*, finalizzati a prevedere l'esito di future decisioni su casi simili<sup>5</sup>.

Tanto premesso, veniamo ora al punto nodale che qui maggiormente interessa, sintetizzandolo con un quesito. Come gestire la risorsa dell'*AI* rispettando la tavola dei valori (costituzionali e sovranazionali) che caratterizza il "giusto processo" (parità di armi tra le parti, libertà personale, diritto di difesa, presunzione d'innocenza, contraddittorio, imparzialità del giudice)? Riaffiora, così, nella sua dirimpente problematicità, anche in questo innovativo settore, la forte tensione che attraversa il processo penale moderno: conciliare efficienza e garanzie (in questo caso, innovazione tecnologica e principi). Con una peculiarità di cui occorre essere pienamente consapevoli, perché cela più

---

tra l'intelligenza artificiale: luci ed ombre dei risk assessment tools tra Stati Uniti ed Europa, in *Diritto penale contemporaneo*, 2019, 1 s.; KOSTORIS, *Predizione decisoria, diversione processuale e archiviazione*, in *Sistema penale*, 2021, 1 s.; LUPARIA, *Prova giudiziaria e ragionamento artificiale: alcune possibili chiavi di lettura*, in *Il concetto di prova alla luce dell'intelligenza artificiale* a cura di Sallantin e Szczeciniarz, Milano, 2005, XIV; QUATTROCOLO, *Equità del processo penale e automated evidence alla luce della Convenzione europea dei diritti dell'uomo*, in *Revista Italo-Espanola de Derecho Procesal*, 2019, 2 s. EAD, *Artificial Intelligence. Computational Modelling and Criminal Proceedings. A Framework for a European Legal Discussion*, Springer, Berlin, 2020; MANES, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in *Discrimen*, 15 maggio 2020, 1. RICCIO, *Ragionando su intelligenza artificiale e processo penale*, in *questa Rivista*, 2019, n. 3, 1 s.; SIGNORATO, *Giustizia penale e intelligenza artificiale. Considerazioni in tema di algoritmo predittivo*, in *Riv. dir. proc.*, 2020, 605. EAD., *Il diritto a decisioni penali non basate esclusivamente su trattamenti automatizzati: un nuovo diritto derivante dal rispetto della dignità umana*, ivi, 2021, 101; UBERTIS, *Intelligenza artificiale, giustizia penale, controllo umano significativo*, in *Sistema penale*, 2020, 1 s.

<sup>2</sup> Cfr., per tutti, VIOLA, voce *Giustizia predittiva*, in *Enc. Giur. Treccani, Diritto on line* (2018), 1.

<sup>3</sup> Lo sottolinea GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale*, cit., 3 nota 10, richiamando, in particolare, l'esperienza francese della *startup Predictive*.

<sup>4</sup> Restringendo l'analisi allo specifico settore giurisdizionale che qui interessa.

<sup>5</sup> Cfr. KOSTORIS, *Predizioni decisorie*, cit., 3.

di una insidia. Sappiamo che, nel campo dello sviluppo tecnologico, l'intelligenza artificiale ambisce, per la sua stessa natura, a porsi come surrogato del pensiero (critico) umano. Sotto il profilo strettamente giuridico, poi, lo scenario che si staglia all'orizzonte è quello di un «sistema oracolare *legal tech*»<sup>6</sup>: un sistema che, nel processo penale, può incentivare il trasferimento della *responsabilità del giudizio e dell'errore* dall'uomo all'algoritmo. Un passaggio di consegne suscettibile di realizzarsi in alcuni snodi procedurali particolarmente esposti alla fallacia valutativa: le prognosi cautelari (artt. 274 e 275 c.p.p.); il controllo del pubblico ministero sulla fondatezza della notizia di reato ai fini della richiesta di archiviazione (ci si riferisce, in particolare, al giudizio prognostico *ex art. 125 disp. att. c.p.p.*); il controllo del giudice sulla solidità degli elementi acquisiti in udienza preliminare ai fini della sentenza di non luogo a procedere (la prognosi imposta dall'art. 425 c.p.p.); l'applicazione delle regole di giudizio dibattimentali; i controlli sulla fondatezza/infondatezza dei motivi d'impugnazione. Di qui una possibile, problematica, conseguenza: l'appiattimento sul *cogito* artificiale di chi amministra la giustizia penale. Con il pericolo che in questo modo si finisca per dar vita a una sorta di torsione tecnologico/efficientista del processo a scapito delle sue peculiari funzioni di controllo e di garanzia: funzioni che dovrebbero restare invece saldamente e prudenzialmente affidate a «giudizi di valore» umani, orientati e sorretti da una forte componente intuitiva, dal beneficio del dubbio, dalla consapevolezza dell'errore. Parametri difficilmente replicabili da una macchina e che sembrerebbero peraltro sottesi al solido legame tra libero convincimento e motivazione (artt. 111 Cost; artt. 192 e 546 c.p.p.) che, come è noto, costituisce la cifra stilistica del nostro sistema processuale penale. Prima il giudice si auto-convince liberamente secondo i criteri della logica (principio di non contraddizione), dell'esperienza, della ragionevolezza, della proporzionalità, dell'equità (due giudici diversi, basandosi sulle stesse prove, potrebbero legittimamente giungere a conclusioni opposte). Dopodichè, lo stesso giudice, attraverso la motivazione, deve convincere le parti (nel processo) e la collettività (all'esterno del processo) circa la correttezza del ragionamento adottato (premesse e conclusioni). Questo stretto legame tra libero convincimento e motivazione parrebbe insomma esprimere plasticamente

---

<sup>6</sup> L'espressione è di MANES, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, cit. 6. S'interrogano sulla possibilità che in futuro la macchina possa sostituire integralmente il giudice, TRAVERSI, *Intelligenza artificiale applicata alla giustizia: ci sarà un giudice robot?*, in *Quest. giust.*, 10 aprile 2019, 1 s; nonché LUCIANI, *La decisione giudiziaria robotica*, in *Rivista AIC*, 2018, 872 s.

l'idea di un giudizio umano, fallibile, ma sempre controllabile *ex post*. Una idea di giudizio che, in fondo, sembrerebbe la più adatta a gestire e a salvaguardare i valori di prima grandezza coinvolti nel processo penale.

Eppure, ora che l'intelligenza artificiale incrocia sempre più spesso la giustizia penale, viene spontaneo domandarsi se quel legame tra libero convincimento e motivazione possa risulterne in qualche modo alterato, indebolito. Detto altrimenti: attribuire in chiave decisoria un peso eccessivo ai dati forniti dall'algoritmo potrebbe, nella prassi giudiziaria, segnare (più o meno surrettiziamente) il passaggio verso una sorta di moderna *prova legale tecnologica*? È possibile che quest'ultima venga troppo spesso considerata la *miglior prova* (date le informazioni disponibili) apparsa sulla scena del processo, portando così a sottostimare il rischio dell'errore giudiziario? Emerge qui, infatti, uno dei nodi più complessi in materia di *AI*. Ci si riferisce, in particolare, alla concreta possibilità (soprattutto in chiave difensiva) di confutare il risultato elaborato dalla macchina, posto che una operazione di questo tipo, per risultare davvero efficace, presuppone l'accesso al codice sorgente che gestisce l'algoritmo. Accesso che può trovare un serio ostacolo nelle istanze di tutela del segreto commerciale del programma informatico. Come contestare il risultato prodotto dall'algoritmo se non si ha la piena contezza del metodo computazionale adottato<sup>7</sup>? Diffusa è ormai, a questo proposito, l'espressione *black box decision*<sup>8</sup>, per denotare il rischio che la motivazione di una pronuncia basata su algoritmi possa risultare incompleta, oscura, non facilmente intellegibile da parte di chi fosse interessato a criticarla. Di qui la difficoltà di ricostruire esattamente a posteriori il ragionamento del giudice di fronte al dato elaborato dal sistema di *AI*.

