

# — ***Risk assessment: una nuova sfida per la giustizia penale?***

La pericolosità criminale al vaglio algoritmico delle probabilità nell'esperienza statunitense

***Risk assessment: a new challenge for Criminal Justice?***

*Future criminal behavior under algorithmic examination of the likelihood in the American experience*

*di Diana Zingales*

---

**Abstract.** Perché il legislatore dovrebbe affidare ad un sistema algoritmico la valutazione della pericolosità di un imputato? Quali sono le prospettive di imparzialità e precisione che un output ci promette nel "prevedere" il rischio di recidiva? E quali limiti potrà porre il diritto penale all'evoluzione tecnologica, rappresentata dai risk assessment tools basati sull'Artificial Intelligence? I risultati prodotti dal ricorso a tali strumenti nel sistema penale statunitense ci offrono le prime risposte a questi interrogativi, delle quali converrà tener conto se e quando ci troveremo, anche in Italia, a confrontarci con questi "sistemi predittivi", tanto innovativi quanto pericolosamente confliggenti con i diritti costituzionali a tutela dell'imputato.

**Abstract.** Why should the legislator reentrust to an algorithmic system the assessment of the dangerousness of the defendant? What are the perspectives of neutrality and accuracy that an output promises in "predicting" the risk of recidivism? What limits will Criminal Law establish to the technological evolution represented by AI-based risk assessment tools? The outcomes produced by the use of these tools in the American criminal law system offer the fundamental answers to such questions, whose will be convenient to be aware of, if and when, in Italy as well, we will manage these "predictive systems" as innovative as opposing to the constitutional principles which protect the defendant.

SOMMARIO: 1. Il *criminal risk assessment* della modernità: algoritmi e predizione del rischio. – 2. La California sceglie di limitare il ricorso ai sistemi algoritmici di valutazione del rischio: l'eloquente esito della *Proposition 25* sul *replacement* del *cash bail* con i *pretrial risk assessment tools*. – 3. Il *First Step Act* e il nuovo *recidivism risk assessment system* del *Department of Justice*: PATTERN. – 4. Le insidie nascoste dei *risk assessment tools* sotto la lente dei principi costituzionali del giusto processo: l'antesignana prospettiva statunitense. – 5. Gli orientamenti della giurisprudenza statunitense e le soluzioni prospettate dalla dottrina per un uso "costituzionalmente orientato" degli *AI-based risk assessment tools*.

SUMMARY: 1. Criminal risk assessment in the modern era: algorithms and risk prediction. – 2. California chooses to limit the use of algorithmic risk assessment tools: the eloquent outcome of Proposition 25 on the cash bail replacement by pretrial risk assessment tools. – 3. The First Step Act and the new recidivism risk assessment system of the Department of Justice: PATTERN. – 4. The hidden dangers of risk assessment tools in the light of the Due process clause. – 5. The trends of the American jurisprudence and the solutions proposed by scholars to use AI-based risk assessment tools in compliance with the Due Process Clause.

## 1. Il *criminal risk assessment* della modernità: algoritmi e predizione del rischio.

Anelare ad una sentenza "giusta" o ad una valutazione prognostica "esatta" della pericolosità criminale sembra costituire non solo l'auspicio di chi si trovi coinvolto in un procedimento penale nella qualità di imputato, ma l'ambizioso obiettivo dell'essere umano che protende incessantemente verso il superamento di confini che vengono spostati sempre più in là.

I progressi, fino a qualche decennio fa inimmaginabili, che le scienze informatiche hanno raggiunto in vari ambiti si manifestano in misura sempre più prepotente anche nel campo del diritto, grazie allo sviluppo di *software* in grado di consegnare "decisioni algoritmiche" in sostituzione di quelle del giudice umano. Gli algoritmi di oggi possono persino conferire maggiore certezza al diritto e garantire la prevedibilità delle decisioni, come accade nel sistema penale cinese, dove i sistemi algoritmici più evoluti generano un *warning* che segnala la difformità della decisione giurisdizionale adottata rispetto all'orientamento giurisprudenziale consolidato – limitando in tal modo la discrezionalità del giudice – con l'obiettivo finale di assicurare la "ragionevolezza" della sentenza e l'uniformità delle pronunce<sup>1</sup>.

