

Umberto Fiore

Lo stile oracolare degli algoritmi: non motivato, con scarse capacità semantiche e ambizioni penal-predittive

Sommario: 1. Proposta di premi fatta dall'Accademia dei sillografi. – 2. Riflessioni preliminari in ambito penale. Predizioni decisorie e limiti normativi. – 3. La valutazione attuariale e gli algoritmi predittivi in materia di pericolosità. – 4. Prime applicazioni in Italia per gli autori di reati sessuali. 5. Conclusioni.

1. Proposta di premi fatta dall'Accademia dei Sillografi.

*“L'Accademia dei Sillografi attendendo di continuo, secondo il suo principale istituto, a procurare con ogni suo sforzo l'utilità comune, e stimando niuna cosa essere più conforme a questo proposito che aiutare e promuovere gli andamenti e le inclinazioni del fortunato secolo in cui siamo, come dice un poeta illustre, ha tolto a considerare diligentemente le qualità e l'indole del nostro tempo, e dopo lungo e maturo esame si è risolta di poterlo chiamare l'età delle macchine, non solo perché gli uomini di oggidì procedono e vivono forse più meccanicamente di tutti i passati, ma eziandio per rispetto al grandissimo numero delle macchine inventate di fresco ed accomodate o che si vanno tutto giorno trovando ed accomodando a tanti e così vari esercizi, **che oramai non gli uomini ma le macchine**, si può dire, trattano le cose umane e fanno le opere della vita...; perciò l'Accademia dei Sillografi reputa essere espedientissimo che gli uomini si rimuovano dai negozi della vita il più che si possa, e che a poco a poco diano luogo, **sottentrando le macchine in loro scambio**. E deliberata di concorrere con ogni suo potere al progresso di questo nuovo ordine delle cose, propone per ora tre premi a quelli che troveranno le tre macchine infrascritte.*

L'intento della prima sarà di fare le parti e la persona di un amico, il quale non biasimi e non motteggi l'amico assente; non lasci di sostenerlo quando l'oda riprendere o porre in giuoco; non anteponga la fama di acuto e di mordace, e l'ottenere il riso degli uomini, al debito dell'amicizia; non divulghi, o per altro effetto o per aver materia da favellare o da ostentarsi, il segreto commessogli; non si prevalga della familiarità e della confidenza dell'amico a soppiantarlo e sopraffonderlo più facilmente; non porti invidia ai vantaggi di quello; abbia cura del suo bene e di ovviare o di riparare á suoi danni, e sia pronto alle sue domande e á suoi bisogni, altrimenti che in parole...

La seconda macchina vuol essere un uomo artificiale a vapore, atto e ordinato a fare opere virtuose e magnanime...

*La terza macchina debbe essere disposta a fare gli uffici di una donna conforme a quella immaginata, parte dal conte Baldassar Castiglione, il quale descrisse il suo concetto nel libro del Cortegiano, parte da altri, i quali ne ragionarono in vari scritti che si troveranno senza fatica, e si avranno a consultare e seguire, come eziandio quello del Conte. Né anche **l'invenzione di questa macchina dovrà parere impossibile agli uomini dei nostri tempi...**” (Leopardi, 1968, pp. 188-192).*

Non vi è molto da aggiungere alla satira di Leopardi che si presenta a dir poco attuale in un periodo in cui la sospensione per un breve periodo di ChatGPT, da parte del Garante della Privacy, ha destato un sentimento avvertito come un danno grave e irreparabile nel

non poter fare domande e ottenere risposte in tempo reale da una macchina che sequenzia le parole in termini di probabilità.

Il punto di svolta è capire se la macchina algoritmo possa essere una risorsa o credendo che sia una risorsa venga utilizzata in processi decisionali in cui i pregiudizi non possono che aumentare in modo esponenziale attraverso una serie di bias la cui visibilità non sempre è a portata di mano.

In effetti Leopardi non svalutava affatto le scoperte scientifiche e tecnologiche; semmai egli rinunciava all'eccessivo entusiasmo che la visione ottimistica del suo tempo riversava sul progresso scientifico. Esattamente lo stesso ottimismo che si riversa sulla macchina algoritmo di oggi; siamo arrivati a pensare di poter affidare tutto alle macchine in virtù di uno spudorato liberismo che per arrivare al risultato finale non fa altro che produrre razzie sociali. Inoltre, una volta per tutte, bisogna ammettere che la spiritualità insita in ogni decisione o azione non è replicabile e se solo si pensa che ciò sia possibile, dobbiamo avere sempre chiara una domanda: siamo noi o è la macchina?

2. Riflessioni preliminari in ambito penale. Predizioni decisorie e limiti normativi.

Diversi sistemi di giustizia penale in tutto il mondo fanno oggi uso di sistemi di intelligenza artificiale per sostenere il processo decisionale umano di diversi attori, come forze dell'ordine, avvocati, giudici.

Lo scenario di base è perlopiù il seguente: il sistema classifica gli individui in classi di riferimento. Tali classi possono esprimere previsioni sul comportamento degli individui, o di gruppi di individui (ad esempio, tasso di recidiva individuale basso/alto/medio; rischio di criminalità in una particolare area geografica). Queste previsioni vengono successivamente impiegate nel processo decisionale algoritmico, cioè per elaborare e suggerire strategie su come trattare tali soggetti in base alla loro classificazione. Un primo modo per utilizzare i sistemi di IA a fini preventivi è quello di applicarli per supportare (o sostituire) l'esperienza investigativa umana con un'analisi integrata dei dati già disponibili, al fine di identificare potenziali modelli criminosi e ridurre la vittimizzazione in ambienti digitali, come i *social media*. Un altro approccio è quello che mira a prevedere le circostanze (tempo e luogo) di possibili reati. Questo approccio rispecchia le tradizionali metodologie investigative di mappatura delle attività criminali in un'area determinata, sulla base dell'analisi di fattori sociali, demografici, economici, ambientali, nonché dei dati relativi ai precedenti penali.

Tuttavia, preliminarmente, rivolgiamo l'attenzione agli obiettivi della predizione decisoria; in effetti il termine è certamente improprio e fuorviante e viene utilizzato solo perché è, ormai, entrato nell'uso. Sostanzialmente non c'è in gioco alcuna 'predizione', ma, solo la possibilità di prevedere quali potranno essere le decisioni future dei giudici su casi analoghi.

D'altronde, è proprio questa prevedibilità a costituire un'aspirazione di fondo della giustizia penale, dove, in ragione della posta in gioco, il valore della certezza giuridica appare ancor maggiormente avvertito che in altri settori dell'ordinamento in quanto insidiato dal pluralismo giuridico contemporaneo, dal venir meno della struttura gerarchica delle fonti e da un diritto a larga matrice giurisprudenziale (Fiandaca, 2011, pp. 79 ss.).