Per quanto concerne, poi, lo specifico profilo delle predizioni decisorie, non sfugge come i dati forniti dall'*AI* siano in grado di incidere sullo sviluppo del

---

<sup>7</sup> Sul punto, QUATTROCOLO, *Quesiti nuovi e «soluzioni antiche»? Consolidati paradigmi normativi vs. rischi e paure della giustizia digitale «predittiva»*, in *Cass. pen.* 2019, 1758, riprendendo le riflessioni di BURREL, *How machine think: Understanding opacity in machine-learning algorithms*, in *Big Data and Society*, 2016, col. I, 1 s. Senza indugiare troppo in dettagli tecnici, del resto, se si guarda al peculiare settore del *machine learning*, che si occupa di creare sistemi che autonomamente apprendono e migliorano le *performance* in base ai dati utilizzati (il sistema li elabora generando valutazioni e classificazioni sui nuovi dati che vengono via via successivamente inseriti), c'è un dato significativo da considerare: il codice sorgente rivela il funzionamento dell'algoritmo di autoapprendimento, è possibile conoscere i dati di *input* e *output*, ma risulta difficile comprendere la configurazione finale del modello creato, che costituisce la base delle decisioni assunte dal sistema.

<sup>8</sup> Cfr. sul punto FERGUSON, *Illuminating Black Data Policing*, in *Ohio St.J. Crim. L.*, 2018, 503; v anche CAIANIELLO, *Criminal Process faced with the Challenges of Scientific and Technological Development*, in *European Journal of Crime, Criminal Law and Criminal Justice*, 27, (2019), 265 s.

procedimento in chiave deflattiva (intesa in senso lato): basti pensare alla scelta tra l'azione e l'inazione penale (archiviazione); all'alternativa tra la sentenza di non luogo a procedere e il decreto di citazione a giudizio; alle strategie delle parti nella gestione di patteggiamenti, sospensioni con messa alla prova, giudizi abbreviati, decreti penali di condanna, concordati in appello.

In definitiva, l'uso massiccio di strumenti predittivi in chiave decisoria, da un lato, e l'eccessivo ricorso agli strumenti di predizione decisoria, dall'altro, potrebbero stringere il sistema in una morsa: nel primo caso, come s'è detto, la piena cognizione processuale che si realizza nel contraddittorio dibattimentale risulterebbe sì salvaguardata, ma potrebbe dipendere troppo dai dati forniti dalla macchina e sortire decisioni poco chiare: quindi, come si diceva, difficilmente contestabili; nel secondo caso, all'opposto, sarebbe proprio la macchina a favorire la fuga da quella piena cognizione processuale, spingendo verso l'inazione, verso la sentenza di non luogo a procedere, verso i riti speciali deflattivi del dibattimento.

Ci si rende conto che ipotizzare scenari cognitivi pervasi da automatismi valutativo/decisori che, secondo una idea quasi deterministica del processo, inseguono (invano) il falso mito di una giustizia digitale perfetta può forse suonare eccessivo. Soprattutto in un ordinamento, come il nostro, gestito dal principio (fortemente identitario) della soggezione del giudice solo alla legge (art. 101 co. 2 Cost.): un principio che, di per sé, dovrebbe impedire soluzioni automatizzate troppo vincolanti basate su strumenti non riconducibili alla produzione legislativa<sup>9</sup>. Su queste basi, in fondo, si tratterebbe di assicurare il giusto equilibrio tra la dimensione digitale e quella umana, evitando conflitti tra queste due componenti, che dovrebbero invece agire sinergicamente, all'interno di una solida visione antropocentrica<sup>10</sup>. Una visione che sembra appartenere alle stesse fonti normative europee, perlomeno se s'intende il *diritto di accesso al giudice* di cui parla l'art. 5 C.E.D.U. come diritto dell'accusato di interloquire con una persona fisica. Sempre sul piano europeo, conforta peraltro notare come l'art. 11 della Direttiva UE 216/680 vieti decisioni fondate esclusivamente su trattamenti automatizzati, riconoscendo «il diritto di ottenere l'intervento umano da parte del titolare del trattamento»<sup>11</sup>. Ragionevole cogliere, in quest'ultima espressione, sia il dovere del giudi-

<sup>9</sup>Cfr. KOSTORIS, *Predizioni decisorie*, cit., 4, secondo cui l'algoritmo è una «entità esterna ed estranea alla legge».

<sup>10</sup> Cfr. KOSTORIS, *Predizioni decisorie*, cit., 20; MANES, *L'oracolo algoritmico*, cit., 22; RICCIO, *Ragionando su intelligenza artificiale e processo penale*, cit., 9.

<sup>11</sup> Per approfondimenti, v. SIGNORATO, *Il diritto a decisioni penali non basate esclusivamente su trattamenti automatizzati*, cit., 106.

ce di intervenire personalmente, sia il dovere del giudice di valutare l'*output* del sistema automatizzato unitamente ad altri elementi di prova che ne confermino l'attendibilità<sup>12</sup>.

Con tutto ciò, sussistono almeno tre ragioni che orientano a non sottovalutare troppo i rischi di una strisciante degenerazione tecnocratica.

La prima. Sappiamo che la tecnologia informatica possiede una vocazione espansiva e pervasiva. Non è raro, peraltro, in questo specifico settore, riscontrare atteggiamenti di eccessiva fiducia nei confronti dei protocolli adottati e dei risultati ottenuti. Salvo poi accorgersi successivamente della fallacia di certe presunte *best practices*. Difficile pensare che la giustizia penale non ne risulti problematicamente coinvolta. Con l'aggravante che, in questo caso, il prezzo da pagare sarebbe altissimo, perché, giova ribadirlo, sono in gioco valori di prima grandezza.

La seconda ragione. Di più ampio respiro. Predomina ormai, a livello politico (nazionale ed europeo), l'idea che la rapidità della giustizia penale costituisca l'obiettivo primario su cui tarare il livello di affidabilità e di solidità di uno Stato economicamente avanzato. Inquietata, nel dibattito sull'efficienza e sulla semplificazione del passaggio giurisdizionale, l'attuale pensiero *mainstream* che, in nome di una fortissima istanza di accelerazione dei processi (frutto di una visione esasperata del principio della durata ragionevole), tende spesso a collocare indistintamente sullo stesso piano la giustizia penale, la giustizia civile, la giustizia amministrativa, la giustizia tributaria, trascurando, così, le differenze legate al diverso peso specifico dei valori coinvolti. Per rendersene conto, è sufficiente dare uno sguardo alle ultime riforme in materia di processo penale. Il *fattore tempo* costituisce, ormai, anche in questo delicatissimo contesto, un obiettivo assoluto, apicale, assorbente<sup>13</sup>. Ebbene, se la drastica contrazione temporale dei processi costituirà davvero la priorità nelle future politiche criminali, come sembra ormai evidente, ipotizzare un massiccio ricorso al *problem solving* comodamente gestito dall'algoritmo è tutt'altro che irrealistico. Perché assicurerebbe anche un significativo abbattimento delle spese di giustizia.

La terza ragione. Si assiste ormai da tempo ad una progressiva perdita di legittimazione della giustizia penale. Si lamentano errori, abusi, mancanza di pro-

---

<sup>12</sup> *Infra* par. 4.

<sup>13</sup> Basterebbe solo riflettere sul singolare meccanismo dell'improcedibilità per superamento dei termini di durata massima del giudizio di impugnazione (art. 344 *bis* c.p.p.). Criticabile o no, quel meccanismo segna la fine del processo penale per ragioni squisitamente cronologiche; poco importa che l'imputato sia davvero innocente o colpevole: il problema dell'accertamento, ad un certo punto, viene utilitaristicamente rimosso.

fessionalità, scarsa specializzazione, lacune strutturali e organizzative. Il *trend* è in crescita e sembra inarrestabile. In questo clima di sfiducia nel prodotto giustizia, non è escluso che l'opinione pubblica possa trarre maggiori sicurezze da decisioni automatizzate basate sull'*AI*, perché ritenute più neutrali, razionali, prevedibili<sup>14</sup>. Sicurezze che i giudici, dal canto loro, potrebbero essere proclivi a fornire, puntando prudenzialmente sull'*AI* per mettersi al riparo da eventuali censure esterne (ci si riferisce soprattutto ai processi ad alta visibilità mediatica). Se è vero, infatti, che il giudice deve anche *apparire imparziale*, il ricorso all'algoritmo potrebbe utilitaristicamente servire davvero allo scopo.

Fondamentale è, su queste basi, insistere su un punto fermo, che deve possedere la forza dell'assioma: la vigilanza assoluta e costante del giudice sul rispetto di alcuni principi fondamentali, che rappresentano il primo e il più solido argine alle derive tecnocratiche. Principi che vanno certamente armonizzati, adattati, modulati, alla luce del peculiare contesto procedimentale e delle risorse digitali di volta in volta fruibili del caso concreto: mai però piegati allo sviluppo tecnologico. Fornire una visione d'insieme dell'impatto dell'intelligenza artificiale su alcuni snodi cruciali del processo penale utilizzando proprio la chiave di lettura dei principi è lo scopo delle pagine che seguono. Con una avvertenza di tipo metodologico, che nasce dalla consapevolezza di affrontare una materia (per il giurista) tecnicamente complessa e fitta di incognite, perché condizionata dallo sviluppo delle scienze informatiche. Di qui, a volte, come si vedrà, la precisa scelta di seminare dubbi nell'impossibilità, al momento, di raccogliere certezze.

