



Università degli Studi di Padova

Dipartimento di Diritto Pubblico, Internazionale e Comunitario

Corso di Laurea in Diritto e Tecnologia

a.a. 2022/2023

**Giustizia predittiva: l'esperienza statunitense e l'eventuale
ammissibilità dei *risk assessment tools* nell'ordinamento italiano**

Relatore: Chiar.mo Prof. Andrea Pin

Studente: Francesco Magagna

Matricola 2010312

INDICE

INTRODUZIONE	1
CAPITOLO I – INTELLIGENZA ARTIFICIALE E GIUSTIZIA PREDITTIVA 3	
1. Intelligenza artificiale: nozione e caratteristiche	3
2. L'intelligenza artificiale nei sistemi giudiziari	4
2.1. Lo stato dell'arte in Italia e nel resto d'Europa	6
3. La giustizia predittiva	7
3.1. Rischi e vantaggi	9
3.2. L'esperienza statunitense	11
3.3. L'approccio europeo	12
CAPITOLO II – IL CASO <i>STATE V. LOOMIS</i>.....	15
1. Il <i>software COMPAS</i>	15
1.1. L'inchiesta di ProPublica	17
2. I fatti.....	18
3. Le questioni giuridiche sollevate	20
4. La decisione della Corte Suprema del Wisconsin.....	21
5. I risvolti.....	23
CAPITOLO III – L'EVENTUALE AMMISSIBILITÀ DEI <i>RISK ASSESSMENT TOOLS</i> NELL'ORDINAMENTO ITALIANO.....	25
1. Il principio di eguaglianza e non discriminazione	25
2. Il principio del giusto processo	26
3. Il principio di personalità della responsabilità penale.....	28
4. I giudizi prognostici in sede di applicazione delle misure cautelari e di commisurazione della pena.....	29

4.1. L'autonoma valutazione delle esigenze cautelari.....	30
5. Il divieto di perizia criminologica.....	30
6. Il decreto legislativo n. 51 del 2018.....	31
7. I principi della Carta etica della CEPEJ.....	32
8. Considerazioni di insieme.....	34
CONCLUSIONI.....	35
BIBLIOGRAFIA.....	37
APPENDICE – INTERVISTA AL DOTT. GIOVANNI CANZIO.....	43

INTRODUZIONE

L'intelligenza artificiale (IA) sta assumendo un ruolo sempre più pervasivo nella realtà odierna. Tale tecnologia si presta a rivoluzionare il paradigma del rapporto uomo-macchina e trova applicazione in un numero assai vasto di ambiti, ivi compreso quello giuridico. Ne sono esempi il procedimento amministrativo informatico e la giustizia predittiva.

La giurisprudenza amministrativa italiana ha già avuto modo di esprimersi in merito all'uso dell'IA nel procedimento amministrativo: un esempio significativo è rappresentato dalla sentenza del Consiglio di Stato n. 881 del 2020. Essa sancisce una declinazione c.d. rafforzata del principio di trasparenza, in base alla quale la «conoscibilità dell'algoritmo deve essere garantita in tutti gli aspetti». Questi ultimi ricomprendono, *inter alia*, il procedimento utilizzato per l'elaborazione del *software* e il meccanismo decisionale da esso adottato.¹ Viene poi menzionato il principio di conoscibilità e comprensibilità, inteso come diritto dell'interessato «a conoscere l'esistenza di processi decisionali automatizzati che lo riguardino ed in questo caso a ricevere informazioni significative sulla logica utilizzata».² La pronuncia riconosce altresì il principio di non esclusività della decisione algoritmica, secondo cui deve sempre essere presente un contributo umano in grado di controllare e, ove necessario, smentire la decisione automatizzata.³ Da ultimo, il Consiglio di Stato fa riferimento al principio di non discriminazione algoritmica, in base al quale l'amministrazione procedente è tenuta ad adottare misure tecniche e organizzative volte a impedire effetti discriminatori nei confronti delle persone fisiche.⁴

Diversamente, in Italia non vi è ancora un quadro di riferimento né normativo né giurisprudenziale in materia di giustizia predittiva. Lo scopo del presente elaborato è dunque quello di valutare le implicazioni di natura giuridica derivanti dall'impiego degli algoritmi predittivi nel procedimento penale e, a seguito di un'analisi dell'approccio nordamericano, stabilirne l'ammissibilità nell'ordinamento italiano. Si è scelto di fare particolare riferimento ai *risk assessment tools* (o strumenti di valutazione del rischio), in quanto – al momento – appaiono i più idonei a trovare ingresso nelle aule di giustizia.

¹ Cons. Stato, sez. VI, 4 febbraio 2020, n. 881, par. 10.

² *Ivi*, par. 11.1.

³ *Ivi*, par. 11.2.

⁴ *Ivi*, par. 11.3.

Nel primo capitolo viene tracciato il perimetro definitorio dell'intelligenza artificiale e vengono esaminate le applicazioni della stessa nei sistemi giudiziari, considerando sia lo stato dell'arte in Italia sia quello nel resto d'Europa. Quanto alla giustizia predittiva, viene operata una distinzione tra predizioni a fini decisori, a cui si riconducono gli strumenti di valutazione del rischio, e predizioni decisorie.

Il secondo capitolo ha ad oggetto l'esperienza statunitense, con specifico riguardo al noto caso *State v. Loomis* dinanzi la Corte Suprema del Wisconsin. In primo luogo, vengono illustrati il funzionamento e le caratteristiche del *software* di valutazione del rischio di recidiva *COMPAS*. In secondo luogo, viene condotta un'analisi della sentenza: se ne espongono i fatti, le questioni giuridiche sollevate dalla difesa, la decisione della Corte e le ripercussioni.

Nel terzo capitolo si indaga sull'eventuale ammissibilità dei *risk assessment tools* nel processo penale italiano, tenendo conto dei rilevanti principi costituzionali e sovranazionali, della compatibilità con il codice di procedura penale, del decreto legislativo n. 51 del 2018 e della Carta etica europea della CEPEJ.

Infine, in appendice è riportata la trascrizione dell'intervista che Giovanni Canzio, Primo Presidente emerito della Corte di cassazione, ha gentilmente concesso allo scrivente: i temi affrontati concernono la giustizia predittiva, i relativi risvolti e le prospettive future.

CAPITOLO I

INTELLIGENZA ARTIFICIALE E GIUSTIZIA PREDITTIVA

1. Intelligenza artificiale: nozione e caratteristiche

Non vi è, al momento, una definizione univoca e generalmente accolta di intelligenza artificiale.¹ Ai fini della presente trattazione, si è scelto di far riferimento alla nozione contenuta nella *Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, adottata nel 2018 dalla Commissione europea per l'efficienza della giustizia (CEPEJ). Tale atto di *soft law*² definisce l'IA come «insieme di metodi scientifici, teorie e tecniche finalizzate a riprodurre mediante le macchine le capacità cognitive degli esseri umani».³

L'IA si caratterizza innanzitutto per l'utilizzo di notevoli quantità di dati e informazioni, nonché per l'elevata capacità logico-computazionale. Essa implementa sofisticati algoritmi, come quelli di *deep learning* e *machine learning*, che sono in grado di estrarre conoscenza dai dati al fine di prendere decisioni. Inoltre, tali sistemi di autoapprendimento possono modificare ed eventualmente correggere gli algoritmi di origine stessi man mano che ricevono nuove informazioni su ciò che stanno elaborando.⁴

Come diversi autori hanno evidenziato, l'intelligenza artificiale è dotata di una neutralità solo apparente. Infatti, l'algoritmo «è ontologicamente condizionato dal sistema di valori e dalle intenzioni di chi ne commissiona la creazione e/o di chi lo crea».⁵ La discrezionalità umana è rinvenibile dal momento della raccolta dei dati sino all'interpretazione del risultato finale prodotto dall'elaboratore.⁶

¹ A. Santosuosso, *Intelligenza artificiale e diritto. Perché le tecnologie di IA sono una grande opportunità per il diritto*, Mondadori Università, Milano, 2020, 6 ss.; G. Ubertis, *Intelligenza artificiale, giustizia penale, controllo umano significativo*, in *Giurisdizione penale: intelligenza artificiale ed etica del giudizio*, Ed. Giuffrè, Milano, 2021, 10.

² Per atti di *soft law* si intendono gli atti privi di efficacia vincolante.

³ CEPEJ, *Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, 2018, appendice III, 47. Il documento è consultabile al sito <https://rm.coe.int/carta-etica-europea-sull-utilizzo-dell-intelligenza-artificiale-nei-si/1680993348>.

⁴ G. Ubertis, *op. cit.*, 10.

⁵ S. Signorato, *Giustizia penale e intelligenza artificiale. Considerazioni in tema di algoritmo predittivo*, in *Riv. dir. proc.*, 2020, 614.

⁶ G. Ubertis, *op. cit.*, 13.

La complessità dei modelli di apprendimento automatico deriva anche dalla loro tendenziale opacità, la quale rende poco agevole – se non addirittura impossibile – ricostruire il ragionamento seguito dall’algoritmo per giungere a una data previsione.⁷ Tale opacità si deve anzitutto alla sopraccitata capacità dei sistemi di IA di «riprogrammarsi da soli durante l’elaborazione di una enorme massa di dati».⁸ In aggiunta, il codice sorgente dei *software* utilizzati è spesso protetto dal segreto industriale, il quale impedisce ai terzi di visionare l’algoritmo.

2. L’intelligenza artificiale nei sistemi giudiziari

L’IA «applicata al settore della giurisdizione (*legal AI*) può definirsi un sistema digitale che, con specifico riguardo all’ambiente della giustizia, acquisisce, ordina e rielabora una enorme quantità di informazioni (*big data*), di tipo giudiziario o giurisprudenziale, al fine di identificare la soluzione ottimale della questione posta».⁹

È possibile individuare due tipologie di IA in ambito giuridico: un modello c.d. forte e un modello c.d. debole. Il primo emula in modo sconfinato le capacità cognitive dell’uomo e «postula l’automazione del processo decisionale in luogo degli attori tradizionali della giurisdizione»; il secondo è un modello collaborativo che si basa sulla complementarità tra uomo e macchina e consente all’umano di mantenere il controllo sull’algoritmo. Pertanto, nell’IA debole il responso dell’algoritmo costituisce un mero supporto alla decisione del giudice.¹⁰ Secondo la citata Carta etica europea, è necessario verificare la coerenza logica dell’esito dell’algoritmo «in un processo d’integrazione fra le misurazioni quantitative [...] da esso offerte con il percorso cognitivo e decisorio del giudice, nel rispetto dei metavalori dell’ordinamento». Tale imprescindibile integrazione rende evidente come in Europa risulti sinora ammissibile soltanto il modello debole di IA.¹¹

L’intelligenza artificiale si presta ad avere molteplici applicazioni nel campo della giustizia. Innanzitutto, si stanno diffondendo sempre più i sistemi di polizia predittiva, i

⁷ A. Facchini – A. Termine, *Explainable AI: come andare oltre la black box degli algoritmi*, Agenda Digitale, 2022, <https://www.agendadigitale.eu/cultura-digitale/explainable-ai-come-andare-oltre-la-black-box-degli-algoritmi/>.

⁸ G. Ubertis, *op. cit.*, 14.

⁹ G. Canzio, *Intelligenza artificiale e processo penale*, in G. Canzio – L. Lupária (a cura di), *Prova scientifica e processo penale*, CEDAM, Padova, 2022, 903.

¹⁰ *Ivi*, 903 ss.; Appendice, I s.

¹¹ G. Canzio, *Intelligenza artificiale e processo penale*, cit., 907 s.

quali sono in grado di prevedere il momento e il luogo in cui potrebbero essere commessi certi reati oppure chi potrebbe commetterli. Le predizioni sono originate dall'elaborazione di grandi moli di dati riguardanti, in particolare, i reati già commessi, le attività delle persone sospettate, l'origine etnica e il livello di scolarizzazione dei soggetti, i luoghi a maggior tasso di criminalità e persino le condizioni atmosferiche maggiormente correlate alla commissione di reati.¹²

L'IA viene inoltre utilizzata nell'ambito degli *automated decision systems*: si tratta di algoritmi diretti a prevenire o comporre liti e risolvere controversie. Questi nuovi metodi alternativi di risoluzione delle controversie garantiscono un significativo risparmio di tempi e costi rispetto ai sistemi tradizionali e spesso vengono offerti solamente *online*.¹³

L'IA può essere impiegata anche come mezzo di ricerca della prova nella fase delle indagini preliminari, al fine di acquisire «tracce, cose [e] notizie idonee ad assumere rilevanza probatoria». Ne sono esempi i sistemi di riconoscimento facciale automatico, i quali possono essere utilizzati per identificare, a partire da immagini e video, i soggetti coinvolti in un determinato reato.¹⁴

È poi stato prospettato l'uso dell'intelligenza artificiale come mezzo di valutazione dell'affidabilità degli elementi probatori nel processo. Ad esempio, il riconoscimento facciale automatico consentirebbe di analizzare la comunicazione non verbale nel corso delle dichiarazioni dell'imputato e dei testimoni al fine di verificarne l'attendibilità.¹⁵

Infine, l'IA trova applicazione nel campo della giustizia predittiva. Tale espressione fa riferimento a «qualsiasi decisione giudiziale, parziale o totale, presa da un computer che apprende dall'esperienza».¹⁶

¹² F. Basile, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Diritto Penale e Uomo*, 10/2019, 1, 10.

¹³ *Ivi*, 14.

¹⁴ L. Romanò, *Intelligenza artificiale come prova scientifica nel processo penale*, in G. Canzio – L. Lupária (a cura di), *Prova scientifica e processo penale*, CEDAM, Padova, 2022, 925 ss.

¹⁵ *Ivi*, 929 s.

¹⁶ M. Barberis, *Giustizia predittiva: ausiliare e sostitutiva. Un approccio evolutivo*, in *Milan Law Review*, 2/2022, 1, 7.