La IA sembrerebbe promettere di sostituire a questo disordine un paesaggio lineare, riproponendo l'idea tipicamente illuministico-positivista (Caratta, 2020, pp. 491-514) di

un diritto ‘calcolabile’. La fascinazione che può venire da una simile prospettiva non deve, tuttavia, indurre a sopravvalutare le capacità di simili strumenti, coltivando aspettative illusorie nei loro confronti. È dunque necessario cercare anzitutto di comprendere il reale significato e la portata di queste tecniche ‘predittive’ per non domandare ad esse più di quanto siano in grado di offrire.

Il primo elemento da considerare in proposito è rappresentato da un dato strutturale di fondo: la c.d. ‘predizione’ si basa su correlazioni fra enormi masse di decisioni giudiziali, ma non viene ottenuta cercando di riprodurre il pensiero dei giudici, ‘modellizzandone’ cioè in termini informatici il ragionamento. Un approccio di questo tipo era stato seguito inizialmente con i c.d. “sistemi esperti” negli anni 80’ e 90’ (Bourcier, 2011, pp. 37-53), ma non era apparso adeguato alle necessità, in quanto l’applicazione di regole codificate oltre a essere complessa e onerosa, non era comunque in grado di interpretare l’infinita varietà delle situazioni e dei contesti che nel campo della giustizia penale, si manifesta non solo in ragione della complessità, della stratificazione e della disorganicità del quadro normativo, ma anche in ragione delle diverse interpretazioni e delle componenti discrezionali che caratterizzano le decisioni dei giudici.

Così, i sistemi esperti, basati su un ragionamento per regole, sono stati gradualmente sostituiti da sistemi basati su un ragionamento per casi, i quali si caratterizzano per il fatto che i nuovi problemi vengono risolti sulla base di generalizzazioni di soluzioni date in passato a problemi simili. Inoltre, essendo concepiti, a differenza dei sistemi esperti, per lavorare senza regole predefinite, tali sistemi possono operare anche in presenza di dati vaghi, indeterminati, discrezionali (Rossi, 2019, p.67; Quattrocchio, 2020, p.12).

L’elaborazione che compiono attraverso il linguaggio binario della matematica, anche quando impiegano metodi di *machine learning*, si basa sul presupposto che le correlazioni che essi effettuano sui grandi volumi di dati *possa sostituire la comprensione dei nessi causali* (Searle, 1980, pp. 417-424) che caratterizzano una decisione giudiziaria, proponendo una previsione sufficientemente attendibile riguardante le future decisioni di un giudice in situazioni analoghe.

Ciò significa che tali sistemi elaborano con estrema velocità enormi quantità di dati attraverso tecniche di autoapprendimento che consentono loro di risolvere problemi inediti (Rossi, 2019, pp. 60 e 67) non previsti originariamente dal *software*. Tuttavia il punto di caduta sta nel fatto che tali sistemi non comprendono il vero significato delle loro prestazioni in quanto non riescono a elaborare ‘ragionamenti’ che siano di una qualche complessità (Rossi, 2019, pp. 78 e ss.). Viene meno, in questo modo, il tipico strumento di lavoro del giudice rappresentato dal ragionamento giuridico.

Il loro basarsi su classificazioni automatiche, limita anche la possibilità di contestare il risultato ottenuto con argomenti razionali, rischiando paradossalmente di mettere in crisi proprio quella prevedibilità, quella ‘calcolabilità’ del diritto che attraverso di loro si mirava ad ottenere (Lettieri, 2021, p. 90).

Inoltre, pur essendo ormai capaci di raggiungere ottime prestazioni dal punto di vista linguistico, attraverso le generalizzazioni operate su enormi quantità di dati, tali sistemi non comprendono in realtà il significato delle parole; le tecnologie digitali non sono in grado di processare i dati in funzione semantica; e se il trattamento è effettuato solo sulla base del ricorrere dei lemmi, non si deve dimenticare che ogni parola – e anche molte costruzioni grammaticali – assumono un significato diverso in rapporto al contesto nel quale sono

inserite e che, inoltre, esistono parole dai significati plurimi, che possono essere individuati solo attraverso un giudizio di relazione con altri elementi, soprattutto in un settore peculiare come quello giuridico.

Resta però il fatto che – proprio perché non considerano gli aspetti semantici del linguaggio e difettano di capacità raziocinanti – queste tecniche non sono neppure in grado di distinguere le affermazioni e le correlazioni vere da quelle false, come ha confermato anche uno studio effettuato da alcuni ricercatori sulla giurisprudenza della Corte europea dei diritti dell'uomo (Aletras, Tsarapatsanis, Preoțiu-Pietro e Lampos, 2016, p. 93). Nonostante la Carta etica europea sull'uso dell'intelligenza artificiale avverta l'importanza di predisporre “meccanismi per impedire le distorsioni e la discriminazione nel processo stesso di progettazione” degli algoritmi, così da assicurare la loro totale trasparenza e correttezza, qualora un domani essi potessero contribuire all'adozione di decisioni giudiziarie, essa precisa che, allo stato attuale, il problema delle c.d. “false correlazioni” non appare ancora risolto, dato che le medesime “non (sono) discernibili nelle grandi masse di dati”.

Tale carenza si accentua ancor più quando si considera il carattere ‘opaco’ di alcuni sistemi di autoapprendimento, che non consentono di controllare il procedimento attraverso il quale sono giunti a un determinato risultato.

Inoltre, le decisioni dei giudici penali non si riducono a meri ragionamenti giuridici, ma contengono numerose valutazioni discrezionali: da quelle riguardanti la condotta, il nesso causale, le circostanze del reato, alla stessa valutazione delle prove, che si ispira al principio del libero convincimento.

Il giudice deve convincersi al di là di ogni ragionevole dubbio per emettere una decisione di colpevolezza e sulla base dello stesso materiale probatorio due giudici diversi potrebbero anche (legittimamente) pervenire a conclusioni diverse (Ferrua, 2017, p. 95). Siamo dunque di fronte a *valutazioni di tipo qualitativo*, fatte anche di ‘intuizioni’, di ‘emozioni’, di ‘salti di paradigma’ tipicamente e profondamente umani che sfuggono alle *correlazioni meramente quantitative* che è capace di produrre un modello computazionale.

A prescindere dagli impedimenti normativi di cui si dirà dopo, in un sistema come il nostro alla predizione decisoria non può essere attribuito un valore vincolante: l'idea che il giudice debba conformare le sue future decisioni alle ricorrenze statistico-quantitative di un programma di IA (Quattrocchio, 2020, p. 210), indipendentemente e magari anche contro le sue convinzioni personali, ingabbiandolo in un artificioso e rigido determinismo giudiziario (Garapon e Lassègue, 2021, pp. 181 ss.), suonerebbe come uno svilimento del suo ruolo e un attentato alla sua stessa indipendenza; egli diverrebbe ‘soggetto’ all'algoritmo, cioè a un'entità esterna ed estranea alla legge, non prodotta dal potere legislativo, realizzando una sua completa deresponsabilizzazione, perché lo relegherebbe a svolgere un ruolo passivo di tipo esclusivamente notarile. Sarebbe, in definitiva, l'anticamera di un sistema di giustizia amministrata da giudici *robot*.