2. *Segue. Supervisione del giudice. Autonomia del giudice. Trasparenza dei metodi e dei risultati. Motivazioni rafforzate.* Volgendo lo sguardo alla concreta esperienza giudiziaria, è inevitabile porsi un quesito che costituisce il filo rosso del presente lavoro: qual è il confine tra l'uso e l'abuso tecnologico? Come capire se, in un determinato processo, l'intelligenza artificiale rappresenta un prezioso supporto alla decisione giudiziale o ne costituisce, invece, la ragion d'essere? Più precisamente. Come verificare se il sistema di *AI* concretamente adottato serve solo ad affinare e a completare il giudizio umano, (posto che quest'ultimo si regge su dati di esperienza limitati, neppure paragonabili a quelli gestibili da una macchina), o se sospinge invece inesorabilmente il giudice verso un *certo tipo di decisione*? Si tratta, insomma, d'individuare il confine tra la valutazione/responsabilità del giudice e la valutazio-

---

<sup>14</sup> Lo sottolinea MANES, *L'oracolo algoritmico*, cit., 11.

ne/responsabilità della macchina. E posto che, come s'è detto, l'analisi *ex post* della motivazione potrebbe non bastare a fugare i dubbi sul peso specifico dell'intelligenza artificiale nel caso concreto, il primo passo consiste nell'adottare metodi chiari e condivisi. In questa prospettiva, la strada maestra è sintetizzabile in quattro punti cardinali. *Supervisione del giudice. Autonomia del giudice. Trasparenza dei metodi e dei risultati. Motivazioni rafforzate.* La supervisione del giudice deve restare centrale e costituire, quindi, un passaggio ineludibile: perché «le qualità umane non potranno mai essere sostituite da componenti artificiali»<sup>15</sup>.

Nella sua attività di controllo, poi, il giudice deve conservare una piena e costante autonomia rispetto al sistema di intelligenza artificiale di volta in volta adottato, senza risultarne passivamente condizionato. Da questo specifico punto di vista, lo si può rilevare incidentalmente, sembrano davvero aprirsi interessanti prospettive d'indagine sul concetto stesso d'imparzialità del giudice di fronte alle nuove risorse digitali.

Ancora. Va garantita la tracciabilità dei metodi e dei risultati ottenuti: tutti i passaggi tecnici vanno rigorosamente spiegati e documentati; solo così è possibile comprendere, *a posteriori*, i motivi per cui una decisione basata sull'intelligenza artificiale è stata adottata correttamente o erroneamente (in ogni momento può emergere un difetto o un nuovo dato che impone di rivalutare/sostituire/aggiornare il sistema: in fondo, a ben vedere, nella sostanza, si tratta di applicare il metodo documentale della cd. *chain of custody* già ampiamente sperimentato nell'esperienza anglosassone; un metodo finalizzato a consentire alle parti di verificare la genuinità dei dati informatici: ogni singolo passaggio della *digital evidence* deve essere tracciabile e rintracciabile<sup>16</sup>).

Infine, occorre assicurare pronunce giurisdizionali sorrette da *motivazioni analitiche*, chiare, qualitativamente e quantitativamente esaustive (artt. 111 comma 6 Cost., 192 comma 2, 546 co. 1 lett. e). In particolare, il giudice dovrebbe illustrare puntualmente le ragioni che lo spingono verso conclusioni

<sup>15</sup> Così CASONATO, *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in *Dir. publ. comparato ed europeo*, 2019, 124, il quale precisa come «la sana consapevolezza di sapere di non sapere sono caratteristiche che contraddistinguono l'umano e la sua ricerca di senso, le quali mal si attagliano ai ragionamenti dell'AI».

<sup>16</sup> Sulla *chain of custody*, in generale, cfr. BEAUFORT-MOORE, *Crime Scene Management and Evidence Recovery*, Oxford, 2009, 10 s. GIANNELLI, *Chain of Custody and the handling of real evidence*, in *American Criminal Law Review*, 1983, 20; con particolare riferimento anche alle indagini informatiche, CHELO, *Le prime indagini sulla scena del crimine. Accertamenti e rilievi urgenti di polizia giudiziaria*, Padova, 2014, 19 s. e, diffusamente, SIGNORATO, *Le indagini digitali. Profili strutturali di una metamorfosi investigativa*, Torino, 2018, 140 s. Sulla prova digitale, in particolare v. il quadro d'insieme tracciato da DANIELE, *La prova digitale*, in *Riv. dir. proc.* 2011, 286 s.

che l'algoritmo smentisce, o, all'opposto, spiegare perché intende conformarsi ai risultati ottenuti tramite il sistema di *AI*. La comparazione tra giudizio umano e «giudizio artificiale» deve insomma costituire un passaggio ineludibile. E in questa comparazione un ruolo centrale la questione della «scientificità» dell'algoritmo (secondo i noti criteri della verificabilità e della falsificabilità della teoria posta a fondamento del sistema).

Segnali rassicuranti sembrano provenire dalla Carta etica europea per l'impiego giurisdizionale dell'*AI*<sup>17</sup>, dove è presente, anzitutto, il richiamo all'esigenza di assicurare il *diritto di accesso al giudice* e il diritto ad un *processo equo* (parità delle armi, contraddittorio). Viene poi ribadito il divieto di ogni *discriminazione* con riferimento ai dati sensibili (salute, razza, etnia, religione, orientamento sessuale ecc.), genetici, biometrici, considerata la capacità dei sistemi di *AI* di elaborarli (basti qui segnalare un aspetto problematico su cui torneremo in seguito: le possibili discriminazioni in sede di applicazione delle misure cautelari)<sup>18</sup>. Non manca, poi, nella Carta etica, il richiamo *alla qualità e alla sicurezza dei dati* di volta in volta impiegati, che dovrebbero sempre provenire da fonti rigorosamente certificate. Sotto questo profilo, la Carta etica parrebbe implicitamente ma inequivocabilmente alludere proprio alla *chain of custody* precedentemente ricordata: i modelli e gli algoritmi andrebbero memorizzati ed eseguiti in ambienti sicuri per preservare l'integrità del sistema. Ancora. Significativo è il riferimento, presente nella Carta *etica*, alla *completa trasparenza tecnica* (e alla conseguente necessità di creare autorità pubbliche indipendenti che valutino e certifichino anticipatamente gli strumenti, monitorandone il costante funzionamento), e alla *responsabilità degli utenti del sistema*: gli operatori del diritto che si avvalgono dell'algoritmo non dovrebbero risultare troppo vincolati alle soluzioni offerte da questa tecnologia, potendo criticamente discostarsene (come si vede, emerge qui, in tutta la sua problematicità, la consapevolezza del rischio, precedentemente segnalato: l'eccessivo appiattimento dell'uomo sulla macchina); i destinatari delle decisioni fondate sull'*AI*, dal canto loro, dovrebbero sempre ottenere

<sup>17</sup> Adottata dalla Commissione per l'efficienza della giustizia (CEPEJ) e rivolta ai «*public and private stakeholders responsible for the design and deployment of artificial intelligence tools and services that involve the processing of judicial decisions and data*» e ai «*public decision-makers in charge of the legislative or regulatory framework, of the development, audit or use of such tools and services*». Sul punto cfr. le osservazioni di QUATTROCOLO, *Intelligenza artificiale e giustizia: nella cornice della Carta etica europea gli spunti per una urgente discussione tra scienze penali ed informatiche* in [www.la.legislazionepenale.it](http://www.la.legislazionepenale.it), 18 dicembre 2018.

<sup>18</sup> Sul tema della protezione dei dati personali nella Carta etica si rinvia a GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale*, cit., 15 s.

consulenza legale e piena ed effettiva tutela giurisdizionale *ex art. 6 CEDU*. Infine, la Carta etica (appendice) insiste sul rispetto della parità delle armi e della presunzione d'innocenza, riconoscendo all'interessato il diritto, tramite l'accesso al giudice, di *contestare la validità scientifica dell'algoritmo* e il *peso di volta in volta attribuito allo stesso* nella singola vicenda giudiziaria.