2.1. Lo stato dell'arte in Italia e nel resto d'Europa

Vi sono attualmente diversi progetti volti a introdurre l'IA nella giustizia italiana. È il caso della Scuola Superiore Sant'Anna di Pisa, che ha ideato una piattaforma di giustizia predittiva in grado di valutare i possibili esiti di un procedimento penale. Ciò è possibile grazie all'analisi, compiuta da algoritmi di apprendimento automatico, di *big data* concernenti pronunce giurisprudenziali.¹⁷ Parimenti si è mossa la Corte d'appello di Brescia progettando uno strumento di giustizia predittiva che si prefigge di comunicare agli operatori del diritto «la durata prevedibile di un procedimento su di una data materia e gli orientamenti esistenti» nei palazzi di giustizia della città.¹⁸ Infine, in materia tributaria rileva *Prodigit*, un progetto sperimentale avviato dal Consiglio di presidenza della giustizia tributaria allo scopo di creare un sistema di IA in grado di predire, a partire dall'analisi di atti normativi, sentenze e contributi dottrinali, la possibile decisione dei giudici in relazione a una specifica questione. Ciò permetterà dunque al contribuente di valutare l'opportunità di presentare ricorso o meno.¹⁹

Quanto alle attività di polizia predittiva, alcune questure italiane stanno utilizzando *software* quali *Key Crime* e *XLaw* per prevenire la consumazione di reati tramite azioni di pattugliamento mirate.²⁰ Nell'ambito della pubblica sicurezza, la Polizia di Stato fa uso, già da alcuni anni, del *SARI* (acronimo di «Sistema Automatico di Riconoscimento Immagini»). Tale sistema, mediante un sofisticato algoritmo, «associa all'immagine di un individuo sconosciuto le generalità di chi sia già presente nelle banche dati della polizia, con una percentuale di affidabilità estremamente alta».²¹

Anche nel resto d'Europa non manca l'implementazione dell'IA nel campo della giustizia. In Francia, ad esempio, sono divenuti popolari i sistemi *Predictice* e *Case Law Analytics*. Il primo è volto a predire i potenziali esiti delle controversie, mentre il secondo permette di stimare l'ammontare dei danni a cui si potrebbe essere condannati.²²

¹⁷ M. Martorana, *Giustizia predittiva: rischi, vantaggi e un modello italiano*, CFnews.it, 2021, <https://www.cfnews.it/avvocatura/giustizia-predittiva-rischi-vantaggi-e-un-modello-italiano/>.

¹⁸ C. Castelli, *Giustizia predittiva: così a Brescia l'intelligenza artificiale aiuta a prevedere tempi e esiti delle cause*, Agenda Digitale, 2019, <https://www.agendadigitale.eu/cultura-digitale/giustizia-predittiva-cosi-a-brescia-lintelligenza-artificiale-aiuta-a-prevedere-tempi-e-esiti-delle-cause/>.

¹⁹ A. Traversi, *Giustizia predittiva: quale futuro?*, Altalex, 2023, <https://www.altalex.com/documents/news/2023/03/15/giustizia-predittiva-qual-e-futuro>.

²⁰ C. Cavaceppi, *L'intelligenza artificiale applicata al diritto penale*, in G. Taddei Elmi – A. Contaldo (a cura di), *Intelligenza artificiale. Algoritmi giuridici. Ius condendum o "fantadiritto"?*, Pacini Giuridica, Pisa, 2020, 111.

²¹ S.n., *SARI = Sistema Riconoscimento Automatico Immagini. La Polizia di Stato ad Arezzo individua e denuncia i primi 7 ladri*, Questura di Arezzo, 2018, <https://questure.poliziadistato.it/it/Arezzo/articolo/10945bb6148a60216596783584>.

²² G. Meliotta, *Intelligenza Artificiale e giustizia predittiva*, in *Diritto del risparmio*, 2/2022, 205, 217 s.

Vi sono poi i Paesi Bassi, in cui l'applicativo *Rechtwijzer* viene impiegato come metodo alternativo di risoluzione delle controversie in maniera interamente telematica. Esso ha un ampio raggio d'intervento e copre le controversie in materia di diritti di proprietà, lavoro, divorzio e servizi alla persona, nonché le controversie condominiali.²³

Considerevole è anche l'esperienza della polizia britannica, la quale si avvale della tecnologia *Hart* per determinare il rischio di recidiva del reo nei due anni successivi alla commissione dell'illecito. Tale sistema combina tra loro diverse informazioni riguardanti gli imputati, come genere, età, domicilio e storia criminale. Le previsioni si sono rivelate corrette in una percentuale di casi compresa tra l'88% e il 98%.²⁴

3. La giustizia predittiva

Il versante della giustizia predittiva comprende sia le predizioni decisorie sia le predizioni a fini decisorii.

Le predizioni decisorie consistono nella previsione dell'esito di una controversia mediante sistemi di IA in grado di autoapprendere dalla casistica giurisprudenziale.²⁵ Sarebbe «proprio questa prevedibilità a costituire un'aspirazione di fondo della giustizia penale». Oggi, però, il valore della c.d. certezza del diritto è affievolito da un diritto notevolmente elastico – in quanto basato sempre più su principi – oltre che a larga matrice giurisprudenziale. L'intelligenza artificiale parrebbe quindi in grado di riaffermare la concezione spiccatamente illuminista di un diritto calcolabile.²⁶

Tuttavia, gli algoritmi predittivi sinora sviluppati presentano un'importante limitazione poiché non hanno la capacità tecnica di emulare il complesso ragionamento giuridico che è alla base dell'operato del magistrato. Essi sostituiscono la comprensione dei nessi causali tipica delle decisioni giudiziarie con le correlazioni che identificano su ingenti volumi di dati. A ciò si aggiunge l'impossibilità per un modello computazionale di compiere valutazioni qualitative, le quali comprendono anche le emozioni e le intuizioni proprie dell'essere umano.²⁷

²³ E. Nagni, *Artificial Intelligence, l'innovativo rapporto di (in)compatibilità fra machina sapiens e processo penale*, in *Sistema penale*, 7/2021, 5, 12.

²⁴ C. Cavaceppi, *op. cit.*, 107.

²⁵ M. Gialuz, *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa*, in *Diritto penale contemporaneo*, 2019, 1, 3.

²⁶ R.E. Kostoris, *Predizione decisoria e diversion processuale*, in *Giurisdizione penale: intelligenza artificiale ed etica del giudizio*, Ed. Giuffrè, Milano, 2021, 97.

²⁷ *Ivi*, 99-101.

Per quanto concerne gli strumenti predittivi a fini decisori, essi hanno lo scopo di «fornire elementi su cui basare una decisione, sia di tipo parziale, come avviene in materia cautelare, sia tale da involgere l'intera regiudicanda».²⁸ Appartengono a questa categoria i *risk assessment tools*: si tratta di programmi per elaboratore che fanno uso dell'IA per calcolare il rischio che l'accusato compia nuovi atti criminosi oppure si sottragga al processo. Tali *software* utilizzano una serie di fattori (ad es. lo *status* socioeconomico, il contesto familiare, l'origine etnica, lo *status* occupazionale, l'indice di criminalità dell'area di residenza e la presenza di eventuali disturbi psichici) al fine di ottenere una previsione relativa al rischio di delinquenza della persona coinvolta espresso sotto forma di punteggio.²⁹

Grazie all'analisi di grandi quantità di informazioni, gli strumenti di valutazione del rischio sono in grado di individuare *patterns* che trovano fondamento su una solida base statistica, prendendo in considerazione molte più variabili rispetto a quanto potrebbe fare un umano.³⁰ Di conseguenza, è possibile affermare che i *risk assessment tools* abbiano dato origine a una «concezione *evidence-based* di valutazione del rischio individuale di commissione di un (nuovo) reato: una concezione, quindi, basata su riscontri oggettivi, destinata a soppiantare, o quanto meno integrare, le valutazioni intuitive dei giudici, tuttora ampiamente diffuse».³¹

La dottrina statunitense ritiene che l'attendibilità della teoria psico-criminologica adottata dai singoli strumenti di *risk assessment* sia sindacabile, a monte, mediante il c.d. *Daubert test*. Quest'ultimo consiste nell'applicazione dei criteri enunciati dalla Corte Suprema degli Stati Uniti nella celebre pronuncia *Daubert v. Merrell Dow Pharmaceuticals* (1993), la quale concerne l'ammissibilità delle perizie tecnico-scientifiche nel processo penale.³² Il test in questione è volto alla valutazione dell'affidabilità di una teoria o un metodo scientifici sulla base di cinque parametri: la verificabilità attraverso esperimenti; l'esito negativo dei tentativi di smentita; la sottoposizione al controllo della comunità scientifica mediante la c.d. *peer review*; il tasso di errore, accertato o potenziale, dei risultati; la generale accettazione da parte della

²⁸ *Ivi*, 95.

²⁹ M. Gialuz, *op. cit.*, 3.

³⁰ *Ibidem*.

³¹ F. Basile, *op. cit.*, 17.

³² S. Quattrocchio, *Risk assessment: sentencing o non sentencing?*, in *Giurisdizione penale: intelligenza artificiale ed etica del giudizio*, Ed. Giuffrè, Milano, 2021, 79.

comunità scientifica di riferimento.³³ Tali criteri hanno trovato riconoscimento nell'ordinamento italiano con la sentenza della Corte di cassazione n. 43786 del 2010.³⁴

Per mezzo del *Daubert test* andrà valutata anche la validità dell'algoritmo stesso quale trasposizione della teoria psico-criminologica in termini computazionali.³⁵ Tuttavia, come si vedrà nel successivo paragrafo, l'opacità dei sistemi di IA e la loro conseguente natura di *black boxes* rendono piuttosto arduo ricostruire a posteriori i passaggi seguiti dall'algoritmo nell'applicare la teoria.³⁶

3.1. Rischi e vantaggi

I rischi connessi all'impiego degli strumenti di giustizia predittiva derivano dalle caratteristiche intrinseche dell'intelligenza artificiale, alle quali si è già fatto riferimento (*supra*, par. 1).

In primo luogo, alla luce dei recenti utilizzi dei *risk assessment tools* negli USA, sono stati sollevati dubbi circa l'effettiva validità predittiva, nota come *accuracy*, degli algoritmi in questione.³⁷ A ben vedere, «il risultato fornito dagli algoritmi predittivi è necessariamente influenzato dalla qualità dei dati che vengono posti come *input*»: sarebbe doveroso implementare meccanismi in grado di garantire l'accuratezza dei dati, l'indipendenza della fonte da cui essi provengono e la pubblicità degli stessi.³⁸

In secondo luogo, l'IA tende a replicare i pregiudizi umani. Se, infatti, tali sistemi vengono addestrati con dati parziali, incompleti o che riflettono i *biases* da cui sono state influenzate le precedenti decisioni, vi è il rischio che gli stessi pregiudizi siano incorporati nelle decisioni future. Andrebbe così a svanire l'imparzialità che si suole associare alle macchine, le quali non appaiono ancora in grado di esimersi dai comportamenti discriminatori tipici degli esseri umani. Per esempio, a causa dei pregiudizi etnici radicati nella storia nordamericana, gli strumenti di valutazione del rischio in uso alle corti statunitensi sono soliti sovrastimare l'indice di pericolosità sociale dei soggetti afroamericani.³⁹

³³ G. Canzio, *La motivazione della sentenza e la prova scientifica*, in G. Canzio – L. Lupária (a cura di), *Prova scientifica e processo penale*, CEDAM, Padova, 2022, 11.

³⁴ G. Canzio, *La motivazione della sentenza e la prova scientifica*, cit., 11.

³⁵ A.M. Maugeri, *L'uso di algoritmi predittivi per accertare la pericolosità sociale: una sfida tra evidence based practices e tutela dei diritti fondamentali*, in *Archivio Penale*, 1/2021, 1, 24.

³⁶ S. Quattrocchio, *op. cit.*, 79 s.; Appendice, I.

³⁷ F. Basile, *op. cit.*, 22.

³⁸ *Ivi*, 23.

³⁹ L. Romanò, *op. cit.*, 921 s.

Si pone poi la discussa questione della (non) trasparenza: i *tools* di giustizia predittiva si caratterizzano invero per la loro imperscrutabilità. Quest'ultima è dovuta alla natura commerciale della gran parte degli algoritmi predittivi utilizzati negli Stati Uniti, per cui il codice sorgente è coperto dal segreto industriale e non è visionabile né dal giudice né dalle parti. Una possibile soluzione potrebbe essere quella di introdurre l'obbligo di «pubblicazione dei codici sorgente e delle altre informazioni rilevanti ai fini dell'accessibilità dell'algoritmo, attraverso procedure competitive di selezione del fornitore del servizio predittivo».⁴⁰

Cionondimeno, vi è chi ritiene che la conoscibilità del codice sorgente non sia sufficiente a superare l'opacità algoritmica.⁴¹ Infatti, come è stato opportunamente evidenziato nel corso della trattazione, i sistemi di IA tendono a essere delle vere e proprie scatole nere (*black boxes*), in quanto non sono basati su una logica comprensibile agli umani e possiedono, in genere, forme di autoapprendimento (*machine learning*). Sovente anche gli stessi programmatori non sono in grado di spiegare la logica alla base della decisione algoritmica.⁴²

È stato altresì individuato il rischio di «de-individualizzazione delle decisioni», derivante dal fatto che «l'*output* “predittivo” costituisce il prodotto di un'estensione al caso concreto dei risultati di un'elaborazione di carattere statistico condotta su un campione di casi diversi da quello oggetto di giudizio». Qualora il giudice assumesse la propria decisione sulla sola base del risultato dello strumento predittivo, si assisterebbe a una generalizzazione del tutto incompatibile con la necessaria «valutazione individualizzata del caso in questione».⁴³

Infine, vi è il timore che la decisione si appiattisca sull'esito dell'algoritmo, mettendo a rischio i principi di imparzialità e libero convincimento del giudice. Egli potrebbe infatti attribuire un peso eccessivo e ingiustificato all'*output* dell'algoritmo, incorrendo nella c.d. fallacia dell'automazione. Tale fallacia, nota anche come *automation bias*, si verifica quando gli esseri umani tendono a prediligere le soluzioni provenienti da sistemi computazionali, ignorando le informazioni contraddittorie ottenute senza l'ausilio di essi.⁴⁴ È importante, come ha affermato Giovanni Canzio, che «il

⁴⁰ L. Notaro, *Intelligenza artificiale e giustizia penale*, in C. Chinnici (a cura di), *Intelligenza artificiale e giustizia penale*, Paruzzo, Caltanissetta, 2020, 123 ss.

⁴¹ *Ibidem*.