La prevedibilità delle decisioni potrebbe finire così per essere pagata al prezzo di un tendenziale appiattimento delle decisioni future su quelle passate. Un processo di auto-avveramento suscettibile di produrre un duplice effetto negativo. Anzitutto, quello di recidere il circolo vitale e virtuoso della giurisprudenza, la quale ha invece il compito di ‘leggere’ costantemente il mutamento sociale per adeguarvi l'interpretazione delle regole giuridiche, garantendo così l'evoluzione di un ordinamento. Inoltre, quello di inficiare lo

stesso valore dell'imparzialità del giudice, facendo 'apparire' quest'ultimo in partenza appiattito su esiti preformati.

Da non sottovalutare, poi, anche se forse più di prospettiva, il timore che una eccessiva fascinazione per queste tecniche possa finire per avere un impatto negativo anche su un piano più latamente sociale, esponendo al rischio che vengano introiettati moduli di standardizzazione delle stesse relazioni umane che finirebbero per privare gli uomini del senso di giustizia e di ragionevolezza, sostituendolo con la conformità a parametri, quantità, misure; un modello aberrante a cui alla fine ci si aspetterebbe che si conformasse non solo il giudice in tutta la sua attività decisoria, ma, ancor prima, lo stesso legislatore, eliminando via via ogni forma di valutazione discrezionale.

Infine, bisogna guardarsi dal mito della neutralità dell'algoritmo, che sarebbe insensibile ai soggettivismi e alle passioni umane. Sono in molti ad avvertire come si tratti in realtà di un'illusione, perché questo strumento opera nel modo per il quale è stato programmato a funzionare e, dunque, riflette necessariamente la 'visione', anzitutto etica, del suo programmatore. Se questo fattore può avere un impatto maggiore per gli strumenti predittivi in funzione decisoria, non sembra irrilevante neppure in rapporto alle predizioni decisorie. È allora evidente che, in proporzione all'ampiezza degli spazi che si vorranno affidare a simili strumenti nel campo della giustizia penale, assumerà assoluta centralità, da un lato, la conoscibilità e la controllabilità dell'algoritmo e del funzionamento del programma (ad oggi ancora non sempre realizzabile), dall'altro, la figura dello specialista informatico incaricato di crearlo.

Non per nulla, si insiste molto sulla necessità che l'elaborazione del *software* avvenga anche con la partecipazione essenziale dei giuristi, che dovranno fornire tutte le indicazioni necessarie per rendere lo strumento il più possibile fruibile, sicuro, trasparente e rispettoso delle garanzie (Kostoris, 2021, pp. 5-9).

Nonostante tutto, non si può disconoscere una tendenza generale a voler utilizzare un tipo di giustizia "più facile". Almeno per quello che succede in Italia e in Europa, salvo alcune applicazioni in specifici settori di cui si dirà successivamente, alcuni impedimenti normativi non permettono l'uso delle conquiste della nuova rivoluzione industriale all'interno della macchina giudiziaria (per il momento). Il fatto stesso che si pongano degli argini è sintomo che qualcosa si muove verso la direzione che si cerca di ostacolare.

Non è un caso che la presa di coscienza da parte dell'Europa attraverso la Carta Etica assume un significato particolare proprio perché segna le enormi potenzialità dell'IA e la possibilità di sfruttarla anche nella realtà giudiziaria, se declinata in un certo modo (Quattrocchio, 2019, p.203), fissando al contempo i principi etici da osservare.

La Carta è destinata ad attori pubblici e privati che sviluppano servizi e strumenti di intelligenza artificiale ed è corredata da quattro Appendici che presentano una panoramica dello stato dell'arte nei sistemi giudiziari europei. Nell'introduzione viene riconosciuto il potenziale applicativo degli strumenti di IA per implementare l'efficienza del sistema giudiziario (che contribuirebbero a migliorare la prevedibilità della legge e la coerenza delle decisioni giudiziarie) ma viene anche evidenziata l'importanza di un suo uso nel rispetto dei diritti umani, in particolare di quelli sanciti nella CEDU e della Convenzione sulla protezione delle persone rispetto al trattamento automatizzato di dati di carattere personale.

All'interno della Carta vengono enunciati cinque principi:

- 1) Il principio del rispetto dei diritti fondamentali;
- 2) Il principio di non-discriminazione;
- 3) Il principio di qualità e sicurezza;
- 4) Il principio di trasparenza, imparzialità e equità;
- 5) Il principio del controllo da parte dell'utilizzatore.

In virtù del primo principio, il trattamento di decisioni e dati giudiziari deve avere finalità chiare, che rispettino pienamente i diritti fondamentali garantiti dalla Convenzione europea sui diritti dell'uomo (CEDU) e dalla Convenzione sulla protezione delle persone rispetto al trattamento automatizzato di dati di carattere personale. Quando utilizzati in ambito giudiziario, tali strumenti non devono ledere il diritto a un equo processo, a un giudice terzo e imparziale ed il diritto di accesso al giudice.

Il principio di non discriminazione è finalizzato a prevenire "lo sviluppo o l'intensificazione di qualsiasi discriminazione tra persone o gruppi di persone", assicurando metodologie che non riproducano o aggravino dinamiche discriminatorie tra i gruppi sociali che potrebbero verificarsi in virtù della raccolta e classificazione di dati (e che possono comprendere origine etnica, fede religiosa, convinzioni politiche, condizioni socio-economiche ecc.), con particolare riferimento ai *Risk assessment tools* impiegati soprattutto in USA e UK. I soggetti potrebbero essere vittime di *implicit bias* sia nel caso in cui l'input non sia completamente neutro, sia nel caso in cui l'algoritmo riproduca preconcetti sociali (Quattrocchio, 2018).

Il principio di qualità e sicurezza dei dati da utilizzare raccomanda "l'uso di fonti certificate e dati intangibili, attraverso modelli concepiti in modo multidisciplinare, in un ambiente tecnologico sicuro". È necessario impiegare fonti certificate e garantire la completezza e integrità dei dati impiegati, predisponendo modelli funzionali di *machine learning* multidisciplinari che integrino competenze di professionisti del settore (giudici, pubblici ministeri, avvocati, docenti nei campi del diritto). Di fatto, la scelta dei dati da impiegare implica un'attenta analisi dell'integrità della fonte e del dato stesso per evitare che siano modificati. L'intero processo deve pertanto essere tracciabile, al fine di garantire che non abbia avuto luogo alcuna modifica in grado di alterare il contenuto o il significato della decisione trattata. Per gli stessi motivi, i modelli e gli algoritmi su cui si fonda l'elaborazione devono essere custoditi in ambienti sicuri.

Il principio di trasparenza, imparzialità e equità raccomandano "l'accessibilità, la comprensibilità e la verificabilità esterna dei processi computazionali" fissando il *right to explanation* del risultato decisionale (inteso come comprensibilità dello stesso) e il diritto di accesso al meccanismo di funzionamento interno dell'algoritmo, spesso ostacolato dal segreto industriale, coinvolgendo esperti esterni nella certificazione di tali qualità; dovrebbero essere dunque resi noti quantomeno informazioni parziali fondamentali in materia di algoritmi, per esempio quali siano le variabili utilizzate, quali siano gli obiettivi cui è finalizzata l'ottimizzazione degli algoritmi, i dati di apprendimento, i valori medi e gli scarti tipo dei risultati ottenuti, o la quantità e il tipo di dati trattati dall'algoritmo.