3. *Cautele*. Come già anticipato, il settore cautelare è destinato a costituire uno dei principali terreni di elezione dell'*AI*. L'uso di algoritmi predittivi può infatti risultare utile al fine d'individuare la pericolosità individuale e la capacità di recidiva di un determinato oggetto, rappresentando, quindi, anche un valido supporto nella scelta della misura cautelare più idonea a soddisfare le istanze del caso concreto.

Forte è, però, su questo terreno il rischio di iniquità.

Emblematica, da questo punto di vista, l'esperienza statunitense dei cd. *risk assessment tools*, da cui il giurista italiano può e deve ricavare preziosi insegnamenti. Si tratta, come è noto, di strumenti basati su algoritmi che consentono di gestire e di analizzare un numero elevato di dati statistici su una persona o su gruppi di persone, selezionando fenomeni ricorrenti (*pattern*) e fornendo, così, la percentuale di rischio che un soggetto possa sottrarsi al processo o commettere reati. Come è stato opportunamente segnalato<sup>19</sup>, il ricorso a questi algoritmi nelle diverse fasi del procedimento penale statunitense (*pretrial process, sentencing, parole, probation*) sta diventando via via sempre più frequente, la necessità di effettuare valutazioni prognostiche sulla pericolosità degli imputati. Tra gli algoritmi predittivi creati e gestiti da privati, il più noto è il COMPAS (*Correctional Offender Management Profiling for Alternative Sanction*), un sistema di profilazione che analizza le risposte fornite ad un questionario su alcuni specifici temi: *criminal involvement, relationships/lifestyles, personality/attitudes, family and social exclusion*. A dispetto della sua oggettiva utilità, il COMPAS è avversato da chi lo considera alla stregua di una pericolosa *bias machine* foriera di pesanti discriminazioni di

---

<sup>19</sup> Cfr. ancora GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale*, cit., 4, cui si rinvia per approfondimenti sui *risk assesment tools* e per le numerose fonti bibliografiche statunitensi citate (v. in particolare, HUQ, *Racial equity in Algorithmic Criminal Justice*, in *Duke Law Journal*, 2019, 1043 s.; GARRETT, MONAHAN, *Judging Risk*, in *Californian Law Review*, *Forthcoming*, 10-11). Interessante è, sempre in ordine all'esperienza statunitense, la distinzione tra algoritmi predittivi che analizzano fattori di rischio *statici* (genere, età del soggetto al momento del primo arresto) e algoritmi predittivi che riguardano fattori di rischio *dinamici* (lavoro, utilizzo di sostanze stupefacenti o psicotrope), e tra algoritmi predittivi *statali* (elaborati da alcuni governi: ad esempio, Virginia, Alabama, California) e algoritmi predittivi *privati* (ad esempio, il COMPAS).

tipo soprattutto razziale<sup>20</sup>; con l'aggravante che, trattandosi di un metodo brevettato da una azienda privata, quindi segreto, viene meno il requisito di trasparenza che, come s'è detto, dovrebbe sempre accompagnare l'uso dell'*AI* nel contesto giurisdizionale. Nondimeno, nel celebre caso *State v. Loomis*, la Corte Suprema del *Wisconsin* ha escluso che l'uso di quel programma possa di per sé determinare una lesione del diritto dell'imputato all'equo processo<sup>21</sup>: a patto che i dati elaborati *software* non costituiscano l'unica base cognitiva della sentenza. In *State v. Loomis*, infatti, i giudici avevano prudenzialmente attinto ad una ampia gamma di fonti probatorie (*other supporting factors*), evitando così di basare il proprio convincimento solo sui dati digitali.

Volgendo lo sguardo al nostro processo penale, è inevitabile concentrare subito l'analisi sull'art. 274 c.p.p., che, lo sappiamo, è il fulcro delle analisi predittive cautelari: pericolo d'inquinamento probatorio, pericolo di fuga, pericolo di reiterazione del delitto.

Ebbene, il ricorso all'algoritmo predittivo in questo delicato contesto prognostico può anzitutto sollevare problemi di compatibilità con il fondamentale principio di eguaglianza. Coglie nel segno chi avverte come l'algoritmo sia per sua natura antieguagliario<sup>22</sup>, perché evidenzia solo alcuni fattori di rischio, omettendone altri; di conseguenza, considera in via presuntiva alcuni soggetti più pericolosi di altri. Su tutto, incombe poi la grande incognita della cd. *bias machine* già segnalata: e cioè che le decisioni automatizzate possano essere adottate sulla base di pregiudizi (etnico-razziali, morali, sessuali, ambientali, legati al *background* socioeconomico) veicolati all'interno della macchina dagli stessi programmatori. Su queste basi, appaiono fondate le preoccupazioni di chi considera i provvedimenti fondati sulla «calcolabilità del rischio» e della pericolosità individuale suscettibili di «aprire nuovi canali di discriminazione e di incarcerazione o di ostacolo alla scarcerazione per le fasce socialmente più deboli»<sup>23</sup>.

Spingendo l'analisi più a fondo, non v'è dubbio che il naturale terreno di elezione dell'intelligenza artificiale sia costituito dalla finalità cautelare special-

<sup>20</sup> Secondo alcune ricerche svolte negli Stati Uniti, ad esempio, le persone di colore risulterebbero classificate alla stregua di potenziali futuri criminali in misura pari al doppio delle altre etnie.

<sup>21</sup> *State v. Loomis*, 881 NW 2 d 749 (Wis 2016), par. 53-54. L'imputato, accusato di guida di veicolo senza il consenso del proprietario e di tentata violazione di un posto di blocco, era stato condannato a sei anni di reclusione e a cinque anni di *extended supervision* sulla base di una prognosi di alta pericolosità fondata, tra le altre cose, anche sui dati forniti dal sistema COMPAS. Sul punto v. in particolare, QUATTROCOLO, *Quesiti nuovi e «soluzioni antiche»?*, cit, 1758.

<sup>22</sup> Cfr. V. MANES, *L'oracolo algoritmico*, cit., 12.

<sup>23</sup> Così, SGUBBI, *Il diritto penale totale*, Bologna, 2019, 424.

preventiva rappresentata dal pericolo di reiterazione del delitto (art 274 lett. *c* c.p.p.). Qui il discorso si complica, perché, oltre al principio di eguaglianza, entra in gioco anche la presunzione di non colpevolezza intesa come regola di trattamento (art. 27 co. 2 Cost.), la quale, come è noto, vieta di equiparare l'imputato al colpevole prima della condanna definitiva. Infatti, mentre le finalità cautelari costituite dal pericolo di inquinamento probatorio (art. 274 lett. *a* c.p.p.) e dal pericolo di fuga (art. 274 lett. *b* c.p.p.) prendono in considerazione la sola «pericolosità processuale» dell'indagato/imputato, e non implicano necessariamente la formulazione di anticipati giudizi di colpevolezza, il discorso cambia se si guarda all'esigenza cautelare costituita dal pericolo di reiterazione del delitto. Qui si entra nella specifica dimensione specialpreventiva: c'è una prognosi di rischio percentuale in termini di futura e probabile (reiterata) responsabilità penale. Una prognosi che si regge, logicamente, su un duplice, anticipato giudizio di probabile colpevolezza: è probabile che l'indagato/imputato abbia commesso il reato per il quale si procede; è probabile che, durante il processo, l'indagato/imputato possa commettere altri reati. Di qui l'esigenza di limitare la sua libertà personale (art 13 Cost.). Censurabile o no in rapporto alla presunzione di non colpevolezza (non è certamente questa la sede per affrontare un tema così ampio e complesso), bisogna però avere l'onestà intellettuale di riconoscere che l'obiettivo perseguito è e resta la tutela collettiva, la difesa sociale. Un obiettivo che, difficile negarlo, rappresenta qualcosa di singolare, quasi un corpo estraneo al contesto giurisdizionale, che, lo sappiamo, mira verificare se un reato è stato commesso, non ad impedire che se ne commettano altri. Nei fatti, insomma, l'art 274 lett *c* c.p.p. consente di soddisfare istanze di profilassi criminale non prima del processo, non a margine del processo, ma con il processo, e, soprattutto, a causa del processo. Qui la materia delle cautele s'intreccia problematicamente con il profilo dell'*imputato pericoloso*: una endiadi quest'ultima, lo si può rilevare incidentalmente, che sembra sottendere una vera e propria antinomia concettuale, perchè la presunzione di non colpevolezza dovrebbe implicare anche la presunzione di non pericolosità (nel senso che il giudizio di pericolosità non dovrebbe mai precedere il giudizio sulla colpevolezza)<sup>24</sup>. È la stessa funzione cognitiva del processo ad apparire difficilmente compatibile con l'idea di una pericolosità presunta dell'imputato. Come è stato opportunamente rilevato, «quello che va accertato all'esito del processo, la responsabilità per il fatto

---

<sup>24</sup> Cfr. volendo, PAULESU, *La presunzione di non colpevolezza dell'imputato*, Torino (I ed., 2008), II ed. agg., 2009, 80; in precedenza, con riferimento al codice abrogato, ILLUMINATI, *La presunzione d'innocenza dell'imputato*, Bologna, 1979, 43.

tipico, non può essere considerato il presupposto per un trattamento deteriore del soggetto nei cui confronti si procede»<sup>25</sup>.