⁴² V. Cavosi, *Governare l'intelligenza artificiale. Spunti per la progettazione di sistemi di IA legali, etici e robusti*, Ledizioni, Milano, 2022, 27.

⁴³ *Ivi*, 120.

⁴⁴ *Ivi*, 126.

giudice sia consapevole che non può mai diventare un mero consumatore di un algoritmo».⁴⁵

Nonostante la giustizia predittiva sia caratterizzata da non poche insidie, essa presenta diversi vantaggi. Anzitutto, grazie all'automazione di alcune operazioni, si assisterebbe a un risparmio in termini di tempi e costi, nonché a una riduzione dell'incidenza dell'errore umano.⁴⁶ In aggiunta, l'IA – se adeguatamente sviluppata – potrebbe garantire una maggiore oggettività delle decisioni, le quali sono spesso influenzate dai pregiudizi e dall'arbitrio umani. Da ultimo, gli strumenti di predizione decisoria assicurerebbero una tendenziale calcolabilità e uniformità delle decisioni, contribuendo «alla stabilizzazione della giurisprudenza».⁴⁷ Pertanto, purché implementati in modo corretto ed etico, gli algoritmi di giustizia predittiva risponderebbero al principio di efficacia ed efficienza dell'attività amministrativa.

3.2. L'esperienza statunitense

Pionieri nel campo della giustizia predittiva sono gli Stati Uniti, i quali fanno uso di *risk assessment tools* già da una quindicina d'anni. Tali strumenti vengono impiegati nelle giurisdizioni statunitensi quando è necessario compiere un giudizio prognostico: ciò riguarda sia le decisioni in materia di *pretrial release* (concernenti cioè il rilascio dell'indagato o dell'imputato sottoposto a custodia cautelare) sia le decisioni in materia di *sentencing* (ossia riguardanti la commisurazione della pena da parte del giudice).⁴⁸

Con riferimento al *pretrial release*, sono oltre venti i *software* di valutazione del rischio oggi in uso nelle corti statunitensi. Sette Stati⁴⁹ richiedono per legge – almeno in determinati casi – l'uso di tali algoritmi, mentre altri otto⁵⁰ autorizzano e tendono a incoraggiare la loro adozione. Inoltre, alcuni Stati⁵¹ hanno adottato una specifica legislazione in materia, stabilendo «requisiti di imparzialità e non discriminazione, garanzie di trasparenza e accessibilità del sistema e dei dati alla base dell'algoritmo, nonché la previsione di una validazione periodica dello strumento predittivo».⁵²

⁴⁵ Appendice, VI.

⁴⁶ G. Canzio, *Intelligenza artificiale e processo penale*, cit., 906.

⁴⁷ C. Cavaceppi, *op. cit.*, 101.

⁴⁸ L. Notaro, *op. cit.*, 95.

⁴⁹ Alaska, Delaware, Hawaii, Indiana, Kentucky, New Jersey e Vermont.

⁵⁰ Colorado, Illinois, Montana, New York, Pennsylvania, Rhode Island, Virginia e West Virginia.

⁵¹ Idaho, New York e California.

⁵² L. Notaro, *op. cit.*, 97 s.

Tra gli algoritmi maggiormente utilizzati nel campo del *pretrial release* rileva *PSA (Public Safety Assessment)*, il quale è adottato dagli Stati dell'Arizona, del Kentucky e del New Jersey, oltre che da numerose altre giurisdizioni. Si tratta di un *tool* creato da un ente *non-profit* e che si prefigge di escludere discriminazioni derivanti dall'etnia, dal genere e dalle condizioni economiche. A tal proposito, è stato deciso di renderne pubblici i criteri di attribuzione dei punteggi di pericolosità.⁵³

PSA analizza nove informazioni relative all'età del soggetto, all'imputazione e ai suoi precedenti penali, al fine di attribuire un punteggio per ciascuno dei due fattori di rischio da valutare: «da un lato, il pericolo che il prevenuto non si presenti in udienza» e, dall'altro, «la probabilità che questi commetta un reato se rilasciato prima del dibattimento». È interessante notare come, nella contea di Lucas (Ohio), l'impiego del *software* in questione abbia portato alla riduzione dei reati commessi in attesa di giudizio e, al contempo, all'aumento delle persone messe in libertà.⁵⁴

I *risk assessment tools* utilizzati in fase di *sentencing* sono invece più di sessanta. Tra di essi, oltre all'assai noto *COMPAS* che verrà debitamente esaminato nel successivo capitolo, figura *LSI-R (Level of Service Inventory - Revised)*. Quest'ultimo utilizza fattori come i precedenti penali del reo, i caratteri della sua personalità e la sua situazione socioeconomica al fine di determinarne il rischio di recidiva.⁵⁵

3.3. L'approccio europeo

Al momento, gli algoritmi di giustizia predittiva non hanno ancora trovato applicazione nelle giurisdizioni dei Paesi europei. Come ben rileva Giovanni Canzio, l'Europa si sta muovendo in «un terreno complesso e delicato, provando a creare un ordinamento che regoli l'ingresso di queste forme di IA nel processo».⁵⁶ Infatti, il dibattito che è stato «opportunamente avviato a livello europeo ha a oggetto il come i sistemi giudiziari saranno in grado, nel prossimo futuro, di far fronte a tali sviluppi tecnologici, senza divenirne vittime».⁵⁷

Per quanto riguarda la c.d. grande Europa, un considerevole strumento è rappresentato dalla soprammenzionata Carta etica della CEPEJ, la quale sancisce cinque

⁵³ M. Gialuz, *op. cit.*, 7.

⁵⁴ *Ivi*, 7 s.

⁵⁵ *Ivi*, 5.

⁵⁶ Appendice, II.

⁵⁷ M. Gialuz, *op. cit.*, 12.

principi generali che verranno nel prosieguo analizzati (*infra*, cap. III, par. 7). Nell'appendice II del documento si fa specifico riferimento alle esperienze statunitense e britannica in relazione ai *risk assessment tools*: tali strumenti vengono inquadrati tra gli utilizzi dell'IA «da esaminare con le più estreme riserve». In particolare, viene evidenziato come essi abbiano «condotto a risultati errati» e discriminatori nei confronti della popolazione nera.⁵⁸

Nell'ambito della c.d. piccola Europa, rilevano le disposizioni della direttiva (UE) 680/2016, recepita nell'ordinamento italiano con il decreto legislativo 51/2018. Tale direttiva si applica al «trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali».

Inoltre, la Commissione europea ha presentato nel 2021 una proposta di regolamento che mira a stabilire regole armonizzate sull'intelligenza artificiale e che potrebbe essere approvata in via definitiva a breve. L'allegato III della proposta include tra i «sistemi di IA ad alto rischio» quelli «destinati a essere utilizzati dalle autorità di contrasto per effettuare valutazioni individuali dei rischi delle persone fisiche al fine di determinare il rischio di reato o recidiva».⁵⁹ Le attività considerate ad alto rischio sarebbero consentite soltanto nel rispetto di stringenti requisiti: l'implementazione di un «sistema di gestione dei rischi» (art. 9); l'adozione di specifici accorgimenti in materia di «governance dei dati» (art. 10); la redazione di una particolare «documentazione tecnica» (art. 11); la «registrazione automatica degli eventi (“log”) durante il loro funzionamento» (art. 12); la «trasparenza e fornitura di informazioni agli utenti» (art. 13); la presenza di una «sorveglianza umana» (art. 14); un adeguato livello di «accuratezza, robustezza e ciphersicurezza» (art. 15).⁶⁰

⁵⁸ CEPEJ, Carta etica cit., appendice II, 44.

⁵⁹ Commissione europea, *Proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale* (doc. 52021PC0206), 2021, allegato III, par. 6, lett. a).

⁶⁰ *Ivi*, artt. 9-15.

CAPITOLO II

IL CASO *STATE V. LOOMIS*

1. Il software *COMPAS*

Il più celebre tra i diversi strumenti di *risk assessment* in uso alle corti statunitensi è, con ogni probabilità, *COMPAS* (acronimo di «*Correctional Offender Management Profiling for Alternative Sanctions*»). Il sistema si fonda su un *database* contenente più di trentamila valutazioni compiute dall'algoritmo tra gennaio 2004 e novembre 2005 presso strutture quali carceri e centri per la libertà vigilata (*probation sites*).¹ Si stima che l'applicativo, sviluppato nel 1998, sia stato utilizzato in oltre un milione di casi.² Esso è commercializzato dall'azienda privata Northpointe (in seguito rinominata Equivant) ed è entrato al centro del dibattito pubblico con il caso *State v. Loomis*³ celebrato dinanzi la Corte Suprema del Wisconsin.⁴ Sono molteplici le giurisdizioni che si avvalgono di detto *software*, tra cui la contea di Broward (Florida), lo Stato della California, lo Stato del Wisconsin e lo Stato di New York.⁵

I dati che il programma riceve in ingresso comprendono le informazioni presenti nel fascicolo dell'imputato e le risposte al questionario da questi compilato.⁶ Al soggetto vengono poste 137 domande concernenti le seguenti tematiche:

- precedenti penali;
- passato violento;
- violenza attuale;
- frequentazione di pregiudicati;
- abuso di sostanze stupefacenti;
- difficoltà economica o povertà;

¹ *S.n.*, *Practitioner's Guide to COMPAS Core*, Equivant, 2019, 11 (disponibile al sito <https://www.equivant.com/wp-content/uploads/Practitioners-Guide-to-COMPAS-Core-040419.pdf>).

² J. Dressel – H. Farid, *The accuracy, fairness, and limits of predicting recidivism*, in *Science Advances*, 4(1)/2018, 1.

³ *State of Wisconsin v. Eric L. Loomis*, 881 N.W.2d 749 (Wis. 2016).

⁴ L. Notaro, *Intelligenza artificiale e giustizia penale*, in C. Chinnici (a cura di), *Intelligenza artificiale e giustizia penale*, Paruzzo, Caltanissetta, 2020, 100.

⁵ K. Kirkpatrick, *It's not the algorithm, it's the data*, in *Commun. ACM*, vol. 60(2), 2017, 21, 21.

⁶ *State of Wisconsin v. Eric L. Loomis*, cit., par. 13. Un esempio del questionario è consultabile al sito <https://www.documentcloud.org/documents/2702103-Sample-Risk-Assessment-COMPAS-CORE.html>.

- difficoltà nell'istruzione e formazione;
- pensiero criminoso e personalità criminale;
- criminalità in famiglia;
- problemi di natura sociale;
- modalità di impiego del tempo libero;
- instabilità abitativa;
- difficoltà di socializzazione;
- isolamento sociale;
- opportunità criminali.⁷

L'algoritmo adotta un approccio empirico-statistico per valutare il rischio di recidiva del reo e fornisce in *output* un punteggio in decimi relativo a ciascuno dei tre seguenti fattori: il *pretrial risk*, cioè il rischio che, in pendenza del procedimento, l'accusato non si presenti in udienza o torni a delinquere⁸; il rischio di recidiva generico (*general recidivism*); il rischio di recidiva c.d. violenta (*violent recidivism*).⁹ In aggiunta, «accanto alla funzione strettamente “predittiva”, il sistema elabora [...] una *need scale*, idonea a definire il profilo dell'autore e le esigenze di riabilitazione del medesimo».¹⁰

Come è emerso dalla citata sentenza della Corte Suprema del Wisconsin, il *software* non è in grado di fornire la specifica probabilità di recidiva del singolo autore di reato: esso, invece, elabora una previsione basata sul confronto tra le informazioni concernenti l'individuo in questione e quelle afferenti a un gruppo di soggetti con caratteristiche analoghe.¹¹ Inoltre, è stato notato che, sebbene *COMPAS* sia utilizzato nella fase di *sentencing*, tale applicativo è stato originariamente sviluppato per fornire supporto ai giudici nelle decisioni in merito al *pretrial release* e all'esecuzione della pena.¹² Infine, nonostante l'algoritmo non consideri esplicitamente l'etnia come fattore, questa può essere verosimilmente inferita da aspetti quali il livello d'istruzione, l'area di residenza, lo *status* socioeconomico e l'instabilità familiare. Ciò ha sollevato una serie di

⁷ *Practitioner's Guide to COMPAS Core*, cit., 9; F. Basile, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Diritto Penale e Uomo*, 10/2019, 22.

⁸ Cfr. <https://www.lawinsider.com/dictionary/pretrial-risk-assessment>.

⁹ *State of Wisconsin v. Eric L. Loomis*, cit., par. 14.

¹⁰ L. Grossi, *Software predittivi e diritto penale*, in C. Chinnici (a cura di), *Intelligenza artificiale e giustizia penale*, Paruzzo, Caltanissetta, 2020, 172.

¹¹ *State of Wisconsin v. Eric L. Loomis*, cit., par. 15.

¹² L. Notaro, *op. cit.*, 101.

dubbi sul rispetto dei principi costituzionali statunitensi del giusto processo (*due process*) e della c.d. equal protezione da parte della legge (*equal protection of the law*).¹³

1.1. L'inchiesta di ProPublica

Nel 2016 *COMPAS* è stato oggetto di un'approfondita inchiesta di ProPublica, un'organizzazione non governativa con sede a Manhattan che, mediante il giornalismo investigativo, mira a denunciare abusi di potere da parte di istituzioni pubbliche e private.¹⁴ La ricerca ha interessato un campione di oltre settemila individui arrestati tra il 2013 e il 2014 nella contea di Broward e i relativi punteggi di rischio calcolati dal suddetto *risk assessment tool*.¹⁵

Secondo lo studio, il sistema sarebbe notevolmente inaffidabile nel prevedere la commissione di reati violenti nei due anni successivi alla condanna. Infatti, soltanto il 20% degli individui a cui era stato attribuito un alto tasso di recidiva violenta avrebbe effettivamente compiuto crimini violenti nell'arco temporale in esame. Le previsioni si sarebbero rivelate inaccurate anche in riferimento al tasso di recidiva generale, con un livello di precisione pari al 61%. Dalla ricerca sarebbe altresì emerso che, a causa della formula utilizzata, l'algoritmo ha alte probabilità di etichettare erroneamente come futuri criminali gli imputati afroamericani, con una percentuale quasi doppia rispetto agli imputati bianchi.¹⁶ Questi ultimi, al contrario, tenderebbero ad essere considerati meno rischiosi di quanto realmente siano. Invero, il 48% dei bianchi etichettati come soggetti a basso rischio avrebbe bensì recidivato nei due anni successivi. Per quanto riguarda gli afroamericani, invece, tale cifra ammonterebbe al 28%.¹⁷

Infine, ProPublica evidenzia come – sul piano teorico – i giudici non dovrebbero infliggere una pena più lunga sulla mera base del tasso di recidiva elaborato dall'algoritmo. Piuttosto, esso dovrebbe essere utilizzato per valutare l'idoneità del reo alla libertà vigilata o a specifici programmi di trattamento. Ciò, tuttavia, non sempre accade: nella sentenza di primo grado del caso *Loomis*, ad esempio, il giudice Scott Horne

¹³ H. Liu *et al.*, *Beyond State v. Loomis: Artificial Intelligence, Government Algorithmization, and Accountability*, in *International Journal of Law and Information Technology*, vol. 27(2), 13.