Il quinto principio prevede un uso degli algoritmi "garantendo che gli utilizzatori agiscano come soggetti informati, nel pieno controllo delle loro scelte". L'utilizzatore del sistema deve avere la possibilità di rivedere in qualsiasi momento la decisione giudiziaria e i dati utilizzati per ottenere il risultato senza essere vincolati ad esso.

Nell'Appendice II della carta si esaminano gli utilizzi dell'IA nei sistemi giudiziari europei distinguendo tra *utilizzi che devono essere incoraggiati* (comprendenti lo sviluppo di motori di ricerca giurisprudenziali avanzati, tecniche di apprendimento automatico per migliorare le banche dati esistenti, chatbot per facilitare l'accesso alle informazioni esistenti), *usi possibili che esigono notevoli precauzioni metodologiche* (ad esempio per quanto riguarda il supporto a misure alternative di risoluzione delle controversie in materia civile, l'impiego di strumenti di polizia predittivi per individuare luoghi in cui saranno commessi i reati, piattaforme di risoluzione online di *small claims*), *utilizzi da esaminare al termine di supplementari studi scientifici* (per la profilazione dei magistrati e l'anticipazione delle decisioni dei tribunali) e *utilizzi da esaminare con le più estreme riserve*.

In particolare, quest'ultima categoria si riferisce all'utilizzo di strumenti di valutazione del rischio in materia penale, come l'algoritmo COMPAS negli Stati Uniti di cui si parlerà successivamente, che sfruttando un approccio meramente statistico conducono spesso a risultati discriminatori o errati. La Carta deve essere intesa come "uno spazio in cui sono stati tracciati dei confini": saranno poi gli attori a delineare politiche di accountability, strategie, meccanismi di funzionamento e tutele adeguate, attraverso un sistema di *checks and balances* (Galoppa, 2021, pp. 34-37).

Circa la realtà italiana, l'ipotesi che un algoritmo possa sostituire o anche semplicemente affiancare il giudice nella valutazione del rischio di un reo è preclusa da una serie di "paletti" costituzionali: oltre ai limiti contenuti negli artt. 25 («nessuno può essere distolto dal giudice naturale precostituito per legge») e 102 («la funzione giurisdizionale è esercitata da magistrati ordinari istituiti e regolati dalle norme sull'ordinamento giudiziario»), l'art. 101, comma 1, Cost., nel disporre che i giudici sono soggetti soltanto alla legge, sostanzialmente, esclude che il giudice possa essere vincolato dall'esito di procedure algoritmiche (Donati, 2020, p.428). Inoltre, un limite fondamentale all'introduzione degli algoritmi è garantito dall'art. 111, comma 4, Cost. che garantisce il contraddittorio nella formazione della prova, impedendo al giudice di acquisire o di valutare elementi diversi da quelli oggetto di contraddittorio tra le parti.

Anche dal punto di vista della pericolosità sociale possiamo riscontrare degli argini; infatti, se ai sensi dell'art. 203, co.1 c.p., la pericolosità sociale può essere definita come la *probabilità* che il soggetto in futuro commetta nuovi reati e il suo accertamento è condizione per l'erogazione di misure di sicurezza personale, a seguito degli interventi della Corte Costituzionale (Corte Costituzionale 27 luglio 1982, n.139; Corte Costituzionale 28 luglio 1982, n.249), non sono più ammissibili presunzioni di pericolosità: la pericolosità sociale deve essere sempre accertata *in concreto* dal giudice. Tale giudizio si articola in due momenti: quello dell'*analisi della personalità* del soggetto e della *prognosi criminale* conseguente. La qualità di «persona socialmente pericolosa» si desume dalle circostanze indicate nell'art. 133 c.p.: bisogna tener conto di:

- *Gravità del reato* desunta:

1. dalla natura, dalla specie, dai mezzi, dall'oggetto, dal tempo, dal luogo e da ogni altra modalità dell'azione;
2. dalla gravità del danno o del pericolo cagionato alla persona offesa dal reato;
3. dall'intensità del dolo o dal grado della colpa.

- *Capacità a delinquere* desunta:

1. dai motivi a delinquere e dal carattere del reo;
2. dai precedenti penali e giudiziari e, in genere, dalla condotta e dalla vita del reo, antecedenti al reato;
3. dalla condotta contemporanea o susseguente al reato;
4. dalle condizioni di vita individuale, familiare e sociale del reo.

Il giudizio di capacità a delinquere si proietta necessariamente nel futuro ai fini di una prognosi sulle probabilità di commissione di un nuovo reato o di un reato dello stesso tipo, **ma non ha valore esplicitamente predittivo**. La valutazione del carattere del reo richiede una complessa valutazione della sua personalità e delle caratteristiche *innate* del soggetto idonee a orientare i suoi comportamenti (es. capacità di autocontrollo e stabilità emotiva). Sono oggetto di valutazione anche le condizioni sociali e familiari del soggetto (emarginazione, disoccupazione, adesione a bande criminali ecc.). Proprio per la complessità dei fattori da valutare è stabilito che il giudice effettui l'esame autonomamente, sulla base di un ragionamento razionale e logico che tenga conto delle fragilità del reo, senza poter ricorrere a valutazioni tecnico-scientifiche. Di fatto l'algoritmo fornisce una maggiore certezza e oggettività rispetto alle valutazioni umane solo *apparentemente* (Manes, 2020).

Altro baluardo a difesa dell'ingresso degli algoritmi nel processo è l'art 220 c.p.p. proprio perché afferma l'inammissibilità di perizie per «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».

La perizia sarebbe finalizzata a delineare un profilo della personalità e del carattere del reo per individuare la pena o la misura di sicurezza più adatta al caso di specie. Il legislatore si è fortemente opposto al ricorso a tale tecnica "automatizzata" di analisi: la *ratio* del divieto si può rinvenire nella necessità di tutelare la libertà morale dell'imputato, giacché si delinerebbe il rischio di cedere ai pregiudizi inerenti ad aspetti particolari del carattere dell'imputato che potrebbero condizionare l'organo giudicante. Si vuole dunque evitare che il giudice, nell'assumere le sue determinazioni, si basi essenzialmente sull' *identità dell'imputato* tracciata dalle perizie psicologiche e non sui *fatti commessi* in concreto (Di Prisco, 2018).

Corollario logico di tale previsione è il divieto di servirsi di strumenti algoritmici di valutazione del rischio, sostanzialmente equiparabili a perizie criminologiche. La prognosi sulla capacità a delinquere è un giudizio *intuitu personae* e nessuna valutazione statistica può supportare o sostituire un giudizio di questo tipo, anche perché «il punteggio di rischio verrebbe in tal modo calcolato incrociando i dati relativi a situazioni simili o vicende analoghe, facendo cadere il giudizio sulla pericolosità sociale del reo in un labirinto di inevitabili generalizzazioni empiriche» (Di Prisco, 2018).