Ma veniamo al punto che qui interessa. La finalità cautelare specialpreventiva in questione presenta un taglio spiccatamente personalistico: non si chiede al giudice di valutare solo le specifiche modalità e circostanze del reato che giustificano la domanda cautelare del pubblico ministero, ma, contestualmente, si pretende un vaglio di tipo soggettivo che, per sua natura, può risultare troppo discrezionale e aleatorio. Si parla, infatti, testualmente, di valutazione della «personalità» dell'indagato/imputato, ricavabile da «comportamenti o atti concreti o dai suoi precedenti penali»: quest'ultima espressione si riferisce ad *elementi diversi* da quelli che integrano il fatto-reato per cui si procede, solitamente *anteriori* a quello stesso fatto e rivelatori di una indole particolarmente proclive alla commissione di delitti. Vero è che il giudice è tenuto ad effettuare una valutazione complessiva di tutti gli elementi, non potendo la sua predizione di pericolosità fondarsi solo sul fatto criminoso già ascritto, o solo sulla personalità dell'indagato<sup>26</sup>; ma non sfugge come quest'ultimo profilo, nel complesso di quella valutazione, tenda spesso ad assumere un peso significativo, a volte addirittura assorbente<sup>27</sup>.

Quale ruolo può giocare, qui, l'intelligenza artificiale?

Due le prospettive. La più radicale. Sbarrare la strada agli algoritmi predittivi basati su analisi statistiche fondate sulla personalità del soggetto, perché incompatibili con la presunzione di non colpevolezza. Il ragionamento potrebbe essere impostato nel seguente modo. Se un soggetto è presunto non colpevole, non dovrebbero pesare ai fini cautelari fattori di rischio di tipo strettamente psicologico/criminologico, come il carattere e la personalità, utilizzati in chiave predittiva di futuri delitti. A maggior ragione se frutto di analisi statistiche elaborate da un asettico algoritmo<sup>28</sup>. Diversamente si rischierebbe di riproporre, in chiave moderna, attraverso la tecnologia informatica, la nota teoria del «tipo criminologico d'autore»<sup>29</sup>. Nulla di più lontano dalla *ratio* sot-

<sup>25</sup> Cfr. MAZZA, *Le persone pericolose. In difesa della presunzione d'innocenza*, in *Dir. pen. cont.*, 2011, 1.

<sup>26</sup> Cfr. ad esempio, Cass., Sez. I, 14 novembre 1995, Armeli, in *Arch. nuova proc. pen.*, 1996, 640.

<sup>27</sup> In argomento, MONTAGNA, *I confini dell'indagine personologica nel processo penale*, Roma, 2013, 23, 91.

<sup>28</sup> Vale la pena sottolineare come la prevenzione cautelare (quindi la dimensione giudiziaria) possa essere fortemente influenzata dai risultati della profilassi di polizia (cd. *predictive policing*), che sfrutta *software* di riconoscimento facciale e di identificazione biometrica per mappare il rischio criminale. Così, ad esempio, il programma *Keycrime* frutto di esperienze investigative presso la questura di Milano, utilizzato con riferimento a crimi seriali (rapine, truffe, violenze sessuali ecc.).

<sup>29</sup> L'ordinamento vieta la perizia psicologica ma ammette quella psichiatrica (indagini su patologie psi-

tesa alla presunzione di non colpevolezza, che, legandosi al principio di legalità (art. 25 Cost), proietta nello specifico contesto processuale l'idea garantista e moderna di un diritto penale del fatto. L'imputato è sotto processo per aver tenuto una precisa condotta penalmente rilevante, non per ciò che è o è stato. Bisogna quindi evitare ogni sovrapposizione tra valutazione della personalità e la valutazione della responsabilità (in fondo, come si vede, è la stessa logica che sorregge il divieto di perizia psicologico/criminologica sull'imputato *ex* art. 220 c.p.p., frutto di una certa diffidenza verso alcune tipologie di sapere scientifico ritenute meno attendibili di altre).

Una soluzione di questo tipo, però, non convince pienamente, perché un approccio troppo rigido alla regola di trattamento sottesa alla presunzione di non colpevolezza, la quale, di per sé, non sembra costituire un ostacolo all'impiego di algoritmi predittivi nelle prognosi cautelari. Infatti, delle due l'una. O si coltiva la prospettiva draconiana di espungere totalmente dal sistema la finalità cautelare di cui all'art. 274 lett *c* c.p.p., perché ritenuta costituzionalmente eccezionale<sup>30</sup> o, più semplicemente, perché la si considera foriera di errori, iniquità, abusi da parte dei giudici. Oppure, una volta accolta utilitaristicamente l'idea (come è accaduto finora) secondo cui gli scopi cautelari special-preventivi non sarebbero di per sé incompatibile con la presunzione costituzionale (e «convenzionale»: art 6 C.E.D.U.), appare irragionevole impedire al giudice di attingere a tutte le risorse disponibili, compresa l'intelligenza artificiale. Se la profilassi criminale può e deve essere svolta (anche) in ambito cautelare, allora dovrebbe realizzarsi nel modo più completo. Peraltro, non è detto che gli algoritmi predittivi siano finalizzati ad effettuare una sorta di subdola, asettica e meccanica indagine psicologica sull'imputato. Si tratta di strumenti in grado di analizzare e di interpretare rapidamente soprattutto i fatti esterni all'uomo, per ricavarne giudizi predittivi: strumenti che possono non solo costituire un valido alleato alla valutazione del giudice, ma, ed è il punto nodale, fungere da efficace contrappeso tecnologico proprio ai temuti e criticati errori umani. Se il rischio temuto è quello di una massiccia

---

chiche). Poiché però il confine tra le due aree risulta davvero molto sottile (basti pensare alle differenze tra anomalie caratteriali e autentiche malattie mentali), ne deriva che il giudice gode di un significativo margine di discrezionalità. Non solo. Il divieto di perizia psicologica viene sostanzialmente aggirato nella prassi dalle cd. perizie psichiatriche allargate, dove al perito è richiesto d'indagare non solo su eventuali patologie, ma anche su disturbi psichici rilevanti per una migliore comprensione degli eventi. Cfr. per questi rilievi, CAMON, *Le prove*, in *Fondamenti di Procedura penale*, III ed., Padova, a cura di Camon-Cesari-Daniele-Di Bitonto-Negri-Paulesu, 2021, 342.

<sup>30</sup> Per approfondimenti sul punto, cfr, ancora, volendo, PAULESU, *La presunzione di non colpevolezza dell'imputato*, cit., 83 s.

profilazione automatica di tipo spiccatamente individualistico, troppo sbilanciata sulla persona, basterebbe bandire dalle predizioni cautelari solo la specifica tipologia di algoritmi ideati per questo specifico scopo e gli eventuali *bias* cognitivi che vi sono sottesi (ad esempio, i *software* che classificano i gusti sessuali, le abitudini di vita, o fenomeni di devianza in sé penalmente irrilevanti). Concludendo. Vietare aprioristicamente l'uso dell'*AI* nell'ambito delle prognosi cautelari è una prospettiva che non sembra trovare giustificazioni né sul piano della ragionevolezza, né sul piano dei principi fondamentali del processo. Si tratta, sostanzialmente, non lo si sottolinea mai abbastanza, di assicurare il giusto equilibrio tra il contributo della macchina e la supervisione critica del giudice, prestando attenzione ad evitare le *generalizzazioni statistiche anti-guagliarie* fornite dall'algoritmo: sia quelle che tendano a far risaltare troppo la figura dell'autore del reato a scapito di un vaglio approfondito del fatto storico descritto nella notizia di reato o nella imputazione (sbilanciando, così, la valutazione di cui all'art. 274 lett c c.p.p. troppo sul piano soggettivo); sia quelle suscettibili di dare troppo peso alle condotte di gruppi a cui l'interessato appartiene, è appartenuto, o in cui si riconosce; perché l'eccessivo rilievo dato all'interazione tra una determinata persona e il contesto socio-ambientale potrebbe sconfinare in una sorta di *labelling approach* sicuramente estraneo agli scopi cautelari. Questo sì, davvero incompatibile non solo con la presunzione di non colpevolezza, ma, ancor prima, con il principio della necessaria individualizzazione del trattamento cautelare, che quella stessa presunzione, in combinato disposto con gli artt. 2 e 13 Cost, indubbiamente sottende.