¹⁴ Cfr. <https://www.propublica.org/about>.

¹⁵ J. Larson *et al.*, *How We Analyzed the COMPAS Recidivism Algorithm*, ProPublica, 2016, <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>.

¹⁶ J. Angwin *et al.*, *Machine bias. There's software used across the country to predict future criminals. And it's biased against blacks*, ProPublica, 2016, <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

¹⁷ *Ibidem*.

ha condannato l'imputato a una lunga pena detentiva, facendo esplicito riferimento all'alto tasso di rischio calcolato dal *software COMPAS*.¹⁸

Non è mancata la replica di Northpointe all'indagine di ProPublica. L'azienda produttrice respinge con forza la conclusione che gli indici di rischio di *COMPAS* siano discriminatori nei confronti della popolazione nera e ritiene che ProPublica abbia commesso diversi errori tecnici e statistici nella sua ricerca.¹⁹ Secondo Northpointe, l'applicativo ha previsto correttamente i casi di recidiva sia dei bianchi sia dei neri e, per qualsiasi punteggio in decimi attribuito da *COMPAS*, un bianco e un nero hanno la stessa probabilità di commettere nuovamente un reato.²⁰

2. I fatti

Prima di ripercorrere i principali snodi argomentativi della sentenza *State v. Loomis*, sembra opportuno riassumere brevemente i fatti. Nel 2013 il cittadino statunitense Eric Loomis venne arrestato mentre era alla guida di un'automobile utilizzata in una sparatoria. Gli vennero contestati cinque capi d'imputazione: messa in pericolo dell'incolumità pubblica; tentata fuga dalle forze di polizia; guida di un veicolo a motore senza il consenso del proprietario; possesso di arma da fuoco da parte di un pregiudicato; possesso di fucile o carabina.²¹ L'imputato sostenne di essere estraneo alla sparatoria e si dichiarò colpevole soltanto per i due reati meno gravi, ossia la tentata fuga e la guida di un veicolo senza l'autorizzazione del proprietario.²²

Ai fine dell'irrogazione della pena, la *circuit court* di La Crosse (Wisconsin) – ovvero il tribunale di originaria giurisdizione – richiese il *Presentence Investigation Report* (PSI).²³ Si tratta di un rapporto, preparato dai funzionari del dipartimento per la libertà vigilata, nel quale figurano informazioni sull'imputato e sul suo passato, nonché

¹⁸ *Ibidem*.

¹⁹ W. Dieterich *et al.*, *COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity*, Northpointe, 2016, 1 (disponibile al sito http://go.volarisgroup.com/rs/430-MBX-989/images/ProPublica_Commentary_Final_070616.pdf).

²⁰ E. Yong, *A Popular Algorithm Is No Better at Predicting Crimes Than Random People*, The Atlantic, 2018, <https://www.theatlantic.com/technology/archive/2018/01/equivant-compas-algorithm/550646/>.

²¹ *State of Wisconsin v. Eric L. Loomis*, cit., par. 11. La denominazione originaria dei capi d'accusa è la seguente: *recklessly endangering safety; attempting to flee or elude a traffic officer; operating a motor vehicle without the owner's consent; possession of a firearm by a felon; possession of a short-barreled shotgun or rifle*.

²² *Ivi*, par. 12.

²³ *Ibidem*.

qualsiasi altro elemento idoneo a incidere sulla gravità della pena.²⁴ Nel caso in questione, il PSI conteneva, tra le altre cose, la valutazione di rischio compiuta dall'applicativo *COMPAS*²⁵: essa presentava un elevato tasso di rischio in riferimento a ciascuno dei tre parametri considerati dall'algoritmo (*pretrial risk*, *general recidivism* e *violent recidivism*). Il *report* includeva anche una descrizione di come il responso del *software* dovesse essere interpretato e un monito volto a prevenirne un uso improprio. In particolare, il documento sottolineava il fatto che il *risk assessment* non dovesse essere utilizzato in sede di commisurazione della pena.²⁶

Il tribunale circondariale condannò Loomis a un totale di sei anni di reclusione e cinque anni di *extended supervision* (istituto simile alla liberazione condizionale), rifiutando di concedere la libertà vigilata.²⁷ Nella propria sentenza, il giudice giustificò l'entità della pena inflitta in ragione di diversi fattori, tra cui la valutazione di *COMPAS*, la gravità dei reati commessi dall'imputato, la sua storia, il suo profilo criminoso e anche i capi imputazione per cui egli non si era dichiarato colpevole.²⁸

In seguito, la difesa presentò un'istanza di revisione della pena in cui sostenne che, prendendo in considerazione il *risk assessment* dell'algoritmo in fase di *sentencing*, la corte avesse violato il diritto dell'imputato a un giusto processo.²⁹ In relazione a tale istanza, il tribunale circondariale celebrò due udienze, nella seconda delle quali Loomis chiamò a testimoniare, in qualità di esperto, il dott. David Thompson. Quest'ultimo ribadì che le valutazioni di *COMPAS* non dovessero essere utilizzate al fine di determinare la pena, giacché il *software* non era destinato a quell'uso.³⁰ Secondo il parere del consulente, «la corte che prende in considerazione i risultati forniti dall'algoritmo incorre fortemente nella probabilità di sovrastimare il rischio di recidiva individuale e di determinare la pena dell'imputato sulla scorta di fattori ininfluenti».³¹ Inoltre, il dott. Thompson mise in risalto il fatto che i tribunali non fossero in possesso di sufficienti informazioni riguardo a come l'algoritmo analizzasse il rischio di recidiva: a titolo di esempio, non era chiaro il modo in cui la storia individuale dell'imputato venisse comparata con quella della fascia

²⁴ Cfr. <https://www.mnp.uscourts.gov/presentence-investigation>.

²⁵ Come riporta ProPublica nella soprammenzionata inchiesta, il Dipartimento di correzione del Wisconsin ha adottato, a partire dal 2012, l'uso dell'applicativo *COMPAS* in tutto lo Stato.

²⁶ *State of Wisconsin v. Eric L. Loomis*, cit., par. 16 ss.

²⁷ *Ivi*, par. 22, nota 18.

²⁸ *Ivi*, par. 19 s.

²⁹ *Ivi*, par. 23; S. Carrer, *Se l'amicus curiae è un algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin*, in *Giurisprudenza Penale Web*, 4/2019.

³⁰ *State of Wisconsin v. Eric L. Loomis*, cit., par. 26.

³¹ S. Carrer, *op. cit.*

di popolazione di riferimento, né se detta popolazione appartenesse allo stesso Stato statunitense.³²

La *circuit court* negò l'istanza di revisione, ritenendo che la pena sarebbe stata la medesima a prescindere dalla considerazione dei punteggi di rischio elaborati dallo strumento predittivo. Conseguentemente, Loomis impugnò la decisione dinanzi la Corte d'appello, la quale rimise la questione alla Corte Suprema del Wisconsin.³³

3. Le questioni giuridiche sollevate

Nel giudizio avanti la Corte Suprema del Wisconsin, Loomis sostenne che, avvalendosi della valutazione di rischio di *COMPAS* in fase di *sentencing*, il giudice di primo grado avesse violato il diritto dell'imputato a un giusto processo. Nel dettaglio, il ricorrente menzionò i seguenti profili di illegittimità: (i) la violazione del diritto dell'imputato a essere giudicato sulla base di informazioni accurate, in quanto la natura proprietaria dell'algorithm non consentiva di verificarne l'accuratezza; (ii) la violazione del diritto a una sentenza individualizzata, derivante dalle generalizzazioni statistiche compiute dall'algorithm; (iii) l'uso improprio del parametro del sesso da parte del *software*.³⁴ Come notò la Corte, l'appellante non stava contestando né l'utilizzo della valutazione di rischio di *COMPAS* per le decisioni diverse dalla commisurazione della pena, né l'impiego della citata funzionalità di *needs assessment* in fase di *sentencing*. Diversamente, Loomis contestava soltanto l'uso della funzionalità di *risk assessment* dello strumento predittivo nell'attività di *sentencing*.³⁵

Con particolare riferimento all'opacità dell'algorithm, il ricorrente evidenziò come lo sviluppatore – facendo leva sul segreto industriale – non avesse reso note le modalità di determinazione dei punteggi di rischio, tra cui i pesi attribuiti ai diversi fattori presi in considerazione. L'assenza di tali informazioni non avrebbe quindi permesso di sindacare l'accuratezza del *risk assessment*.³⁶ Inoltre, Loomis citò alcuni studi a supporto della scarsa affidabilità di *COMPAS*.³⁷

³² *State of Wisconsin v. Eric L. Loomis*, cit., par. 27.

³³ *Ivi*, par. 28.

³⁴ *Ivi*, par. 34.

³⁵ *Ivi*, par. 33.

³⁶ *Ivi*, par. 51.

³⁷ *Ivi*, parr. 59 ss.

Relativamente al secondo profilo di violazione, la difesa enfatizzò il fatto che, poiché gli indici di rischio elaborati dall'applicativo si basavano su dati statistici di gruppo, il sistema non fosse in grado di determinare la specifica pericolosità del singolo individuo, bensì quella del gruppo di appartenenza.³⁸

Infine, per quanto concerne la terza ipotesi di violazione, Loomis asserì che, siccome *COMPAS* considerava il dato del genere nell'attribuire i punteggi di rischio, la *circuit court* avesse violato il diritto dell'imputato a non essere giudicato sulla base del sesso. Secondo l'appellante, infatti, il *software* tendeva a stimare rischi maggiori per gli individui di sesso maschile rispetto a quelli di sesso femminile: questo perché, in media, gli uomini avevano un tasso di recidiva più elevato rispetto alle donne.³⁹

4. La decisione della Corte Suprema del Wisconsin

Nella sua sentenza, la Corte Suprema del Wisconsin esaminò le questioni di diritto poste da Loomis, respingendo ciascuna delle tre ipotesi di violazione delle norme processuali.⁴⁰

Quanto alla prima eccezione sollevata, i giudici citarono una precedente pronuncia della Corte d'appello del Wisconsin in cui era stato riconosciuto che il diritto a essere giudicati sulla base di informazioni accurate comprendeva anche il diritto a esaminare e verificare le informazioni contenute nel PSI.⁴¹ Tuttavia, sebbene non fosse noto come *COMPAS* utilizzasse i dati ricevuti in *input* per calcolare i punteggi di rischio, il manuale operativo dell'algoritmo chiariva come tali punteggi si basassero prevalentemente su dati statici (come i precedenti penali dell'imputato), con un uso ridotto di fattori dinamici (come l'abuso di sostanze stupefacenti e le frequentazioni di soggetti pregiudicati). Inoltre, dal momento che la valutazione di rischio si basava sulle risposte al questionario fornite da Loomis, oltre che sui dati – pubblicamente disponibili – del suo casellario giudiziario, la Corte sostenne che il ricorrente avesse avuto la possibilità di verificare l'esattezza delle domande e delle risposte presenti sul *report* del *software*.⁴² Conseguentemente, i giudici ritennero che non vi fu, da parte del tribunale circondariale, alcuna violazione del diritto dell'imputato a essere giudicato sulla base di informazioni

³⁸ *Ivi*, parr. 67 s.

³⁹ *Ivi*, parr. 75-79.

⁴⁰ *Ivi*, par. 29.

⁴¹ *Ivi*, par. 47.

⁴² *Ivi*, parr. 54 s.

accurate. Ciononostante, essi sottolinearono l'esigenza di ulteriori cautele in relazione all'uso delle valutazioni di *COMPAS*. In particolare, i PSI contenenti i *risk assessments* dell'applicativo in questione dovevano informare l'organo giudicante riguardo le seguenti limitazioni: l'assenza di precisi dettagli circa il funzionamento dell'algoritmo; il fatto che le valutazioni venissero effettuate confrontando gli imputati con un campione nazionale anziché locale; il fatto che alcuni studi avessero identificato la tendenza del *software* a sovrastimare il rischio di recidiva di alcune minoranze.⁴³

Per quanto concerne l'ipotesi di violazione del diritto a una sentenza individualizzata, la Corte affermò che la considerazione dei risultati di *COMPAS* in fase di determinazione della pena, assieme a tutti gli altri fattori a disposizione, mettesse in realtà i giudici in possesso di un maggior numero di informazioni al fine di giungere a una sentenza individualizzata.⁴⁴ Inoltre, la Corte esortò i tribunali di primo grado a «esercitare la propria discrezionalità nel tenere conto dei risultati *COMPAS* con riguardo peculiare ad ogni specifico individuo».⁴⁵

In riferimento all'ultima ipotesi di violazione contestata da *Loomis*, il collegio giudicante rilevò come il ricorrente non fosse stato in grado di dimostrare se la *circuit court* avesse effettivamente dato rilievo al genere dell'imputato in sede di condanna. In aggiunta, ad avviso della Corte, la considerazione del fattore-sesso nella determinazione del tasso di recidiva trovava fondamento su una base fattuale e, per tale motivo, garantiva l'accuratezza dei risultati dell'algoritmo.⁴⁶

In conclusione, la Corte Suprema ritenne che l'uso di *COMPAS* non avesse violato il diritto dell'imputato a un giusto processo, stabilendone così la legittimità anche in fase di *sentencing*. I giudici, tuttavia, non rinunciarono a definire alcuni vincoli rivolti ai tribunali che ricorrevano all'uso di tale strumento predittivo.⁴⁷ In primo luogo, i punteggi di rischio non potevano essere utilizzati autonomamente rispetto ad altri elementi per stabilire se il reo dovesse scontare la propria pena in carcere o meno, né per determinare l'entità della pena stessa. In secondo luogo, era necessario che i tribunali circondariali spiegassero quali fattori, oltre alla valutazione di rischio di *COMPAS*, avessero rilevato – in modo indipendente – nella commisurazione della pena.⁴⁸ Si trattò di un approccio non

⁴³ *Ivi*, parr. 65 s.