Fondando il giudizio sulla condotta posta in essere dal soggetto, viene scongiurato il rischio che la valutazione si fondi in esclusiva su condotte poste in essere, in passato, da soggetti che si siano trovati in situazione analoghe all'imputato in esame (Polidoro, 2020).

Ed è proprio quello che si affronterà nel paragrafo successivo con tutte le conseguenze discriminatorie del caso; inoltre bisognerà capire anche quello che effettivamente sta

succedendo in Italia a prescindere da talune garanzie che vengono aggirate in modo più o meno elegante.

3. La valutazione attuariale e gli algoritmi predittivi in materia di pericolosità.

Prima di fare alcune osservazioni in merito all'analisi degli "algoritmi predittivi" della pericolosità criminale, può risultare utile anteporre alcune considerazioni sulla valutazione c.d. attuariale di siffatta pericolosità, che costituisce il presupposto teorico per l'utilizzo degli algoritmi predittivi.

Negli ultimi anni si sta facendo sempre più strada una concezione *evidence-based* di valutazione del rischio individuale di commissione di un (nuovo) reato: una concezione, basata su riscontri oggettivi che potrebbe o può integrare le valutazioni intuitive dei giudici, tuttora ampiamente diffuse.

La valutazione *evidence-based* della pericolosità criminale presuppone la previa individuazione di una serie di *fattori di rischio* (o *predittori*) direttamente coinvolti nel comportamento criminoso, fattori che possono, tra l'altro, riguardare:

- l'età;
- il sesso;
- l'origine etnica;
- il livello di scolarizzazione;
- la situazione familiare e lavorativa;
- la posizione sociale;
- i precedenti penali;
- le precedenti esperienze carcerarie;
- i luoghi e le persone frequentati;
- la presenza di autori di reato nella cerchia familiare o nella rete di conoscenze;
- il luogo di residenza;
- le difficoltà di regolazione della rabbia e aggressività;
- il discontrollo degli impulsi;
- una storia di ospedalizzazione;
- un pensiero pro-criminale;
- alcune variabili contestuali (quali, ad esempio, la mancanza di sostegno familiare e sociale);
- il consumo di sostanze stupefacenti o alcoliche;
- le psicopatie.

Ebbene, tutti questi fattori, possono consentire un *approccio di tipo attuariale (o statistico)* alla valutazione della pericolosità criminale. Attraverso, infatti, una loro combinazione, si possono predisporre delle 'scale' che consentono l'attribuzione di un punteggio (*score*) al soggetto preso in esame (Castelletti, Rivellini, Straticò, 2014, pp. 153 ss.)

Peraltro, le "scale", utilizzate per la valutazione attuariale della pericolosità criminale, si differenziano tra di loro, in base, tra l'altro, alla *popolazione* (ad esempio, popolazione di adulti, di minori, di maschi, di femmine, di pazienti psichiatrici, di detenuti o ex-detenuti); in base alla *tipologia di reati implicati* (scale relative a tutti i reati, scale relative a singole tipologie di reati, come i reati sessuali o i reati violenti); in base alla *temporalizzazione del rischio* (immediato, o a medio o a lungo termine); in base, infine, al *contesto applicativo*

(comunità civile, istituti di pena, centri di salute mentale, ospedali psichiatrico-giudiziari) (Zara, 2016, pp. 17 ss.).

Inoltre non tutti i fattori di rischio sono uguali e non tutti impattano in modo univoco e nello stesso modo. I fattori di rischio, infatti, hanno un differente tasso di *dinamicità*, nel senso che esistono: a) fattori statici, non modificabili (ad esempio, il sesso e l'origine etnica); b) fattori dinamici stabili, che sono modificabili grazie al trattamento terapeutico (ad esempio, il discontrollo degli impulsi); c) infine, fattori di rischio acuti, che cambiano rapidamente e sono associati ad una condizione facilitante la reazione violenta (ad esempio, l'uso di sostanze stupefacenti).

A questo punto pensiamo alla possibilità che queste valutazioni attuariali – e prima ancora la raccolta e rielaborazione dei dati che consentono la predisposizione delle “scale” – possono essere gestite, come di fatto già avviene negli Stati Uniti, da sistemi di intelligenza artificiale, quindi ad algoritmi predittivi, forniti di procedure di autoapprendimento (*machine learning*) e straordinaria capacità e rapidità nel far emergere *relazioni, coincidenze, correlazioni*; non sarà difficile immaginare che ai pochi vantaggi in termini di velocità dell'informazione si contrappongono i grandi rischi in termini di discriminazione. E proprio perché a volte l'immaginazione conta poco, è necessario dettagliare quello che succede in alcuni paesi “democratici”.

Negli Stati Uniti, già da oltre venti anni sono in fase di diffusione algoritmi predittivi della pericolosità criminale. Essi sono, ad esempio, usati nella fase del “*parole*” (per decidere se un individuo, nelle more della celebrazione del processo, possa essere rilasciato dietro il pagamento di una eventuale cauzione), o per misurare il rischio di recidiva del condannato, ai fini della sua ammissibilità al “*probation*” o ad altra misura alternativa alla detenzione.

Tuttavia l'algoritmo predittivo di gran lunga più famoso, utilizzato (e controverso) negli Stati Uniti è COMPAS - *Correctional Offender Management Profiling for Alternative Sanctions*, un *software* elaborato e commercializzato da una società privata, la Northpointe (da gennaio 2017, ridenominata Equivant). Solo per inciso, da quando è stato sviluppato nel 1998, COMPAS è stato utilizzato in più di un milione di casi.

Per avere un'idea di come funzioni COMPAS è possibile sfogliare *online* una versione del relativo “Manuale operativo”, risalente a marzo 2015, da cui tra l'altro si apprende che:

«COMPAS è uno strumento di valutazione dei rischi (di commissione di reato) e delle esigenze (trattamentali) di quarta generazione. Le agenzie di giustizia penale in tutto il Paese utilizzano COMPAS per assumere decisioni in merito al collocamento, alla supervisione e alla gestione degli autori di reati. COMPAS è stato sviluppato empiricamente con un focus sui predittori di cui è nota l'influenza sulla recidiva. Esso prende in considerazione anche fattori di rischio dinamici e fornisce informazioni su una varietà di fattori di rischio ampiamente convalidati (dalla ricerca scientifica) al fine di agevolare gli interventi correttivi rivolti a ridurre le probabilità di recidiva [...]. COMPAS è stato sviluppato per la prima volta nel 1998 e da allora è stato rivisto più volte man mano che la base di conoscenze fornite dalla criminologia e dalla prassi correzionale si è evoluta [...]. Continuiamo a apportare miglioramenti a COMPAS sulla base dei risultati di ricerche empiriche e di studi sulla recidiva, condotti in carcere o presso agenzie preposte a seguire la “*probation*”. COMPAS viene periodicamente aggiornato per stare al passo con le migliori pratiche emergenti e i progressi tecnologici [...].