Merita in ogni caso aggiungere come, almeno con riferimento ad alcune tipologie di decisioni cautelari, l'eventuale spazio concesso ai sistemi di *AI* dovrebbe necessariamente sottostare ad una supervisione umana ancor più penetrante e accurata. Ci si riferisce, in particolare, alle diagnosi/prognosi in termini di adeguatezza e di proporzionalità della misura cautelare (art. 275 c.p.p.), che, per loro natura, richiedono supplementi valutativi in termini di "equità" che nessun sistema di intelligenza artificiale sembra in grado di replicare.

Accolta dunque l'idea che l'intelligenza artificiale possa trovare spazio nelle prognosi cautelari, il passo successivo consiste nel responsabilizzare il giudice sul piano della motivazione. L'uso di algoritmi predittivi dovrebbe quindi accompagnarsi ad una sorta di motivazione *ad hoc* particolarmente rafforzata ai sensi dell'art. 292 c.p.p. Sicché il giudice dovrebbe indicare nel dettaglio gli elementi di fatto a sostegno del quadro indiziario forniti dal sistema d'intelligenza artificiale di volta in volta adottato, ma, soprattutto, ed è que-

stione dirimente, i motivi per cui gli stessi assumono rilevanza in rapporto agli altri dati non digitali. C'è peraltro un altro aspetto da considerare. Sappiamo che il sistema cautelare pretende oggi dai giudici valutazioni autonome (art. 292 c.p.p.), per assicurarne l'effettiva imparzialità. Concepita per evitare appiattimenti dei giudici sui pubblici ministeri, quell'autonomia dovrà essere assicurata anche con riferimento al sistema di *AI* di volta in volta impiegato in sede cautelare.

Particolare attenzione dovrebbe essere infine dedicata alla spiegazione delle ragioni per cui sono stati considerati non rilevanti gli elementi forniti dalla difesa (art. 292 lett. *c bis* c.p.p.), confrontandoli con i dati di segno opposto. Qui viene in gioco, infatti, l'effettiva tutela del *diritto di difesa nei confronti del sistema AI* concretamente fruibile dal pubblico ministero e del giudice, ma anche il principio dell'effettiva *parità delle armi* (art. 111 Cost. e 6 C.E.D.U.). Come s'è detto, la Carta etica prevede che la validità scientifica dell'algoritmo e il peso di volta in volta attribuito allo stesso nella singola vicenda giudiziaria possano sempre essere contestati davanti ad un organo giurisdizionale. Ebbene, posto che il fattore sorpresa che sorregge le misure cautelari impone di posticipare il contraddittorio rispetto al momento della loro applicazione, la previsione di una motivazione rafforzata che illustri puntualmente il sistema di *AI* adottato, la sua affidabilità, i risultati ottenuti, consentirebbe al difensore di acquisire una base cognitiva tale da consentirgli di reagire efficacemente in sede di interrogatorio di garanzia e di riesame (artt. 294 e 309 c.p.p.), per confutare gli elementi a carico digitali.

4. *Prove*. Non meno problematico è l'impatto dell'intelligenza artificiale sulle dinamiche probatorie. Così come accade per ogni sapere scientifico, anche quello fornito dall'intelligenza artificiale andrebbe veicolato nel processo penale attraverso la perizia, nel contraddittorio tra le parti, davanti ad un giudice terzo e imparziale. Anche alla luce dell'alto rischio di manipolazione che riguarda tutti i sistemi digitali<sup>31</sup>.

Naturalmente, l'esigenza di assicurare l'effettiva garanzia del contraddittorio nella formazione della prova «artificiale» presenta tutti i noti problemi legati all'impiego della scienza nel processo. Posto che non è possibile in questa sede neppure accennare ad un tema così ampio e complesso, ci limiteremo solo ad alcuni brevissimi rilievi di carattere generale. Sappiamo che la scienza costituisce un prezioso alleato per il giudice nella ricerca della verità. E in

---

<sup>31</sup> Cfr. sul punto, SIGNORATO, *Le indagini digitali*, cit., 259.

questa prospettiva va da sé che la crescita della cultura scientifica tra gli operatori del diritto andrebbe sensibilmente implementata. Ma sappiamo anche che ogni medaglia ha il suo rovescio. La scienza può infatti costituire per il giudice un falso alleato, incanalando pericolosamente il processo verso l'errore giudiziario. La questione è nota. La giustizia penale attinge ormai ampiamente ai risultati della scienza, ma deve fare i conti con la sua conclamata fallibilità. Spesso il processo diventa il luogo dove i periti e i consulenti tecnici svelano le lacune di quelle stesse discipline che dovrebbero invece orientare in maniera corretta le decisioni degli organi giudicanti. Di qui i noti protocolli: verifica del metodo scientifico impiegato, falsificabilità del metodo, conoscenza della percentuale di errore, accreditamento della teoria nella comunità scientifica.

La complessa questione dell'affidabilità del sapere scientifico si ripropone ovviamente anche con riferimento all'intelligenza artificiale. In questo settore, però, emerge un problema particolare. Come è stato puntualmente osservato<sup>32</sup>, quando si ha a che fare con un *software predittivo*, occorre, per un verso, esaminare una serie di profili differenti: quello dell'algoritmo che trasforma un *input* in un certo *output*; quello della sua natura automatizzata; quello della validità scientifica della teoria; quello della sua tradizione in linguaggio digitale.

Su questa basi, andrebbe salvaguardato il *principio di accesso immediato al giudice* di cui parla la Carta etica, valorizzando l'incidente probatorio. Più precisamente. L'indagine tecnica sull'affidabilità del *software* andrebbe ricondotta all'interno della macro-categoria delle cd. perizie complesse, integrando quindi, per l'appunto, un caso di incidente probatorio ai sensi dell'art. 392, co. 2, c.p.p. Una soluzione di questo tipo potrebbe simultaneamente soddisfare due istanze. Permetterebbe, anzitutto, di anticipare le verifiche tecniche rispetto ad un dibattito spesso lontano nel tempo, nel pieno rispetto del principio della parità di armi e del contraddittorio, davanti ad un giudice terzo e imparziale. Consentirebbe inoltre alla difesa di contestare, subito, davanti al giudice, eventuali difetti del *software*: una mossa strategicamente interessante, perché innescerebbe una tempestiva *disclosure* degli elementi d'accusa che poggiano su dati digitali, facendo così emergere anticipatamente il rischio di una eventuale, futura, pericolosa, *black box decision* dibattimentale<sup>33</sup>.

Sempre con riferimento alla perizia, c'è un altro aspetto da considerare. Sappiamo che una parte della giurisprudenza la considera alla stregua di un

<sup>32</sup> Cfr. QUATTROCOLO, *Questioni nuove e soluzioni antiche?*, cit. 1758.

<sup>33</sup> *Supra*, par. 1.

mezzo di prova “neutro”, cioè né a carico né a discarico. Per di più, l’alto coefficiente di discrezionalità del giudice nella scelta se diporla o no renderebbe la perizia estranea all’area del diritto alla prova delle parti<sup>34</sup>. Un diritto che accusa e difesa potrebbero comunque sempre esercitare attraverso la consulenza tecnica. Come si vede, un ragionamento di questo tipo si regge sul seguente sillogismo: la perizia è una prova lunga e laboriosa; costituisce la prova scientifica *del giudice*; spetta dunque a quest’ultimo decidere discrezionalmente se effettuarla o no.