⁴⁴ *Ivi*, parr. 72 s.

⁴⁵ S. Carrer, *op. cit.*

⁴⁶ *State of Wisconsin v. Eric L. Loomis*, cit., parr. 83 e 86.

⁴⁷ *Ivi*, parr. 86 s.

⁴⁸ *Ivi*, parr. 98 s.

dissimile dal principio europeo di non esclusività della decisione algoritmica, il quale verrà analizzato nel prosieguo (*infra*, cap. III, par. 6).

La Corte elencò inoltre alcuni ambiti in cui i tribunali potevano utilizzare, insieme ad altri fattori, i *risk assessment tools* a fondamento delle proprie decisioni: l'applicazione di misure alternative alla detenzione per gli individui con un basso rischio di recidiva; la concessione della messa alla prova; l'imposizione di termini e presupposti per la liberazione condizionale.⁴⁹

5. I risvolti

A seguito dell'esito negativo della sentenza della Corte Suprema del Wisconsin, il difensore di Loomis presentò un *writ of certiorari* (domanda di revisione giudiziaria) alla Corte Suprema degli Stati Uniti. Quest'ultima respinse l'istanza, rifiutando così di esaminare il caso.⁵⁰ Si trattò dell'ulteriore conferma della legittimità dell'uso non esclusivo dei *risk assessment tools* da parte delle corti statunitensi in sede di commisurazione della pena.

La sentenza *Loomis* non fu esente da critiche: nella dottrina internazionale vi è chi sostenne che, invece di risolvere le sfide imposte dalla tecnologia in relazione al giusto processo, alla egual protezione da parte della legge e alla trasparenza, la Corte Suprema del Wisconsin avesse preferito limitarsi a fornire un avvertimento circa i rischi legati all'uso degli strumenti di giustizia predittiva. Ciò avrebbe causato l'esternalizzazione – almeno parziale – del processo decisionale del giudice ad algoritmi connotati da un'affidabilità discutibile, oltre che da una noncuranza delle norme fondamentali dell'ordinamento statunitense.⁵¹

In riferimento al contesto europeo, Leonardo Romanò si esprime come segue.

Quanto emerso nel caso *Loomis* del 2016 rappresenta l'emblema del timore che la dottrina maggioritaria nel nostro Paese (e non solo) mostra nei confronti dello stato dell'arte degli algoritmi predittivi: il trattamento automatizzato offerto da *COMPAS* è indecifrabile, generalizzato, discriminatorio, invasivo della sfera personale dell'imputato e pericolosamente prescrittivo per il giudice. Ad ognuna di queste caratteristiche corrisponde la violazione di un diritto o di una garanzia espressa, non solo al livello costituzionale interno, ma anche sovranazionale europeo.⁵²

⁴⁹ *Ivi*, par. 88

⁵⁰ *Certiorari denied, Loomis v. Wisconsin*, 137 S. Ct. 2290 (2017).

⁵¹ H. Liu *et al.*, *op. cit.*, 14 s.

⁵² L. Romanò, *Intelligenza artificiale come prova scientifica nel processo penale*, in G. Canzio – L. Lupària (a cura di), *Prova scientifica e processo penale*, CEDAM, Padova, 2022, 933.

Un ulteriore ammonimento è mosso da Jordi Nieva-Fenoll, il quale ritiene che il rischio più grande sia «quello di utilizzare i risultati dell'applicazione per accertare direttamente, benché solamente in via cautelare o [...] incidentale, la colpevolezza dell'accusato». Infatti, «per emanare misure tanto gravi come la carcerazione preventiva, nel processo penale si richiede sostanzialmente la prova certa della colpevolezza»: non sarebbe affatto sufficiente basarsi su una semplice percentuale di rischio.⁵³

⁵³ J. Nieva-Fenoll, *Intelligenza artificiale e processo*, Giappichelli, Torino, 2019, 62.

CAPITOLO III

L'EVENTUALE AMMISSIBILITÀ DEI *RISK ASSESSMENT TOOLS* NELL'ORDINAMENTO ITALIANO

1. Il principio di eguaglianza e non discriminazione

A seguito dell'analisi del *leading case* statunitense, vi è da interrogarsi sull'eventuale ammissibilità dei *risk assessment tools* nel processo penale italiano.

Tra i principi potenzialmente coinvolti nell'uso di tali algoritmi, assume fondamentale rilievo quello di eguaglianza. Ai sensi dell'art. 3 della Costituzione, «tutti i cittadini hanno pari dignità sociale e sono eguali davanti alla legge, senza distinzione di sesso, di razza, di lingua, di religione, di opinioni politiche, di condizioni personali e sociali». Ne consegue che, sul piano formale, a ciascun consociato è garantita in egual maniera l'astratta titolarità dei diritti fondamentali. A ciò si accosta la proiezione sostanziale del principio di eguaglianza, la quale è sancita nel secondo comma del menzionato articolo. Tale norma impone che i soggetti più deboli (ad es. coloro che si trovano in situazioni di difficoltà economica o di emarginazione sociale) siano messi nella condizione di poter effettivamente godere dei diritti che astrattamente detengono.¹ Si tratta di un'eguaglianza c.d. di risultato: al fine di giungere al medesimo esito, è necessario che la legge tratti in modo uguale situazioni giuridiche uguali e in modo diverso situazioni diverse. Questo enunciato sembra essere espressione della logica matematica secondo cui, effettuando le medesime operazioni sui medesimi dati, il risultato sarà sempre lo stesso.² Per tale motivo è possibile ritenere che, se fossero esenti da *bias* e programmati in maniera etica, gli algoritmi assolverebbero appieno detta funzione.

Rileva inoltre il principio di non discriminazione *ex art.* 14 della Convenzione europea dei diritti dell'uomo (CEDU), il quale vieta qualunque forma di discriminazione nel godimento dei diritti e delle libertà riconosciuti dalla Convenzione (nel caso in esame interessa soprattutto il diritto a un equo processo). Un analogo divieto è rinvenibile all'art.

¹ Q. Camerlengo – F. Furlan, *Lezioni di diritto costituzionale vivente*, CEDAM, Padova, 2021, 301 s.

² L. Viola, *Giustizia predittiva*, in *Diritto on line*, Treccani, https://www.treccani.it/enciclopedia/giustizia-predittiva_%28Diritto-on-line%29/.

21 della Carta dei diritti fondamentali dell'Unione europea (CDFUE), che prevede ulteriori motivi di discriminazione quali la disabilità e l'orientamento sessuale.

Pertanto, alla luce dei riferimenti normativi considerati, non è ammissibile l'uso di algoritmi di valutazione del rischio che operino discriminazioni fondate su elementi come l'etnia, il sesso e lo *status* socioeconomico dell'indagato o imputato. Ciò assume particolare rilevanza se si considerano le violazioni lamentate da ProPublica e Loomis in riferimento all'applicativo *COMPAS*.³

2. Il principio del giusto processo

Come si è visto, la Corte Suprema del Wisconsin ritenne che l'utilizzo di *COMPAS* non pregiudicasse il diritto dell'imputato a un equo processo. Ai fini di stabilire l'idoneità dei *risk assessment tools* ad essere ammessi nell'ordinamento italiano, è necessario considerare il principio del giusto processo di cui all'art. 111 Cost., così come modificato dalla legge costituzionale n. 2 del 1999. Ivi viene sancito che «ogni processo si svolge nel contraddittorio tra le parti, in condizioni di parità, davanti a giudice terzo e imparziale» (secondo comma). Dal dettato normativo è individuabile il principio del contraddittorio, che impone al giudice di non pronunciarsi su alcuna domanda se la parte contro cui essa è proposta non viene regolarmente citata oppure non compare in udienza.⁴ Vi è poi uno specifico riferimento al processo penale, il quale è regolato dal «contraddittorio nella formazione della prova» (quarto comma). Ciò significa che, nella fase delle indagini preliminari, il difensore dell'indagato ha diritto ad essere presente agli atti di indagine o a conoscerne il verbale e che, nella fase dibattimentale, la prova è ritenuta valida soltanto se è stata formata oralmente all'interno dell'esame incrociato.⁵ Ai fini della presente indagine rileva altresì il terzo comma, il quale, nel riconoscere il diritto di difesa, riproduce sostanzialmente le garanzie espresse nell'art. 6 CEDU: il diritto dell'accusato di essere informato «riservatamente della natura e dei motivi dell'accusa elevata a suo carico», di disporre «del tempo e delle condizioni necessari per preparare la sua difesa», di poter «interrogare o di far interrogare le persone che rendono dichiarazioni a suo carico», di ottenere «l'acquisizione di ogni mezzo di prova a suo favore» e di essere

³ Vedasi *supra*, cap. II, parr. 1.1 e 3.

⁴ *S.n.*, *Giusto processo*, in *Enciclopedia on line*, Treccani, <https://www.treccani.it/enciclopedia/giusto-processo/>.

⁵ *S.n.*, *Principio del contraddittorio. Diritto processuale penale*, in *Enciclopedia on line*, Treccani, <https://www.treccani.it/enciclopedia/principio-del-contraddittorio-diritto-processuale-penale>.

assistito da un interprete qualora non conosca la lingua impiegata nel processo.⁶ Infine, dal sesto comma dell'art. 111 cit. si ricava l'obbligo di motivazione di tutti i provvedimenti giurisdizionali.⁷

Un importante corollario del giusto processo è il principio di parità delle armi che la Corte europea dei diritti dell'uomo ha da tempo inferito dalla nozione di «processo equo» ex art. 6 della Convenzione. Detto principio può essere ricavato anche dal secondo comma dell'art. 111 Cost., il quale lo riconduce espressamente al principio del contraddittorio.⁸ La c.d. parità delle armi impone che debba essere garantito un giusto equilibrio tra accusa e difesa. Nello specifico, è richiesto che ciascuna parte abbia la possibilità di conoscere le allegazioni e le argomentazioni della controparte, al fine eventuale di contestarle e falsificarle.⁹

Le garanzie del giusto processo sono riconosciute anche dal diritto dell'Unione europea, in particolare all'art. 47 CDFUE.

Per quanto concerne i *risk assessment tools*, i diritti maggiormente in pericolo sembrano essere quello di difesa e quello di parità delle armi. Infatti, qualora il codice sorgente dell'algoritmo e il suo funzionamento non fossero resi noti (vedasi il *software COMPAS*), l'indagato o imputato non avrebbe la possibilità di seguirne il ragionamento ed eventualmente confutare la validità del risultato.¹⁰ Alcune problematiche residuerebbero anche nel caso in cui lo sviluppatore dell'applicativo decidesse di pubblicarne il codice. In primo luogo, la difesa necessiterebbe di un'elevata specializzazione tecnica ai fini dello studio e dell'analisi dell'algoritmo: tale attività sarebbe accessibile solo ai soggetti in possesso di considerevoli mezzi economici.¹¹ In secondo luogo, in relazione agli algoritmi di *machine learning*, «non pare del tutto chiaro se l'accesso al codice sorgente sia sufficiente a consentire una completa conoscibilità dei meccanismi di funzionamento dello strumento».¹²

⁶ P. Ferrua, *Il processo penale dopo la riforma dell'articolo 111 della Costituzione*, in *Questione giustizia*, 1/2000, 1.

⁷ La motivazione è stata definita come «l'espressione dell'iter logico-giuridico attraverso il quale il giudice è pervenuto alla decisione» (V. Denti, *sub Art. 111 Cost.*, in *La magistratura*, vol. IV, in G. Branca (a cura di), *Commentario alla Costituzione*, Zanichelli, Bologna, 1987, 12).

⁸ E. Lupo, *Le garanzie di contesto: la parità tra le parti*, in *La legislazione penale*, 10/2020, 20, 20.

⁹ S. Quattrocchio, *Processo penale e rivoluzione digitale: da ossimoro a endiadi?*, in *MediaLaws*, 3/2020, 121, 128 s.

¹⁰ L. Grossi, *Software predittivi e diritto penale*, in C. Chinnici (a cura di), *Intelligenza artificiale e giustizia penale*, Paruzzo, Caltanissetta, 2020, 179 s.

¹¹ Appendice, V.

¹² L. Notaro, *Intelligenza artificiale e giustizia penale*, in C. Chinnici (a cura di), *Intelligenza artificiale e giustizia penale*, Paruzzo, Caltanissetta, 2020, 125.

Inoltre, vi è il rischio – come si è visto nel caso *Loomis* – che il risultato dell’algoritmo condizioni eccessivamente il giudice, pregiudicandone così l’imparzialità.¹³ Ciò sarebbe dovuto all’eccessiva fiducia che l’uomo tende a riporre nelle macchine, ossia alla citata fallacia dell’automazione.¹⁴

Da ultimo, l’impiego dei *risk assessment tools* da parte dei tribunali è idoneo a incidere anche sull’obbligo di motivazione delle pronunce giurisdizionali, in virtù del quale i provvedimenti debbono essere sorretti da motivazioni analitiche, chiare ed esaustive. A tal proposito, si ritiene che l’organo giudicante sarà tenuto a illustrare perché intende conformarsi al risultato dello strumento di giustizia predittiva.¹⁵

3. Il principio di personalità della responsabilità penale

Un ulteriore profilo di violazione che potrebbe derivare dall’impiego degli algoritmi di valutazione del rischio riguarda il principio di personalità della responsabilità penale. Tale principio trova fondamento nell’art. 27, commi 1 e 3, Cost. e stabilisce che soltanto l’autore del reato possa essere chiamato a rispondere dello stesso.¹⁶ La Corte costituzionale, nella sentenza n. 50 del 1980, si è espressa sull’importanza del principio in questione, affermando che l’individualizzazione della pena è volta a «tenere conto dell’effettiva entità e delle specifiche esigenze dei singoli casi» e «si pone come naturale attuazione e sviluppo di principi costituzionali».¹⁷ Parallelamente, per quanto concerne le fasi precedenti all’emanazione della sentenza, è stato individuato – a partire dagli artt. 13 e 27, comma 2, Cost. – il principio di individualizzazione delle misure cautelari.¹⁸

Poiché, come si è già visto¹⁹, i *risk assessment tools* determinano il tasso di recidiva del singolo individuo sulla base di generalizzazioni statistiche, vi è la possibilità che si assista alla violazione della garanzia di individualizzazione del trattamento sanzionatorio. Si aprirebbe così la strada a un inammissibile diritto penale del profilo

¹³ L. Grossi, *op. cit.*, 179.