COMPAS prende in considerazione – nella sua configurazione base – la risposta a 137 domande, concernenti le seguenti voci:

- precedenti criminali;
- precedenti illeciti e infrazioni;
- passato di violenza;
- violenza attuale;
- frequentazioni con criminali;
- abuso di sostanze;
- problemi economici;
- difficoltà nell'istruzione e nella formazione professionale;
- ambiente familiare delinquenziale;
- contesto sociale;
- modo di utilizzo del tempo libero;
- instabilità residenziale;
- adeguamento sociale;
- difetti di socializzazione;
- opportunità criminali;
- isolamento sociale;
- pensiero pro-criminale;
- personalità criminale.

[...] Le risposte alle domande vengono fornite dal soggetto sotto valutazione, oppure vengono ricercate negli archivi o nei registri a disposizione delle procure e della polizia. La razza non è un fattore preso in considerazione. In sede di intervista con l'imputato gli vengono rivolte domande del seguente tipo: "Uno dei tuoi genitori è mai stato in prigione o è attualmente in prigione?". Il questionario chiede inoltre alle persone di essere d'accordo o in disaccordo con affermazioni del seguente tipo: "Una persona affamata ha il diritto di rubare" e "Se le persone mi fanno arrabbiare o perdere la calma, posso essere pericoloso".

[...] COMPAS si distingue da altri *software* di calcolo attuariale, in quanto tiene in considerazione anche i fattori di rischio dinamici, oltre a quelli statici, e in quanto fornisce indicazioni non solo sul rischio di recidiva, ma anche sul trattamento più adatto per la singola persona per ridurre tale rischio».

Nei confronti dell'impiego di COMPAS sono state tuttavia sollevate – sulla scorta di ricerche indipendenti – alcune critiche in ordine alla sua effettiva validità predittiva (*accuracy*) e alla sua imparzialità (*fairness*).

In particolare, nel maggio del 2016 un gruppo di ricercatori ha pubblicato, su incarico di una Organizzazione Non Governativa (ProPublica), una ricerca in cui si analizzavano le prestazioni di COMPAS su un campione di oltre 7000 persone arrestate nella contea di Broward, in Florida, tra il 2013 e il 2014. Da questa ricerca emergerebbe che le previsioni formulate da COMPAS erano inaffidabili e risentivano di distorsioni su base razziale, favorendo gli imputati bianchi rispetto agli imputati neri, dal momento che si rilevava una sottostima del rischio di recidiva dei primi e una sovrastima del rischio di recidiva dei secondi (Angwin, Larson, Mattu, Kirchner, 2016).

Anche uno studio successivo ha sollevato gravi perplessità, evidenziando due profili problematici ulteriori rispetto al rischio di pregiudizio razziale:

- l'algoritmo COMPAS fornirebbe previsioni sostanzialmente equivalenti a quelle fornite da persone prive di conoscenze specifiche in materia: quindi, la sua utilità sarebbe estremamente discutibile;
- il grado di affidabilità delle valutazioni, fornite da COMPAS dopo aver preso in considerazione ben 137 voci, **sarebbe sostanzialmente equivalente al grado di affidabilità di valutazioni fornite sulla scorta di solo 2 voci.**

A ciò basta aggiungere che, al momento, non mancano ulteriori e più approfonditi studi di verifica delle prestazioni di COMPAS, anche a causa del fatto che molte delle informazioni utili per effettuare tali studi sono, in realtà, coperte da segreto industriale, ben custodito dalla società privata che commercializza il relativo *software*.

Critiche ancor maggiori nei confronti di COMPAS hanno riguardato il suo possibile utilizzo in sede di “*sentencing*”, vale a dire a fini di commisurazione della pena dell'imputato riconosciuto colpevole.

Tali critiche sono in particolare emerse in relazione al c.d. caso Loomis, dal nome di un imputato che aveva fatto ricorso alla Corte Suprema del Wisconsin per contestare l'entità della pena che gli era stata inflitta dalla Corte locale che, in fase della commisurazione, si era per l'appunto avvalsa di COMPAS: l'algoritmo predittivo veniva contestato dal ricorrente per la sua predisposizione a seguire pregiudizi basati sul genere e sulla razza, nonché per il difetto di trasparenza relativo al suo meccanismo di funzionamento.

La Corte Suprema del Wisconsin, sollecitata da tali rilievi, ha formulato un *warning* in relazione al futuro uso di COMPAS, mettendo in evidenza:

- la sua natura di prodotto coperto da segreto industriale, che impedisce la divulgazione di informazioni relative al suo metodo di funzionamento;
- il fatto che le valutazioni sono effettuate da COMPAS su base collettiva, di gruppo, e non individuale;
- infine, il rischio di una sovrastima del rischio di commissione di reati a carico di talune minoranze etniche.

Ciò nondimeno, nel caso di specie, la Corte ha respinto il ricorso del sig. Loomis, sulla scorta della considerazione che le valutazioni di COMPAS non erano state “decisive”, in quanto erano state pur sempre sottoposte al controllo e alla validazione di un giudice umano.

Per ora, ma solo per ora, quanto meno in Europa, gli algoritmi predittivi della pericolosità e, più in generale, gli *automated decision systems*, non hanno avuto accesso nelle nostre aule penali, anche perché, a precludere loro l'accesso, si erge l'art. 15 della direttiva 95/46/CE, confluito nell'art. 22 del nuovo Regolamento europeo in materia di protezione dei dati personali, entrato in vigore il 25 maggio 2018. Tale articolo stabilisce, infatti, che ogni persona ha il diritto di non essere sottoposta ad una decisione che produca effetti giuridici o abbia effetti significativi nei suoi confronti fondata esclusivamente su *un trattamento automatizzato di dati* destinati a valutare taluni aspetti della sua personalità (Basile, 2019, pp. 17-22).

Come si è avuto modo di evidenziare, in un paese come gli Stati Uniti d'America, culla della democrazia, si sono imposti algoritmi predittivi i cui risultati non sono altro che l'evoluzione matematica di discriminazioni che sono state combattute a viso aperto da milioni di persone e per tanti decenni. È vero; si tratta di sistemi in base ai quali la giustizia funziona in modo molto diverso, ma saremo veramente in grado di lasciare fuori dalle aule

di tribunale quelle armi di distruzione matematica, come le definisce Cathy O'Neil, e non lasciarci ammaliare dal fascino degli algoritmi?