Ebbene, quali potrebbero essere le ricadute di questo indirizzo giurisprudenziale selettivo laddove emergesse la necessità di una verifica tecnica sull’affidabilità di un sistema intelligenza artificiale? Premesso che il discorso sulla neutralità della perizia non regge, per il semplice motivo che ogni mezzo di prova, prima di conoscerne gli esiti, non è classificabile a carico o a discarico<sup>35</sup>, e posto inoltre che il codice prevede la possibilità di disporre la perizia *anche* d’ufficio (lasciando così intendere che il giudice dovrebbe comunque pronunciarsi, motivando adeguatamente, sull’eventuale diniego della perizia sollecitata dalla parte *ex art.* 190 c.p.p.), preme qui segnalare il pericolo che la risorsa della consulenza tecnica possa risultare del tutto inadeguata ad assicurare l’effettività del diritto di difesa (qui inteso come diritto alla prova tecnica), indebolendo il principio della *par condicio*. Per due motivi, logicamente concatenati. Il primo. Le risorse informatiche a cui può attingere il consulente dell’accusa non sono neppure paragonabili a quelle fruibili dal consulente della difesa. Come s’è detto, alcuni sistemi di *AI* sono protetti dal segreto commerciale e non sono facilmente accessibili ai privati. Il secondo. La persistente e irragionevole diffidenza (frutto di una radicata riserva mentale) verso i contributi tecnici provenienti dal difensore. Lo dimostra una recentissima sentenza, dove compare una singolare, atecnica, pericolosa assimilazione della consulenza tecnica del pubblico ministero alla perizia, e dove si sostiene apertamente la maggior affidabilità del responso del consulente tecnico dell’accusa rispetto a quello fornito dal consulente della difesa<sup>36</sup>. Di qui la necessità d’insistere sulla strada della perizia ogni volta in cui emerga l’esigenza di utilizzare forme di intelligenza artificiale nel processo. Imparzialità del giudice, contraddittorio, *par condicio* e diritto di difesa potrebbero così trovare

<sup>34</sup> Così, ad esempio, Cass. Sez. un. 3 marzo 2017, n. 39746, in *Cass. pen.* 2018, 500.

<sup>35</sup> Per una apprezzabile (ma a nostro avviso ancora troppo timida) critica alla tesi della neutralità della perizia v. Cass. Sez. un., 28 gennaio 2019, Pavan, in *Dir. pen. proc.*, 2019, 822, con nota di BONZANO.

<sup>36</sup> Cass., Sez. I, 29 maggio 2020, n. 16458, Barbone, in [www.sistemapenale.it](http://www.sistemapenale.it), con osservazioni critiche di KOSTORIS, *Una grave mistificazione inquisitoria: la pretesa fede privilegiata del responso del consulente tecnico dell’accusa*.

la loro piena e concreta attuazione.

Proseguendo nell'analisi dell'impatto dell'intelligenza artificiale sulle dinamiche probatorie, ci si può interrogare sulla possibilità di sfruttare gli algoritmi per svolgere indagini psicologiche sui testimoni, e quindi per elaborare statistiche sulla percentuale di credibilità del dichiarante.

Premesso che la giurisprudenza è pacifica nel ritenere che il divieto di perizia psicologica di cui all'art. 220 c.p.p. valga solo per l'imputato, non interessando né il testimone né la vittima<sup>37</sup>, va ricordato che si è soliti parlare di "capacità psicologica a testimoniare" per descrivere l'abilità che si manifesta secondo diversi livelli e che, nella sua espressione ottimale, rende la persona più sicura dei propri ricordi, meno suggestionabile, meno incline a giudizi stereotipati, a proiezioni personali, all'influenza esercitata dai *mass media*<sup>38</sup>. Ebbene, la prospettiva di creare algoritmi predittivi, che, elaborando una molteplicità di dati attinenti al testimone, possano agevolare il giudice e le parti nel valutarne anticipatamente il grado di affidabilità e di attendibilità può rivelarsi estremamente interessante. Così, ad esempio, si potrebbe pensare ad un *software* in grado di gestire un numero elevato di dati (si potrebbe ad esempio, a tal fine, riflettere sulla creazione di una sorta di *banca dati dei testimoni*) per facilitare la scoperta di mitomani, persone inclini a rendersi irreperibili, mendaci e falsi calunniatori seriali. Figure di testimoni inutili e pericolose che, non di rado, emergono solo durante il processo e che potrebbero invece essere individuate anticipatamente proprio grazie all'impegno di algoritmi, assicurando, così, una serie di vantaggi: le velocizzazione dei tempi processuali; l'abbattimento delle spese di giustizia; l'attenuazione del rischio dell'errore giudiziario.

Un discorso di questo tipo dovrebbe valere anche per la testimonianza della vittima del reato: a maggior ragione alla luce dell'indirizzo giurisprudenziale che considera sufficiente ai fini della condanna la sola dichiarazione dell'offeso, senza la necessità di corroborarla con riscontri esterni<sup>39</sup>. Una soluzione da cui traspare un atteggiamento quasi fideistico nei confronti della vittima: un atteggiamento a nostro avviso logicamente inconciliabile con le finalità persecutorie che la vittima indubbiamente persegue attraverso il processo. Se davvero non si vuole e non si può rinunciare alla testimonianza della vittima, che almeno si assicurino controlli rigorosi. A maggior ragione quando la vittima è l'unico testimone disponibile. In questo ordine di idee, ad esempio,

<sup>37</sup> Cfr. ad esempio, Cass., Sez. I, 28 settembre 1995, Russo, in *Cass. pen.*, 1997, 2174.

<sup>38</sup> Cfr. DE CATALDO NEUBURGER, *Psicologia della testimonianza e prova testimoniale*, Milano, 1988, 212.

<sup>39</sup> Cass., Sez. III, 13 febbraio 2020, Ciotti, in *Rv.*, n. 2790070.

si potrebbe sfruttare la risorsa dell'*AI* per gestire un numero significativo di dati relativi al ruolo delle vittime nei processi penali. L'idea di creare un *caselario delle vittime dei reati* che (nella assoluta tutela della *privacy*) registri i casi in cui un soggetto è stato coinvolto come persona offesa (maggiorrenne, minorenni, vittima particolarmente vulnerabile *ex art. 90 quater c.p.p.*) in un procedimento penale (denuncia, querela, opposizione alla richiesta di archiviazione, testimonianza, costituzione di parte civile), e la tipologia di decisione che ne è conseguita (assoluzione/condanna), andrebbe seriamente presa in considerazione. Senza riserve di ordine ideologico. Alla prevedibile critica secondo cui in questo caso si avvierebbe una profilazione foriera di possibili, ulteriori stadi di "vittimizzazione", si potrebbe infatti replicare con l'argomento della ragionevolezza e dell'equilibrio. Si tratterebbe, in fondo, di una pragmatica soluzione di compromesso che consentirebbe ai pubblici ministeri e ai giudici di sfruttare l'*AI* per elaborare dati statisticamente attendibili sulla figura della persona offesa, utilizzabili sia nelle fasi genetiche del procedimento (controllo sulla fondatezza della notizia di reato), sia, in seguito, come riscontro positivo/negativo alla testimonianza della vittima, per assicurare condanne giuste e per evitare condanne ingiuste.

5. *Decisioni.* Come già anticipato, l'impiego di algoritmi predittivi potrebbe influire sugli sviluppi del procedimento, condizionando le scelte dei giudici e delle parti. Viene in rilievo, qui, anzitutto, il profilo della predizione decisoria, che può interessare diversi profili: il passaggio dalla fase investigativa alla fase processuale e il passaggio dall'udienza preliminare al dibattimento, da un lato; il rapporto tra la scelta di riti deflattivi e gli epiloghi dibattimentali, dall'altro lato.

Sotto il primo profilo, l'attenzione si concentra su due rigide alternative: richiesta di archiviazione/ rinvio a giudizio; sentenza di non luogo a procedere/ decreto di citazione a giudizio. Ebbene, per effetto della recente riforma<sup>40</sup>, il pubblico ministero e il giudice dell'udienza preliminare sono ora rispettivamente tenuti a chiedere l'archiviazione e a pronunciare la sentenza di non luogo a procedere quando gli elementi acquisiti nelle indagini preliminari o le prove emerse in udienza preliminare *non consentono di formulare una ragionevole previsione di condanna*. Il precedente criterio predittivo, e cioè l'*insostenibilità dell'accusa in giudizio*<sup>41</sup>, viene così sostituito da un parametro

<sup>40</sup> Legge 27 settembre 2021 n. 134 (*Delega al Governo per l'efficienza del processo penale nonché in materia di giustizia riparativa e disposizioni per la celere definizione dei procedimenti giudiziari*).