¹⁴ Vedasi *supra*, cap. I, par. 3.1.

¹⁵ P.P. Paulesu, *Intelligenza artificiale e giustizia penale. Una lettura attraverso i principi*, in *Arch. pen. web*, 1/2022, 8 s.

¹⁶ M. Gialuz, *Quando la giustizia penale incontra l’intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa*, in *Diritto penale contemporaneo*, 2019, 1, 21.

¹⁷ Corte cost., 2 aprile 1980, n. 50, par. 3 (*Considerando in diritto*).

¹⁸ *Ibidem*.

¹⁹ Vedasi *supra*, cap. I, par. 3.1.

d'autore, in cui la pericolosità di una persona viene calcolata prendendo in considerazione le condotte commesse in passato da altri individui.²⁰

D'altro canto, è opportuno «chiedersi se da tale vincolo di individualizzazione si possa desumere una vera e propria regola di esclusione della valutazione di pericolosità» compiuta dagli algoritmi predittivi. A ben vedere, già ora i giudici sono chiamati ad operare valutazioni prognostiche avvalendosi di massime di esperienza che sono frutto della generalizzazione delle vicende di altri soggetti: la differenza è che gli algoritmi hanno la capacità di compiere inferenze statistiche prendendo in considerazione un numero assai più ampio di esperienze passate.²¹

4. I giudizi prognostici in sede di applicazione delle misure cautelari e di commisurazione della pena

Nello stabilire le condizioni secondo cui le misure cautelari possono essere disposte dal giudice, l'art. 274 del codice di rito richiede a questi di svolgere una serie di valutazioni di carattere predittivo circa la pericolosità del prevenuto. In particolare, la disposizione fa riferimento al pericolo d'inquinamento probatorio, al pericolo di fuga e al pericolo di reiterazione del reato.²² Quest'ultimo va valutato – oltre che sulla base delle «specifiche modalità e circostanze del fatto» – in relazione alla «personalità della persona sottoposta alle indagini o dell'imputato, desunta da comportamenti o atti concreti o dai suoi precedenti penali».²³

In materia di esigenze cautelari, rilevano altresì i criteri di scelta delle misure *ex* art. 275 c.p.p. Anche detta norma impone un giudizio prognostico, il quale riguarda sia la sanzione che «si ritiene possa essere irrogata» sia la possibilità che «con la sentenza possa essere concessa la sospensione condizionale della pena».²⁴

Per quel che concerne la fase di quantificazione della pena, il codice penale, all'art. 133, sancisce che il giudice debba tenere conto della gravità del reato e della capacità a delinquere del soggetto. Quest'ultima deve essere desunta «dai motivi a delinquere e dal carattere del reo», «dalla condotta e dalla vita del reo, antecedenti al

²⁰ M. Gialuz, *op. cit.*, 21.

²¹ *Ibidem*.

²² M. Gialuz, *op. cit.*, 19; P.P. Paulesu, *op. cit.*, 11.

²³ Art. 274, lett. c), c.p.p.

²⁴ Art. 275, commi 2 e 2-bis, c.p.p.

reato», «dalla condotta contemporanea o susseguente al reato» e, infine, «dalle condizioni di vita individuale, familiare e sociale del reo».²⁵

A prima vista, gli elementi che l'organo giudicante è chiamato a valutare nelle suddette prognosi non sembrano differire molto da quelli considerati dal *software COMPAS* nell'elaborazione dei punteggi di rischio.²⁶ Non sarebbe dunque ragionevole impedire al giudice di avvalersi dell'IA al fine di compiere le valutazioni predittive richieste dalla legge in sede di applicazione delle misure cautelari e di commisurazione della pena. L'unico profilo di complessità potrebbe porsi in riferimento alla determinazione della proporzionalità della misura cautelare *ex art. 275, comma 2, c.p.p.* Tale attività, infatti, richiede «supplementi valutativi in termini di “equità”» che nessun algoritmo è – al momento – in grado di imitare: si renderebbe ancor più necessaria, in questo senso, la supervisione dell'essere umano.²⁷

4.1. L'autonoma valutazione delle esigenze cautelari

Ai sensi dell'art. 292 c.p.p., l'ordinanza di misura cautelare deve contenere «l'esposizione e l'autonoma valutazione delle specifiche esigenze cautelari e degli indizi che giustificano in concreto la misura disposta».²⁸ Questa prescrizione non solo assicura l'imparzialità del giudice evitando che egli si appiattisca al pubblico ministero, ma consentirebbe altresì di scongiurare un eventuale cieco affidamento al *software* predittivo. Invero, il giudice che si avvarrà della prognosi del *risk assessment tool* al fine di disporre la misura cautelare sarà tenuto a indicare gli elementi a sostegno del risultato algoritmico.²⁹

5. Il divieto di perizia criminologica

È diffusa l'idea che il nostro legislatore processuale abbia sempre avuto una scarsa fiducia nei confronti della scienza psico-criminologica. Con ogni probabilità si deve a ciò il dettato dell'art. 220, comma 2, c.p.p., il quale stabilisce che, «salvo quanto previsto ai fini dell'esecuzione della pena o della misura di sicurezza, non sono ammesse perizie per

²⁵ Art. 133, comma 2, c.p.; M. Gialuz, *op. cit.*, 19.

²⁶ Vedasi *supra*, cap. II, par. 1.

²⁷ P.P. Paulesu, *op. cit.*, 14 s.

²⁸ Art. 292, comma 2, lett. c), c.p.p.

²⁹ P.P. Paulesu, *op. cit.*, 16.

stabilire l'abitudine o la professionalità nel reato, la tendenza a delinquere, il carattere e la personalità dell'imputato e in genere le qualità psichiche indipendenti da cause patologiche». ³⁰ Tale norma mira a impedire la sovrapposizione tra valutazione della personalità e valutazione della responsabilità penale. ³¹

Al fine di determinare se il divieto in questione sia estendibile agli strumenti di valutazione del rischio, è necessario comprendere se l'impiego di questi algoritmi sia classificabile come perizia criminologica. In dottrina si è affermato che una tale assimilazione non sia così scontata, in quanto non sembra che i *risk assessment tools* «siano finalizzati specificamente a scandagliare il foro interiore dell'interessato» onde stabilire gli elementi di cui al secondo comma dell'art. 220 cit. ³² Parrebbe dunque opportuno escludere l'applicabilità della disposizione in questione. Diversamente, vi è chi ritiene che, essendo basati su teorie psico-criminologiche, tali *software* non possano trovare ingresso nella fase cognitiva del processo penale. ³³ Viceversa, alla luce del tenore letterale della norma, può ritenersi pacifica l'ammissibilità di detti *tools* in fase esecutiva. ³⁴

6. Il decreto legislativo n. 51 del 2018

Specifiche garanzie in ambito penale sono dettate dal decreto legislativo 18 maggio 2018, n. 51, il quale recepisce nell'ordinamento italiano le disposizioni della citata direttiva 680/2016. L'atto normativo si applica al trattamento dei dati personali da parte delle autorità a fini di prevenzione, investigazione e repressione di reati. ³⁵ Si tratta di una tipologia di trattamento che esula dall'ambito di applicazione materiale del regolamento 679/2016 (c.d. GDPR). ³⁶

L'art. 8 d.lgs. 51/2018 introduce il divieto generale di «decisioni basate unicamente su un trattamento automatizzato, compresa la profilazione, che producono effetti negativi nei confronti dell'interessato». Tali decisioni sono invece ammissibili qualora «siano autorizzate dal diritto dell'Unione europea o da specifiche disposizioni di

³⁰ M. Gialuz, *op. cit.*, 20.

³¹ P.P. Paulesu, *op. cit.*, 14.

³² In tal senso, M. Gialuz, *op. cit.*, 20.

³³ Sul punto, S. Quattrocchio, *Risk assessment: sentencing o non sentencing?*, in *Giurisdizione penale: intelligenza artificiale ed etica del giudizio*, Ed. Giuffrè, Milano, 2021, 80.

³⁴ L. Notaro, *op. cit.*, 146.

³⁵ Art. 1 d.lgs. 51/2018.

³⁶ Art. 2, par. 2, lett. d), GDPR.

legge». All'interessato deve essere riconosciuto, in ogni caso, il «diritto di ottenere l'intervento umano».³⁷ Inoltre, se le decisioni interamente automatizzate si basano sulle categorie particolari di dati personali *ex art. 9 GDPR*³⁸, debbono essere previste «misure adeguate a salvaguardia dei diritti, delle libertà e dei legittimi interessi dell'interessato». Infine, la norma si riferisce esplicitamente al citato art. 21 CDFUE e scongiura la discriminazione degli individui sulla base delle categorie particolari di dati personali.³⁹

Quanto ai *risk assessment tools*, appare convincente l'idea secondo cui il loro impiego non costituisca una decisione basata unicamente su un trattamento automatizzato, a patto che tali algoritmi siano utilizzati soltanto come supporto decisionale e che la decisione finale sia presa dal giudice umano.⁴⁰ A tal proposito, si è rilevato anche il dovere del giudice di valutare l'*output* dell'applicativo congiuntamente ad altri elementi probatori a supporto dell'attendibilità dello stesso.⁴¹

Un'altra disposizione di rilievo è l'art. 23 del suddetto decreto. Esso richiede che, prima di procedere al trattamento, il titolare⁴² compia una valutazione d'impatto sulla protezione dei dati nel caso in cui «il trattamento, per l'uso di nuove tecnologie e per la sua natura, per l'ambito di applicazione, per il contesto e per le finalità, present[i] un rischio elevato per i diritti e le libertà delle persone fisiche».⁴³ Chi scrive ritiene che, considerati l'idoneità degli strumenti di valutazione del rischio a incidere sulla libertà personale dell'accusato, l'alto rischio di discriminazioni e la natura sensibile dei dati personali utilizzati, il trattamento operato da tali *software* rientri nella fattispecie dell'art. 23 cit.

7. I principi della Carta etica della CEPEJ

La presente analisi non può prescindere dalla disamina dei principi sanciti dalla già accennata *Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, elaborata dalla Commissione per l'efficienza della

³⁷ Art. 8, commi 1 e 2, d.lgs. 51/2018.

³⁸ L'art. 9 GDPR include tra le categorie particolari di dati personali quelli genetici e biometrici e quelli che rivelano l'origine etnica, le opinioni politiche, le convinzioni religiose o filosofiche e l'orientamento sessuale della persona fisica.

³⁹ Art. 8, commi 3 e 4, d.lgs. 51/2018.

⁴⁰ In tal senso, M. Gialuz, *op. cit.*, 16 s.

⁴¹ P.P. Paulesu, *op. cit.*, 6.

⁴² In riferimento al possibile uso dei *risk assessment tools* da parte dei tribunali italiani, il titolare del trattamento potrebbe verosimilmente essere il Ministero della giustizia.

⁴³ Art. 23, comma 1, d.lgs. 51/2018.

giustizia del Consiglio d'Europa. Seppur non cogente, il documento è rivolto sia ai legislatori, i quali dovrebbero tenere conto dei principi ivi contenuti durante l'attività di elaborazione normativa, sia agli operatori giuridici, i quali dovrebbero tenerne conto nella fase di applicazione delle norme.⁴⁴

Il primo principio sancito dalla Carta etica è quello del rispetto dei diritti fondamentali, il quale è volto ad assicurare che lo sviluppo e l'impiego dei sistemi di IA avvengano compatibilmente con i diritti fondamentali garantiti dalla CEDU, nonché dalla c.d. Convenzione 108+.⁴⁵ Si fa riferimento, in particolare, al diritto di accesso al giudice e al diritto a un equo processo. Inoltre, la Carta auspica che, sin dalla fase di progettazione degli algoritmi, siano previste misure tali da impedire la violazione delle garanzie tutelate da dette convenzioni.⁴⁶

Il secondo principio è quello di non discriminazione. Esso assume particolare rilevanza in relazione ai *risk assessment tools* ed esprime la necessità di prevenire lo sviluppo e l'intensificazione di qualsiasi forma di discriminazione. Richiede a tale scopo che siano elaborate misure correttive atte a limitare o, meglio, neutralizzare gli effetti discriminatori.⁴⁷

Il terzo principio riguarda la qualità e sicurezza dei sistemi di IA e dei dati da essi utilizzati. Nello specifico, si raccomanda che siano garantite l'integrità e l'immodificabilità del *dataset* sulla cui base viene addestrato l'algoritmo. L'intero processo deve quindi essere tracciabile, al fine di assicurare che non vi siano state alterazioni idonee a incidere sulla decisione finale.⁴⁸

Il quarto principio riconosciuto dalla Carta è quello di trasparenza, imparzialità ed equità. In virtù di tale canone, è necessario raggiungere il giusto equilibrio tra la tutela della proprietà intellettuale dello sviluppatore e l'esigenza di trasparenza. Quest'ultima è intesa come conoscibilità del codice sorgente e delle modalità con cui opera l'algoritmo. All'uopo viene proposta la creazione di un meccanismo di certificazione gestito dalle

⁴⁴ C. Cavaceppi, *L'intelligenza artificiale applicata al diritto penale*, in G. Taddei Elmi – A. Contaldo (a cura di), *Intelligenza artificiale. Algoritmi giuridici. Ius condendum o "fantadiritto"?*, Pacini Giuridica, Pisa, 2020, 101.

⁴⁵ Per Convenzione 108+ si intende la Convenzione n. 108 sulla protezione delle persone rispetto al trattamento automatizzato di dati a carattere personale, adottata dal Consiglio d'Europa nel 1981 e modificata nel 2018.

⁴⁶ CEPEJ, *Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, 2018, 7.

⁴⁷ *Ivi*, 8.