4. Prime applicazioni in Italia per gli autori di reati sessuali.

Nonostante i limiti anzidetti, l'idea di utilizzare gli algoritmi perlomeno in termini di valutazione del rischio di recidiva con sistemi basati sul Risk-Need-Responsivity (trattamento proporzionale al rischio di commettere un nuovo reato), si incomincia a prospettare anche nell'ordinamento italiano; ad esempio, strumenti di valutazione del rischio di recidiva come STATIC 99R, STABLE 2007 sono utilizzati in un programma condotto dal CIPM (Centro Italiano per la Promozione della Mediazione) nel carcere di Milano Bollate e presso il Presidio criminologico del Comune di Milano dal 2005, per la scelta del tipo di trattamento nei confronti degli autori di reati sessualmente connotati (Xella, 2013, p. 1887 ss.). O ancora in una sentenza del Tribunale di Milano si fa riferimento all'uso di strumenti di radice nordamericana per la valutazione del rischio di recidiva, funzionale alla valutazione della pericolosità sociale ai fini dell'applicazione di una misura di sicurezza nel settore dei reati sessuali: *“Tali elementi sono condivisi dall'uomo in sede di colloqui individuali oppure durante l'indagine del rischio di recidiva, realizzato con strumenti di matrice nord americana, che indicano un livello globale “medio-alto” (circa il 25% di possibilità di commettere un nuovo reato sessuale nei 5 anni dopo la detenzione)”* (Tribunale di Milano, 2016).

Quanto sia forte la voglia o la necessità di strumenti predittivi, ma al tempo stesso i pericoli che ne possono derivare, si rinviene da un articolo di Carla Xella (psicologa, psicodiagnosta e psicoterapeuta) in Rassegna italiana di criminologia, n. 3 del 2020. Nell'abstract si legge *“I trattamenti per la prevenzione della recidiva negli autori di reati sessuali sono una realtà in molti Paesi, mentre in Italia non abbiamo che iniziative isolate e mal finanziate. La risposta del sistema giudiziario è prevalentemente retributiva e la valutazione del rischio di recidiva viene effettuata solo sulla base del giudizio clinico, notoriamente poco predittivo. Poiché i sex offenders non presentano tutti lo stesso rischio di ripetere il reato, in molti Paesi si è cercato di affinare la valutazione del rischio di recidiva usando strumenti validi e predittivi, anche allo scopo di effettuare un trattamento mirato ai fattori criminogeni”*.

Vediamo, tuttavia, di approfondire il funzionamento di tali strumenti predittivi e per farlo prendiamo in esame i fattori considerati dalla Static-99R che sono molto simili ad altri strumenti:

- Convivenze (costituiscono un fattore protettivo);
- Tendenza ad esercitare violenza anche non sessuale: la presenza di reati di violenza non sessuale nell'attuale condanna o reati violenti precedenti;
- Presenza nel fascicolo personale di reati sessuali precedenti e presenza di reati sessuali senza contatto;
- Numero di condanne riportate per qualsiasi reato;
- Tipo di vittime: l'aver fatto vittime extrafamiliari, o sconosciute, o di sesso maschile aumenta il rischio di recidiva.

I dati necessari alla compilazione della scala vengono desunti principalmente dal fascicolo giudiziario del soggetto. È assolutamente necessario che l'operatore si limiti ai dati acquisiti, e non “arricchisca” il responso della scala con sue valutazioni personali, che ne diminuiscono l'attendibilità.

La Static produce 5 categorie di rischio: molto basso, sotto la media, medio, superiore alla media e molto superiore alla media. Data la disparità dei risultati delle diverse ricerche, e dato anche il fatto che il tasso di recidiva varia da un Paese all'altro e anche nel corso del tempo, nell'ultima revisione della scala (2016) si è scelto di optare per una definizione di rischio *relativo*, vale a dire misurato in base allo scostamento dal tasso *medio*, che viene definito sulla base dei tassi effettivi di recidiva dei reati sessuali, aggiornati periodicamente.

Anche se la Static-99R possiede una buona predittività fra i diversi strumenti utilizzati, ha il difetto di essere statica, cioè di non dare alcuna indicazione clinica su come maneggiare il rischio: essendo basata sulla storia passata del soggetto, non si può che prendere atto del suo maggiore o minore rischio di recidiva al momento in cui viene rimesso in libertà.

Nei sistemi giudiziari che prevedono un intervento di reinserimento nella società e di prevenzione della recidiva per gli autori di reato sessuale (*risk management*), è stato quindi indispensabile dotarsi di strumenti diversi: strumenti che consentano di identificare fattori sui quali sia possibile intervenire.

Per questo motivo la nuova generazione degli strumenti di valutazione del rischio è stata quella dei predittori *dinamici*: si intendono con questo termine caratteristiche di personalità relativamente stabili, ma suscettibili di trattamento e quindi di cambiamento. In base al principio RNR (trattamento proporzionale al rischio di commettere un nuovo reato), sono state costruite le scale STABLE-2000 e STABLE-2007. Si tratta di una intervista semi-strutturata che mira a definire un profilo personale di rischio per ciascun soggetto.

I 13 item della scala sono raggruppati in 5 cluster:

- Influenze sociali significative;
- Deficit di intimità (capacità di relazioni stabili, identificazione emotiva con i bambini, ostilità verso le donne, rifiuto sociale/solitudine, mancanza di interesse per gli altri)
- Autocontrollo generale (impulsività, scarse capacità di *problem solving*, emozioni negative/ostilità);
- Autocontrollo sessuale (ipersessualità, interessi sessuali devianti);
- Collaborazione con la supervisione.

Una corretta siglatura della STABLE 2007 richiede che vengano raccolte in modo adeguato le informazioni necessarie. Lo strumento è quindi corredato sia di un manuale, che viene costantemente aggiornato sulla base dei nuovi dati di ricerca, sia di una guida all'intervista strutturata.

Come per la Static, è fondamentale che l'intervista con il soggetto non sia l'unica fonte di informazione a disposizione dell'operatore: per alcuni item (ad esempio il rapporto con le donne, o l'impulsività, o la capacità di *problem solving*) è consigliabile servirsi anche di altre fonti: operatori, familiari, dati di archivio.

La Static-99R e la STABLE (anche se viene detto che sono utilizzate in maniera sperimentale), vista la loro diffusione e i risultati in termini di predittività verificati in molti paesi del mondo, un gruppo di clinici e ricercatori del C.I.P.M. (Centro Italiano per la Promozione della Mediazione), in collaborazione con il Dipartimento dell'Amministrazione Penitenziaria e con il Dipartimento di Giustizia Minorile e di Comunità, ha intrapreso un progetto di traduzione e validazione dello strumento. Il C.I.P.M. gestisce dal 2005 una Unità di Trattamento Intensificato per gli autori di reati sessuali e un Presidio

Criminologico convenzionato con il Comune di Milano, che costituisce idealmente una prosecuzione del trattamento intramurario e un progetto di presa in carico clinico-criminologica. Sono attualmente operanti varie sedi decentrate a Roma, Firenze, Cagliari, Piacenza, Pavia, Genova. Gli operatori del C.I.P.M. hanno effettuato una formazione a distanza sulla siglatura di entrambi gli strumenti e a breve potranno formare altri operatori. Nel progetto è previsto che il DAP possa fornire anche i dati sulla recidiva a tutt'oggi mancanti (insomma in termini di predittività in Italia, siamo molto più avanti di quanto si possa pensare).