<sup>41</sup> Artt. 125 disp. att. e 425 c.p.p. (formulazione originaria).

segno ben diverso: la ragionevole esclusione di un particolare *esito di quel giudizio*, cioè la condanna. La *predizione decisoria* diventa quindi parte integrante del sistema processuale penale. Con l'effetto di innalzare notevolmente l'asticella predittiva: perché escludere ragionevolmente la condanna significa, nei fatti, escludere che si possa raggiungere in dibattimento la prova della colpevolezza oltre ogni ragionevole dubbio. Ecco dunque che la regola di giudizio dibattimentale (la quale, lo sappiamo, è il precipitato tecnico della presunzione di non colpevolezza) si espande, si proietta a ritroso, va a condizionare, nella direzione del *favor rei*, le scelte di pubblici ministeri e giudici per le indagini preliminari (i quali, per decidere tra azione ed inazione, e tra decreto di citazione a giudizio e sentenza di non luogo a procedere, dovranno rispettivamente affidarsi ad indagini complete *ex art 358 c.p.p.*, e a supplementi probatori solidi *ex art. 422 c.p.p.*). In questo complesso scenario, è plausibile ritenere che, soprattutto i pubblici ministeri, dovendo spesso gestire atti investigativi irripetibili (di cui occorrerà pronosticare la tenuta in dibattimento), possano ricorrere spesso a strumenti di predizione decisoria basati sull'*AI*. Con il rischio però, vista l'indubbia difficoltà della prognosi imposta, di appiattirsi troppo sui dati elaborati dalla macchina per sottrarsi (almeno in parte) al peso della responsabilità di una scelta (azione/ inazione) che appare davvero difficile da sostenere (soprattutto nei confronti dell'opinione pubblica).

Come si diceva, gli algoritmi predittivi potrebbero trovare uno spazio significativo anche nell'ambito delle strategie dell'imputato in materia di scelta dei riti speciali deflattivi, in rapporto ai possibili esiti dibattimentali. Si potrebbe forse enunciare la seguente regola di carattere generale: le previsioni più affidabili basate sull'*AI* dovrebbero forse attendersi con riferimento ai riti speciali strutturalmente più semplici, caratterizzati da attività standardizzate, ripetitive. In questa prospettiva, la prognosi più agevole parrebbe forse quella in materia di decreto penale di condanna: qui il ricorso all'algoritmo potrebbe consentire all'imputato di effettuare previsioni abbastanza sicure (in rapporto alla possibile decisione dibattimentale) sull'opportunità di opporsi al decreto o di accettarlo. Strategicamente utile potrebbe risultare, per l'imputato, anche il supporto di algoritmi predittivi in materia di patteggiamento: vuoi per valutare la convenienza del rito in rapporto ai probabili epiloghi dibattimentali (anche se resterebbe l'incognita di un successivo mutamento del quadro probatorio); vuoi per individuare le probabilità di ottenere il consenso del pubblico ministero; vuoi, soprattutto, per verificare la probabilità che il giudice si esprima positivamente sulla sanzione concordata: in questo caso, infatti, basterebbe

esaminare la percentuale delle richieste presentate ed accolte su reati simili dal giudice: quest'ultimo, infatti, non potrebbe modificare l'accordo ma solo decidere sulla congruità della pena individuata dalle parti<sup>42</sup>. Più difficile, invece, immaginare previsioni affidabili basate sull'*AI* in materia di richiesta di sospensione del procedimento con messa alla prova e di giudizio abbreviato. Numerose e articolate le *variabili soggettive* legate all'imputato nel primo rito speciale (la concessione della sospensione è subordinata ad alcune condizioni relative alla vita personale, familiare, sociale ed economica dell'imputato ed alla previsione che il programma di trattamento possa indurre l'imputato ad astenersi dal commettere in futuro altri reati: artt. 464 *bis* e *quater* c.p.p.). Di grande peso le *variabili oggettive* nel giudizio abbreviato: variabili che possono portare ad un significativo cambio di scenario probatorio (prove ammesse d'ufficio parte dal giudice *ex art.* 441 co. 5; contestazione del fatto diverso da parte del pubblico ministero *ex art.* 441 *bis* c.p.p. e conseguente esercizio del diritto alla prova da parte dell'imputato).

Spostando l'attenzione sul piano della diagnosi dibattimentale, ci si può domandare se la risorsa dell'*AI* possa agevolare la decisione sul fatto incerto, riducendo il rischio dell'errore e rafforzando, così, indirettamente, la presunzione d'innocenza (qui intesa come regola di giudizio: *in dubio pro reo*). L'idea di un *software* progettato per elaborare, di volta in volta, sulla base di tutti i dati processuali disponibili, una certa soglia di *ragionevole dubbio* può risultare davvero interessante. Non va infatti dimenticato il problema di fondo sotteso alla decisione sul fatto incerto: il codice individua sì le conseguenze dell'incertezza (ragionevole dubbio→assoluzione; prova della colpevolezza oltre ogni ragionevole dubbio→condanna), ma non chiarisce che cosa s'intende per dubbio ragionevole. Il risultato elaborato dalla macchina su base statistica (ad esempio, incrociando i dati relativi ad una pluralità di sentenze sul medesimo fatto storico e sorrette da mezzi di prova simili) potrebbe fornire una utile base cognitiva, su cui innestare il successivo ragionamento del giudice chiamato a sciogliere l'alternativa assoluzione/condanna (artt. 530 e 533 c.p.p.). Più precisamente. La probabilità statistica elaborata dal sistema di *AI* andrebbe poi integrata dalla probabilità logica sottesa al ragionamento dal giudice<sup>43</sup>. Due, qui, le incognite, legate in particolare alla tutela del principio

<sup>42</sup> Cfr. KOSTORIS, *Predizione decisoria*, cit., 12.

<sup>43</sup> Si tratta, in sostanza, della probabilità logica di matrice baconiana, intesa, ai nostri fini, come grado di conferma che gli elementi di prova sono in grado di assicurare all'ipotesi accusatoria. Per approfondimenti sul punto v. CAPRIOLI, voce *Condanna (dir. proc. pen.)*, in *Enc. dir.* Agg. I, Milano, 2008, 103; FERRUA, *Il giudizio penale: fatto e valore giuridico*, in *La prova nel dibattimento penale*<sup>3</sup>, Torino, 2007 353 (I ed. 1999, 193); ID, *La colpevolezza oltre ogni ragionevole dubbio, Il nuovo regime delle impu-*

dell'imparzialità del giudice, colto da due prospettive differenti: il rischio di un appiattimento interno e quindi la necessità di tracciare un chiaro confine tra il giudizio probabilistico della macchina e il giudizio probabilistico umano; il rischio di un appiattimento «esterno», e cioè la prospettiva di decisioni future troppo condizionate da quelle passate (nei fatti, una sorta di conformismo giudiziario che mal si attaglia alla materia penale e ai valori che vi sono sottesi).

Resta infine da domandarsi quali siano le conseguenze processuali nel caso in cui il giudice disattenda i risultati forniti da un sistema di *AI*. Per quanto riguarda *l'automated evidence* non dovrebbero manifestarsi differenze rispetto a quanto solitamente accade quando il giudice non prende ~~una~~ considerazione un dato scientifico emerso nel processo o lo travisa. Il giudice sarebbe quindi tenuto a motivare analiticamente (art. 192 co. 1 e 546 co. 1 lett. e c.p.p.) sul punto e la sua decisione risulterebbe eventualmente impugnabile sotto il profilo del vizio di motivazione. Più delicato è il problema dei controlli in materia di strumenti di predizione decisoria. Con riferimento a questo profilo, infatti, si ritiene che il comportamento del giudice che ignorasse la predizione decisoria fornita da un sistema di *AI* potrebbe essere assimilato a quello di un giudice che disattende un indirizzo giurisprudenziale consolidato<sup>44</sup>. Ma in un sistema, come il nostro, che (per ora) non contempla la forza vincolante del precedente<sup>45</sup>, le parti non sembrano possedere gli strumenti per ottenere la rimozione della decisione adottata.

---

*gnazioni tra Corte costituzionale e sezioni unite*, a cura di Filippi, Padova, 2007, 150 s; KOSTORIS, *Le impugnazioni penali, travagliato terreno alla ricerca di nuovi equilibri*, in *Riv. dir. proc.*, 2006, 48 s.; e, volendo, PAULESU, *La presunzione di non colpevolezza dell'imputato*, cit., 223 s.

<sup>44</sup> Cfr. KOSTORIS, *Predizione decisoria*, cit., 4.

<sup>45</sup> Sul tema cfr, volendo, PAULESU, *Conflitti interpretativi in Cassazione e valore del precedente nel processo penale*, in *Riv. dir. proc.*, 2019, 1050 s.