⁴⁸ *Ivi*, 10.

autorità. Infine, per quanto concerne l'imparzialità ed equità, esse sono da considerarsi, rispettivamente, come assenza di *bias* e anteposizione degli interessi di giustizia.⁴⁹

L'ultimo principio è quello del c.d. controllo da parte dell'utilizzatore, in forza del quale gli operatori di giustizia non dovrebbero essere vincolati alla soluzione suggerita dall'IA. Viene poi affermato, anche in questa sede, il diritto del destinatario della decisione a ottenere l'intervento umano.⁵⁰

8. Considerazioni di insieme

Tenuto conto delle norme esaminate nel presente capitolo, non è da escludersi aprioristicamente l'ammissibilità dei *risk assessment tools* nel processo penale italiano.⁵¹ Sarebbe «irragionevole – oltre che antistorico – rinunciare alle componenti positive che lo sviluppo tecnologico offre al sistema penale».⁵² Invero, gli algoritmi in questione potrebbero fungere da utile contrappeso ai «temuti e criticati errori umani».⁵³ Sarà doveroso, tuttavia, garantire il giusto equilibrio tra l'apporto della c.d. *machina sapiens* e l'operato del giudice, preservando un approccio antropocentrico all'intelligenza artificiale.

Occorrerà senz'altro assicurare la piena trasparenza dell'algoritmo: solo così l'indagato o imputato avrà l'effettiva possibilità di contestare l'*output* dello strumento predittivo. Sarà poi necessario impedire che si verifichino o acuiscano discriminazioni di ogni natura, al fine di rispettare il principio di eguaglianza. Da ultimo, la valutazione del *software* dovrà costituire un mero supporto informativo alla decisione umana, sì da garantire l'imparzialità del giudice e la sua autonoma valutazione.

⁴⁹ *Ivi*, 11.

⁵⁰ *Ivi*, 12.

⁵¹ Dello stesso avviso è Giovanni Canzio (cfr. Appendice, III e VI).

⁵² V. Manes, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in *Discrimen*, 15 maggio 2020, 19.

⁵³ P.P. Paulesu, *op. cit.*, 14.

CONCLUSIONI

Come è stato evidenziato nel corso della trattazione, l'impiego dei *risk assessment tools* nel processo penale non è affatto esente da criticità. Anzitutto, sussiste il rischio che tali strumenti replichino i pregiudizi già presenti nelle decisioni giudiziarie sulla cui base vengono addestrati. Vi è poi un'intrinseca assenza di trasparenza dovuta all'imperscrutabilità dei sistemi di autoapprendimento, nonché al segreto industriale generalmente invocato dagli sviluppatori per esimersi dalla divulgazione del codice sorgente. Inoltre, è stato riscontrato come i *software* in questione si avvalgano di mere generalizzazioni statistiche al fine di valutare la specifica pericolosità del singolo individuo: ciò determinerebbe una de-individualizzazione delle decisioni. È stato messo in luce anche il rischio che il giudice si allinei passivamente all'esito dell'algoritmo-oracolo.

A differenza di quanto avviene nell'Unione europea, la quale è caratterizzata da un approccio estremamente cauto in relazione all'impiego delle tecnologie di giustizia predittiva, negli Stati Uniti si fa uso dei *risk assessment tools* già da diversi anni. Si è quindi ritenuto necessario analizzare in questa sede la sentenza *Loomis*, decisione di essenziale riferimento nella materia in esame. Con detta pronuncia, i supremi giudici del Wisconsin hanno legittimato l'uso dell'applicativo *COMPAS* in fase di irrogazione della sanzione penale, non riscontrando alcuna violazione del diritto dell'imputato a un giusto processo. Essi hanno nondimeno delineato una significativa limitazione secondo cui l'*output* dell'algoritmo predittivo non può costituire il solo fattore considerato dal giudice nel commisurare la pena.

Il fine del presente studio è stato quello di valutare la potenziale ammissibilità dei *risk assessment tools* nell'ordinamento processualpenalistico italiano. In tale direzione, si è ritenuto doveroso analizzare la compatibilità con i principi di eguaglianza e del giusto processo, a livello sia costituzionale sia convenzionale. Ne sono risultati, in primo luogo, il divieto d'uso di algoritmi che possano condurre a effetti discriminatori e, in secondo luogo, la necessità di pubblicazione del codice sorgente e della logica adottata dall'applicativo – come peraltro statuito dal Consiglio di Stato nella citata sentenza 881/2020. Vi è inoltre la verosimile esigenza che, nella motivazione del provvedimento, il giudice indichi le ragioni che lo abbiano indotto a conformarsi alla prognosi dello strumento predittivo: ciò parrebbe idoneo a salvaguardare l'indipendenza dell'organo

giudicante. Infine, si è visto come il ricorso all'algoritmo predittivo in funzione di semplice supporto alla decisione del giudice non integrerebbe quel processo decisionale completamente automatizzato proibito dall'art. 8 d.lgs. 51/2018.

Alla luce delle summenzionate considerazioni, chi scrive ritiene che i *risk assessment tools* siano idonei a trovare ingresso nel processo penale italiano. Chiaramente, ciò dovrà avvenire all'interno della cornice che si è cercato di delineare con attenzione nelle precedenti pagine: ineludibile sarà il rispetto dei principi fondamentali posti a salvaguardia dei diritti della persona. Solo così sarà possibile scongiurare il rischio che si giunga a un deleterio assolutismo tecnologico.

BIBLIOGRAFIA

A. Facchini – A. Termine, *Explainable AI: come andare oltre la black box degli algoritmi*, Agenda Digitale, 2022, <https://www.agendadigitale.eu/cultura-digitale/explainable-ai-come-andare-oltre-la-black-box-degli-algoritmi/>.

A. Santosuosso, *Intelligenza artificiale e diritto. Perché le tecnologie di IA sono una grande opportunità per il diritto*, Mondadori Università, Milano, 2020.

A. Traversi, *Giustizia predittiva: quale futuro?*, Altalex, 2023, <https://www.altalex.com/documents/news/2023/03/15/giustizia-predittiva-qual-e-futuro>.

A.M. Maugeri, *L'uso di algoritmi predittivi per accertare la pericolosità sociale: una sfida tra evidence based practices e tutela dei diritti fondamentali*, in *Archivio Penale*, 1/2021, 1.

C. Castelli, *Giustizia predittiva: così a Brescia l'intelligenza artificiale aiuta a prevedere tempi e esiti delle cause*, Agenda Digitale, 2019, <https://www.agendadigitale.eu/cultura-digitale/giustizia-predittiva-cosi-a-brescia-lintelligenza-artificiale-aiuta-a-prevedere-tempi-e-esiti-delle-cause/>.

C. Cavaceppi, *L'intelligenza artificiale applicata al diritto penale*, in G. Taddei Elmi – A. Contaldo (a cura di), *Intelligenza artificiale. Algoritmi giuridici. Ius condendum o "fantadiritto"?*, Pacini Giuridica, Pisa, 2020.

CEPEJ, *Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, 2018 (consultabile al sito <https://rm.coe.int/carta-etica-europea-sull-utilizzo-dell-intelligenza-artificiale-nei-si/1680993348>).

Commissione europea, *Proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale* (doc. 52021PC0206), 2021.

E. Lupo, *Le garanzie di contesto: la parità tra le parti*, in *La legislazione penale*, 10/2020, 20.

E. Nagni, *Artificial Intelligence, l'innovativo rapporto di (in)compatibilità fra machina sapiens e processo penale*, in *Sistema penale*, 7/2021, 5.

E. Yong, *A Popular Algorithm Is No Better at Predicting Crimes Than Random People*, *The Atlantic*, 2018, <https://www.theatlantic.com/technology/archive/2018/01/equivalent-compass-algorithm/550646/>.

F. Basile, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Diritto Penale e Uomo*, 10/2019, 1.

G. Canzio, *Intelligenza artificiale e processo penale*, in G. Canzio – L. Lupária (a cura di), *Prova scientifica e processo penale*, CEDAM, Padova, 2022.

G. Canzio, *La motivazione della sentenza e la prova scientifica*, in G. Canzio – L. Lupária (a cura di), *Prova scientifica e processo penale*, CEDAM, Padova, 2022.

G. Meliota, *Intelligenza Artificiale e giustizia predittiva*, in *Diritto del risparmio*, 2/2022, 205.

G. Ubertis, *Intelligenza artificiale, giustizia penale, controllo umano significativo*, in *Giurisdizione penale: intelligenza artificiale ed etica del giudizio*, Ed. Giuffrè, Milano, 2021.

H. Liu *et al.*, *Beyond State v. Loomis: Artificial Intelligence, Government Algorithmization, and Accountability*, in *International Journal of Law and Information Technology*, vol. 27(2).

J. Angwin *et al.*, *Machine bias. There's software used across the country to predict future criminals. And it's biased against blacks*, ProPublica, 2016, <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

J. Dressel – H. Farid, *The accuracy, fairness, and limits of predicting recidivism*, in *Science Advances*, 4(1)/2018.

J. Larson *et al.*, *How We Analyzed the COMPAS Recidivism Algorithm*, ProPublica, 2016, <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>.

J. Nieva-Fenoll, *Intelligenza artificiale e processo*, Giappichelli, Torino, 2019.

K. Kirkpatrick, *It's not the algorithm, it's the data*, in *Commun. ACM*, vol. 60(2), 2017, 21.

L. Grossi, *Software predittivi e diritto penale*, in C. Chinnici (a cura di), *Intelligenza artificiale e giustizia penale*, Paruzzo, Caltanissetta, 2020.

L. Notaro, *Intelligenza artificiale e giustizia penale*, in C. Chinnici (a cura di), *Intelligenza artificiale e giustizia penale*, Paruzzo, Caltanissetta, 2020.

L. Romanò, *Intelligenza artificiale come prova scientifica nel processo penale*, in G. Canzio – L. Lupária (a cura di), *Prova scientifica e processo penale*, CEDAM, Padova, 2022.

L. Viola, *Giustizia predittiva*, in *Diritto on line*, Treccani, https://www.treccani.it/enciclopedia/giustizia-predittiva_%28Diritto-on-line%29/.

M. Barberis, *Giustizia predittiva: ausiliare e sostitutiva. Un approccio evolutivo*, in *Milan Law Review*, 2/2022, 1.

M. Gialuz, *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa*, in *Diritto penale contemporaneo*, 2019, 1.

M. Martorana, *Giustizia predittiva: rischi, vantaggi e un modello italiano*, CFnews.it, 2021, <https://www.cfnews.it/avvocatura/giustizia-predittiva-rischi-vantaggi-e-un-modello-italiano/>.

P. Ferrua, *Il processo penale dopo la riforma dell'articolo 111 della Costituzione*, in *Questione giustizia*, 1/2000.

P.P. Paulesu, *Intelligenza artificiale e giustizia penale. Una lettura attraverso i principi*, in *Arch. pen. web*, 1/2022.

Q. Camerlengo – F. Furlan, *Lezioni di diritto costituzionale vivente*, CEDAM, Padova, 2021.

R.E. Kostoris, *Predizione decisoria e diversion processuale*, in *Giurisdizione penale: intelligenza artificiale ed etica del giudizio*, Ed. Giuffrè, Milano, 2021.

S. Carrer, *Se l'amicus curiae è un algoritmo: il chiacchierato caso Loomis alla Corte Suprema del Wisconsin*, in *Giurisprudenza Penale Web*, 4/2019.

S. Quattrocolo, *Processo penale e rivoluzione digitale: da ossimoro a endiadi?*, in *MediaLaws*, 3/2020, 121.

S. Quattrocolo, *Risk assessment: sentencing o non sentencing?*, in *Giurisdizione penale: intelligenza artificiale ed etica del giudizio*, Ed. Giuffrè, Milano, 2021.

S. Signorato, *Giustizia penale e intelligenza artificiale. Considerazioni in tema di algoritmo predittivo*, in *Riv. dir. proc.*, 2020.

S.n., *Giusto processo*, in *Enciclopedia on line*, Treccani, <https://www.treccani.it/enciclopedia/giusto-processo/>.

S.n., *Practitioner's Guide to COMPAS Core*, Equivant, 2019, <https://www.equivant.com/wp-content/uploads/Practitioners-Guide-to-COMPAS-Core-040419.pdf>.

S.n., *Principio del contraddittorio. Diritto processuale penale*, in *Enciclopedia on line*, Treccani, <https://www.treccani.it/enciclopedia/principio-del-contraddittorio-diritto-processuale-penale>.

S.n., *SARI = Sistema Riconoscimento Automatico Immagini. La Polizia di Stato ad Arezzo individua e denuncia i primi 7 ladri*, Questura di Arezzo, 2018, <https://questure.poliziadistato.it/it/Arezzo/articolo/10945bb6148a60216596783584>.

V. Cavosi, *Governare l'intelligenza artificiale. Spunti per la progettazione di sistemi di IA legali, etici e robusti*, Ledizioni, Milano, 2022.

V. Denti, *sub Art. 111 Cost.*, in *La magistratura*, vol. IV, in G. Branca (a cura di), *Commentario alla Costituzione*, Zanichelli, Bologna, 1987.

V. Manes, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in *Discrimen*, 15 maggio 2020.

W. Dieterich *et al.*, *COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity*, Northpointe, 2016, http://go.volarisgroup.com/rs/430-MBX-989/images/ProPublica_Commentary_Final_070616.pdf.

APPENDICE

INTERVISTA AL DOTT. GIOVANNI CANZIO

In data 19 marzo 2023 ho avuto l'opportunità di intervistare Giovanni Canzio, Primo Presidente emerito della Corte di cassazione. Di seguito è riportata la trascrizione integrale dell'intervista. Le parti in corsivo si riferiscono alle domande che ho posto all'intervistato.

In Italia si sta già facendo largo uso dell'intelligenza artificiale nell'ambito delle cc.dd. indagini digitali, ma l'impiego di strumenti di giustizia predittiva a fini decisori non appare ancora così vicino. Da cosa deriva, a Suo avviso, questa doppia velocità?