Gli strumenti appena descritti, consentono, sostanzialmente, di tracciare, per ogni soggetto, un personale profilo di rischio. Non tutti i fattori, però, sui quali è costruito lo strumento predittivo sono presenti in egual misura in ciascun soggetto: per alcuni, la presenza di interessi sessuali devianti sarà il principale fattore di rischio, per altri lo sarà l'isolamento sociale, per altri ancora l'impulsività generalizzata. Sia nel trattamento individuale che in quello di gruppo, questi saranno gli obiettivi sui quali indirizzare il lavoro.

Il trattamento "taglia unica" si è dimostrato infatti non solo poco utile, ma addirittura controproducente. Oggi i programmi sono differenziati in base ai diversi livelli di rischio dei partecipanti e tengono conto in modo assai più deciso dei fattori protettivi presenti nella vita di ognuno. Gli strumenti di valutazione del rischio sono quindi un utile strumento a disposizione di chi si occupa del trattamento di questo tipo di rei.

I rischi di un loro uso improprio sono però sempre presenti. Innanzi tutto, per applicare le scale di valutazione è necessaria una formazione specifica, effettuata da trainer riconosciuti con vasta esperienza nel campo e, almeno per i primi tempi, è opportuna una supervisione periodica. È necessario poi avere *tempo*, sia per raccogliere i dati dal fascicolo giudiziario, sia per effettuare l'intervista semi-strutturata.

Soprattutto, è necessario utilizzare i dati così raccolti in maniera corretta, per lo scopo che ha guidato gli autori delle scale e che ispira anche la nostra legislazione in materia di esecuzione penale (L.354/75): la pianificazione di un trattamento individualizzato, misurato sui reali fattori criminogeni dell'autore di reato.

Tuttavia queste condizioni sono difficilmente realizzabili in Italia.

Il primo rischio è che la valutazione, una volta che si riesca a farla, finisca per essere null'altro che un ulteriore stigma per l'autore di reato (Xella, 2020, pp. 195-201).

Ed ecco che si ritorna alle origini; qualsiasi strumento predittivo che fonda le radici nel passato, e non può fare altrimenti, non produce predizioni, ma solo pregiudizi e discriminazioni.

5. Conclusioni

Alla luce di quanto detto, non si può escludere, in modo categorico, che i meccanismi predittivi di cui all'intelligenza artificiale non troveranno una qualche applicazione anche, indirettamente, nell'ambito del settore penale. In effetti nell'ambito dei reati sessuali già ci siamo portati avanti, almeno per quanto concerne il pericolo di recidiva.

Se poi pensiamo al settore investigativo, in Italia trovano ampia applicazione di algoritmi come KeyCrime e XLAW i quali, se da un lato, mettono a confronto i crimini in una determinata area, dall'altro, ne prevedono l'insorgenza sul territorio.

Tuttavia onde evitare dicotomie, polarizzazioni o estremismi connessi ad una tecnologia che se non si può fermare è meglio governare, è necessario, comunque, rifuggire dall'idea di contrapporre il digitale all'umano. Sarebbe un continuum oltremodo disastroso; semmai i principi etici dovrebbero solo guidarci a riconoscere alle tecniche di intelligenza artificiale la sola funzione ausiliare e complementare alle attività umane, potendone fare a meno qualora la benché minima discriminazione dovesse essere a portata di mano o, meglio, a portata di bit.

Bibliografia

Aletras Nikolaos, Tsarapatsanis Dimitrios, Preotiuc-Pietro Daniel and Lampos Vasileios: "Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective", *PeerJ Computer Science* (<https://peerj.com/articles/cs-93>), 2016.

Angwin Julia, Larson Jeff, Mattu Surya, Kirchner Lauren: Machine Bias, in www.propublica.org, 23 maggio 2016.

Basile Fabio: Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine, 2019.

Bourcier Danièle: L'acte de juger est-il modélisable? De la logique à la justice, in *Archives de philosophie du droit*, pp. 37-53.

Caratta Antonio: Decisione robotica e valori del processo, in *Rivista di diritto processuale*, pp. 491-514, 2020.

Castelletti Luca, Rivellini Giulia, Straticò Ettore: Efficacia predittiva degli strumenti di Violence Risk Assessment e possibili ambiti applicativi nella psichiatria forense e generale italiana, in *Journal of Psychopathology*, 2014.

Di Prisco Alessia: Elementi di criticità sulla perizia psicologica nel processo penale, in *Ius in Itinere*, 2018.

Donati Filippo: Intelligenza artificiale e giustizia in *Rivista Associazione italiana dei Costituzionalisti*, n.1/2020.

Ferrua Paolo: La prova nel processo penale, *Struttura e procedimento*, 2° edizione, Torino, Giappichelli, 2017.

Fiandaca Giovanni: Crisi della riserva di legge e disegno della democrazia rappresentativa nell'età del protagonismo giurisdizionale, in *Criminalia*, pp. 79-98, 2011.

Galloppa Elisa: L'uso degli algoritmi nel processo penale e la sua conformità ai principi dell'equo processo, *LUISS Guido Carli*, 2021.

Garapon Antoine et Lassègue Jean: La giustizia digitale. Determinismo tecnologico e libertà (edizione italiana di Maria Rosaria Ferrarese, Bologna, Il Mulino, 2021).

Kostoris Roberto: Predizione decisoria, diversion processuale e archiviazione, in *Sistema Penale*, pp. 5-9, 2021.

Leopardi Giacomo: *Canti – Operette Morali*, Milano, Fratelli Fabbri Editori, 1968.

Lettieri Nicola: Contro la previsione. Tre argomenti per una critica del calcolo predittivo e del suo uso in ambito giuridico, in *Ars Interpretandi*, pp. 83-96, 2021.

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

Polidoro Donato: Tecnologie informatiche e procedimento penale: la giustizia penale "messa alla prova" dall'intelligenza artificiale in *Archivio Penale*, 2020, n. 3.

Quattrocolo Serena: Intelligenza artificiale e giustizia: nella cornice della Carta Etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche in *La Legislazione Penale*, 2018.

Quattrocolo Serena: *Artificial Intelligence, Computational Modelling and Criminal Proceedings. A Framework for a European Legal Discussion*, Berlino, Springer, 2020.

Quattrocolo Serena: Quesiti nuovi e soluzioni antiche? Consolidati paradigmi normativi vs rischi e paure della giustizia digitale "predittiva" in *Cassazione penale* n. 4/2019.

Rossi Francesca: *Il confine del futuro. Possiamo fidarci dell'intelligenza artificiale?* Milano, Feltrinelli, 2019.

Searle John: *Minds, Brains and Programs*, in *Behavioral and Brain Sciences*, 1980.

Tribunale di Milano: 19 aprile 2016, in www.dejure.it.

Xella Carla Maria: La valutazione del rischio di recidiva per gli autori di reati sessuali, in *Riv. it. med. leg. e dir. In campo sanitario* 2013.

Xella Carla Maria: Sex Offender: valutazione del rischio e rischi della valutazione in *Rassegna Italiana di Criminologia*, n.3/2020.

Zara Georgia: Tra il probabile e il certo. La valutazione dei rischi di violenza e di recidiva criminale in *Diritto penale contemporaneo*, 2016.