Il sistema di giustizia predittiva di cui oggi si avvalgono le corti statunitensi (in particolare per il calcolo della recidivanza del condannato) ai fini della fissazione della pena o della *probation* non è ancora disciplinabile e gestibile né in Italia né nel resto d'Europa. La prima ragione è che l'applicazione pratica di queste tecniche di IA è stata sottoposta a critiche serrate per quanto riguarda eventuali esiti di tipo discriminatorio. La seconda ragione è l'assenza di trasparenza legata all'impossibilità di controllare il codice sorgente e le informazioni che sono state immesse. Questo perché i *software* utilizzati [negli USA] sono prodotti da imprese private, le quali godono, oltre che del diritto di brevetto, dell'esclusiva. Le corti statunitensi non sono quindi obbligate a rivelare il come si è pervenuti a quel risultato, ovvero quali e quante informazioni di base sono presenti nel *database*. Si tratta inoltre di macchine che, nel tempo, si arricchiscono di informazioni (autoapprendimento). Manca quindi il momento – fondamentale nell'Europa continentale – del controllo della macchina e delle tecniche con cui sono stati elaborati gli algoritmi da cui si inferiscono determinati risultati applicabili al processo.

Nel campo del diritto amministrativo (ad es. nella gestione dei bandi di gara) la Pubblica Amministrazione sta già utilizzando algoritmi e modelli matematici per valutare l'affidabilità di un'impresa, escluderla e così via. Le aziende che hanno fatto ricorso ai vari TAR hanno sempre vinto sotto il profilo della controllabilità e della trasparenza: esse hanno diritto di conoscere come si è formato il *database*, da quali elementi si è giunti a un certo risultato, ecc.

Nell'Europa continentale, quello che io chiamo «modello forte di IA», cioè quello che simula in modo sconfinato il processo cognitivo dell'uomo, non ha possibilità di essere immesso. Questo perché, in Europa, esistono regole di origine costituzionale e

sovranaZIONALE per cui il giusto processo non può non essere celebrato e deciso che da uomini che si confrontano in un contraddittorio per l'accertamento del fatto attraverso le prove. Ciò comporta la rinuncia dell'Europa continentale allo sviluppo della tecnologia e della scienza? No, perché facendo così commetterebbe un grande errore nei confronti del futuro. L'Europa si sta muovendo, piuttosto, in un terreno complesso e delicato, provando a creare un ordinamento che regoli l'ingresso di queste forme di IA nel processo: si tratta di quello che io definisco «modello debole di IA», cioè di tipo collaborativo. Normalmente il giudice decide in base alle informazioni che riceve dal mondo reale. Se, però, l'algoritmo matematico-statistico arricchisce le informazioni del giudice, noi ne traiamo un vantaggio. Il giudice acquisisce così più informazioni che però deve elaborare: non può essere il mero consumatore di tali informazioni. In particolare, egli non può bypassare il contraddittorio, il quale è costituzionalmente doveroso (art. 111 Cost.). Esistono il contraddittorio *per la* prova e il contraddittorio *sulla* prova. Nel contraddittorio *per la* prova si discute dell'ammissibilità della prova in giudizio; nel contraddittorio *sulla* prova si discute della valutazione della prova, che avviene attraverso il dialogo tra accusa e difesa all'interno del processo. Per poter esercitare il contraddittorio le parti devono conoscere le prove in tutta la loro composizione (cioè come si sono formate, se si sono legittimamente formate, ecc.). Non è possibile contraddire il risultato di un'elaborazione di tipo algoritmico se non si sa quali informazioni sono state immesse.

In Europa continentale il giudice non decide con un verdetto [come negli USA], bensì con una sentenza che va motivata: deve quindi spiegare com'è giunto alla decisione, non può semplicemente richiamarsi come “consumatore” al risultato matematico-statistico.

Tutto ciò comporta che, nel momento dell'ammissione e in quello della valutazione della prova, il giudice debba controllare il risultato probatorio. L'interrogativo inquietante di fronte al quale ci troviamo è come organizzare tutto questo all'interno del processo.

Come Lei ha anticipato, la giurisprudenza amministrativa ha già avuto modo di esprimersi sull'uso dell'IA nel procedimento amministrativo informatico. In particolare, ha declinato il principio di conoscibilità dell'algoritmo: ciò differisce dagli Stati Uniti, dove – anche in questo caso – viene tutelato il segreto industriale dello sviluppatore.

Ritiene che una simile logica verrà estesa anche agli algoritmi predittivi nel processo penale?

Questo è il nodo cruciale a cui ci troviamo di fronte e riguarda, a maggior ragione, il processo penale, il quale è – per definizione – il «giusto processo». È sufficiente pensare alla presunzione d'innocenza dell'imputato, all'onere della prova che è solo a carico dell'accusa, al diritto di difesa nel contraddittorio *per la prova e sulla prova*, alla motivazione che il giudice deve dare alla sentenza, alla regola conclusiva del giudizio dell'«al di là di ogni ragionevole dubbio». A maggior ragione c'è un'esigenza di trasparenza informativa sulla consistenza e sul funzionamento di questa c.d. *machina sapiens*, cioè del modello statistico-matematico che elabora l'algoritmo. Questo è un nodo cruciale perché questi principi, oltre ad essere sanciti a livello costituzionale, sono riconosciuti dalla Convenzione europea dei diritti dell'uomo: non possiamo quindi esimerci dal loro rispetto.

Nel 2021 la Commissione europea ha elaborato una proposta di regolamento relativo all'uso dell'IA negli ambienti di giustizia. Dalle notizie che ho, questa proposta (già corredata da una serie di emendamenti derivanti dalla discussione già avvenuta in Consiglio e in Parlamento) dovrebbe essere definitivamente adottata nei prossimi mesi. Lo schema di base di questo testo legislativo è tratto dal GDPR. Anche in questo ambito saranno previste un'*authority* nazionale e un'*authority* europea. Si immagina inoltre una distinzione tra forme di IA a basso, medio e alto rischio. Ci saranno comunque una serie di presidi volti a garantire che l'amministrazione della giustizia rimanga antropocentrica, garantendo che il controllo finale nei confronti dei dati sia affidato all'uomo.

A mio avviso, la giustizia c.d. predittiva non può essere fermata a priori con il “*diktat*” di una qualsiasi autorità, ma è necessario – a garanzia dei diritti fondamentali della persona – che sia disciplinata e regolamentata.

Nel settembre del 2021, il Consiglio per i diritti umani delle Nazioni Unite ha pubblicato un documento molto articolato che mette in guardia l'umanità dei rischi per la tutela dei diritti umani derivanti da un uso sconfinato dell'IA.

Il celebre software COMPAS basa il proprio “verdetto” su un questionario di 137 domande che viene chiesto di compilare all'imputato. A mio avviso, un rischio che potrebbe emergere è che la stessa difesa possa avvantaggiarsi della conoscibilità del

codice sorgente, analizzandolo per poi “pilotare” le risposte che l’imputato fornirà. Si genererebbe così un tasso di rischio alterato. Secondo Lei, ciò potrebbe verosimilmente accadere?

Se pensiamo alla patologia, è chiaro che le risposte sono difficili. Tuttavia, non possiamo immaginare un mondo patologico: dobbiamo immaginare la fisiologia, perché di fronte a qualsiasi prova c’è il rischio di interferenze, frodi e abusi. Ad ogni modo, quei tipi di test pongono una serie di domande che attengono a circostanze della vita, dell’ambiente e della società in cui il soggetto vive.

Nella sentenza *Loomis* c’è un passaggio molto importante in cui la Corte Suprema del Wisconsin afferma che il *software COMPAS* «*should always constitute merely one tool available to a Court and needs to be confirmed by additional sound information*». Questo è molto vicino alla nostra concezione del rapporto collaborativo macchina-uomo e della necessità di riscontri nel mondo reale.

Nell’opera Prova scientifica e processo penale, Lei ha fatto riferimento agli evidenti risultati pratici di questi algoritmi, soprattutto «in termini di risparmio di tempi e costi, di semplificazione delle procedure e di tendenziale calcolabilità e uniformità delle decisioni» (p. 906). Crede che l’uso dei risk assessment tools nel nostro ordinamento possa pertanto rispondere alle esigenze di buon andamento dell’attività amministrativa, con particolare riferimento ai principi di efficacia ed efficienza?

Il ricorso a questi strumenti diventerà sempre più ineludibile; non possiamo eludere lo sviluppo scientifico. Dobbiamo però, come esseri umani, provare a controllare questo fenomeno, a gestirne i confini. Lo abbiamo fatto con la c.d. prova scientifica e dobbiamo farlo anche con questa prova digitale “estrema” che è la prova algoritmica. Questo ci porta sul terreno dell’implementazione della professionalità degli operatori (avvocati e magistrati). Gli operatori dovranno essere in grado di controllare gli schemi epistemologici e logici che nascono dall’utilizzo di questi modelli matematico-statistici. È fondamentale comprendere come funzionano per stabilire se il passaggio dall’informazione di base alla risposta fornita dall’algoritmo è corretto e trasparente. Certamente è un lavoro duro, però io sono ottimista e non credo che stiamo assistendo a

un passaggio di testimone dall'uomo alla macchina. Dipenderà – e io sono sicuro che l'uomo ce la farà – dalla capacità dell'uomo di controllare questi sviluppi così prepotenti.

Qualora nel nostro ordinamento venissero utilizzati i risk assessment tools, come potrebbe difendersi un imputato al quale viene assegnato un alto tasso di probabilità di recidiva? È evidente come lo squilibrio di mezzi a disposizione sia notevole: per la difesa risulterebbe estremamente difficile sostenere l'infondatezza dei complessi calcoli statistico-probabilistici dell' algoritmo o, addirittura, l'esistenza di bias nel database stesso. L'algoritmo è destinato ad essere pressoché insindacabile?

Ha ragione. Infatti, uno dei temi fondamentali sarà proprio l'esercizio del diritto di difesa di fronte a un apparato tecnologico che può essere a disposizione soltanto dello Stato e di istituzioni molto ricche. Però, la difesa non è completamente priva di armi: una difesa povera può avere a disposizione o lo Stato che la sorregge (ma con costi altissimi) o lo spazio della confutazione alternativa. Con quest'ultima si intende la possibilità per l'imputato di confutare, attraverso il contraddittorio, il riscontro fornito dall'algoritmo. Naturalmente, se ho molti soldi posso rivolgermi anche io a una società privata e commissionare un *software* simile che elabora informazioni contrastanti tali da mettere in dubbio le risposte fornite dal *software* dell'accusa.

Io ricordo che, venti o trenta anni fa, con l'introduzione della prova scientifica e della prova digitale nel processo, i timori che il giudice soccombesse alla potenza di quelle nuove prove non mancavano. Poi si è visto che epistemologia, logica e regole costituzionali riuscivano a governare ancora il processo. Bisogna essere fiduciosi della capacità dell'uomo di formarsi professionalmente in modo da saper controllare questi inediti esiti della tecnologia e della scienza. Non bisogna arrendersi: questo è il messaggio che lascio. Vedo comunque che l'Europa continentale ha già accolto la sfida. Leggeremo con attenzione il nuovo regolamento di cui parlavo e in quali forme si esplicherà.

Un'altra possibile problematica è che la decisione si appiattisca sull'esito dell'algoritmo predittivo. L'impiego dell'IA mette quindi in pericolo anche il principio del libero convincimento del giudice?

Io non credo molto al tradizionale libero convincimento: il giudice dovrebbe essere libero da pregiudizi e vincoli, ma non è libero rispetto alle regole dell'ordinamento. Io parlo più di discrezionalità tecnica del giudice, il quale si muove sui binari dettati della Costituzione, dal codice di rito e da tutta una serie di regole epistemologiche e logiche. Detto ciò, io rimango dell'idea che questo ordinamento sia in grado di fronteggiare anche la nuova sfida che proviene dall'IA. Ci troveremo ad interagire con un nuovo sistema di fonti probatorie con cui dovremo saper fare i conti. È vero che, già oggi, il giudice lavora con probabilità e prognosi. Quindi, non si tratta di un giudice "sprovvisto" dal punto di vista della probabilità. Tuttavia, egli dovrà essere molto più provveduto nel capire come si arriva a quell'esito probabilistico. L'importante è che il giudice sia consapevole che non può mai diventare un mero consumatore di un algoritmo.

Dunque, Lei accoglierebbe con favore l'uso di questi sistemi?

Con quei limiti e in quei termini di cui ho parlato, sì. Ciò significa nei confini del giusto processo, così come disegnato dalla Costituzione e da tutte quelle regole del processo penale che vanno dalla valutazione della prova, al contraddittorio, alla decisione «al di là di ogni ragionevole dubbio», ecc. Da questi binari il giudice non può e non deve uscire.

RINGRAZIAMENTI

Giunto al termine di questo elaborato, ritengo doveroso ringraziare il prof. Pin per aver accolto la mia proposta di tesi e per avermi assistito nella stesura della stessa. È stato un piacere poter iniziare e concludere con lui il mio percorso di studi. A seguito di questi tre anni da Rappresentante degli studenti, desidero altresì esprimere la mia sincera gratitudine per la sua costante disponibilità e dedizione, nonché per l'impegno che egli ha profuso nella buona riuscita di questo innovativo corso di laurea. In particolar modo, la sua continua attenzione alle esigenze degli studenti ha sempre reso possibile l'instaurazione di un proficuo dialogo costruttivo.

Desidero poi ringraziare il presidente Canzio per avermi offerto la prestigiosa opportunità di intervistarlo. Le tematiche affrontate e gli stimolanti spunti di riflessione che ne sono emersi hanno rappresentato un prezioso valore aggiunto per il mio lavoro di tesi.

Un imprescindibile ringraziamento spetta ai miei genitori: il loro incessante supporto e la fiducia in me riposta mi hanno sempre incoraggiato a guardare oltre e non abbattermi. Inoltre, vorrei esprimere la mia riconoscenza nei loro confronti per avermi dato la possibilità – purtroppo non ancora garantita a tutti – di proseguire gli studi, oltre che per avermi permesso di vivere l'impagabile esperienza dell'Erasmus, la quale ha indubbiamente contribuito a rendermi ciò che sono ora.

Mi preme ringraziare mia sorella per il suo esempio di inarrestabile determinazione e audacia e per avermi supportato in questi mesi dedicati alla scrittura della tesi. Sono certo che, ora, anche lei sarebbe in grado di discutere circa l'ammissibilità dei *risk assessment tools* nell'ordinamento italiano.

Infine, un sentito ringraziamento va ai miei amici di sempre e ai compagni di avventure (e sventure) conosciuti all'università: sono davvero felice di poter condividere con loro questo traguardo. Insieme abbiamo vissuto momenti indelebili che hanno contribuito ad allietare questi anni ed è bello sapere di aver trovato le persone giuste su cui poter contare.