Un'area che è risultata particolarmente florida per lo sfruttamento delle potenzialità degli algoritmi è quella del *risk assessment*<sup>2</sup>, ovvero della valutazione prognostica del rischio-ricidiva di un imputato e della sua pericolosità sociale<sup>3</sup>: ad oggi

---

<sup>1</sup> T. Sourdin, *Judges, Technology and Artificial Intelligence. The Artificial Judge*, Elgar, 2021, pp. 13-14; M. Yu, G. Du, *Why Are Chinese Courts Turning to AI?*, in *The Diplomat*, 19 gennaio 2019.

<sup>2</sup> Nella dottrina penale statunitense si evidenzia che il "risk assessment" implica l'utilizzo di sistemi attuariali e algoritmici per elaborare previsioni sulla probabilità di futuri reati. Sul punto, tra gli altri, M. Stevenson, *Assessing Risk Assessment in Action*, *Minnesota Law Review*, vol. 103, No. 58, 2018, p. 314: «The term "risk assessment," however, usually refers to the use of formal, actuarial, and algorithmic methods of predicting the likelihood of future crime or misconduct».

<sup>3</sup> Per una disamina degli strumenti di *risk assessment* e del loro utilizzo nel sistema della giustizia penale statunitense si vedano B.L. Garrett, J. Monahan, *Judging Risk*, in *California Law Review*, vol. 108, 2020, p. 450.

risultano in uso su scala mondiale circa quattrocento *risk assessment tools*<sup>4</sup>, la cui evoluzione risulta scandita da quattro fasi temporali<sup>5</sup>.

Mentre in una prima fase il *risk assessment* si basava su procedure informali lontane da logiche di *scores* e di calcoli statistici operati su base collettiva e si incentrava su interviste al soggetto da valutare, nella seconda generazione di *tools*, risalente ai primi anni '70, le metodologie della valutazione diventavano di natura attuariale e si basavano su fattori "statici" (come l'età, il genere e i precedenti penali).

È tra la fine degli anni '70 e gli inizi degli anni '80, con l'introduzione dei *tools* di terza generazione caratterizzati dall'inserimento anche di fattori "dinamici" di rischio (come l'impiego e l'*educational status*), che gli strumenti statistici di *risk assessment* divengono "*RNA tools* (*Risk and Needs Assessment tools*), in quanto volti non solo ad individuare il rischio-reato, ma anche a diminuire il rischio di recidiva del soggetto e a favorire il suo reinserimento sociale<sup>6</sup>. I *risk assessment tools* contemporanei sono quelli di quarta generazione: si tratta di modelli RNA che presentano maggiore complessità perché dotati di un *dataset* con un più elevato numero di *items* da elaborare rispetto ai modelli precedenti, caratteristica che dovrebbe garantire una maggiore predittività. Il *risk assessment* di quarta generazione si caratterizza inoltre per la massimizzazione della prevenzione e gestione del rischio attraverso la predisposizione di specifici programmi di trattamento del soggetto sottoposto a valutazione<sup>7</sup>.

La quinta generazione di *risk assessment tools* è infine rappresentata dai sistemi algoritmici di *machine learning*<sup>8</sup>, caratterizzati da *software* programmati per eseguire certi compiti e che aggiornano i loro codici man mano che "imparano" dai risultati osservati<sup>9</sup>, anche se ad oggi l'esiguità del numero di informazioni disponibili – dovuta alla segretezza dei dati sul loro funzionamento – non permette di concludere nel senso di una loro più efficace predittività<sup>10</sup>. Secondo le definizioni della dottrina statunitense, solo questi ultimi sistemi rientrano nell'*Artificial Intelligence*, in quanto dotati della capacità di autoapprendimento<sup>11</sup>: se è vero infatti che parlare di *AI* significa necessariamente far

---

<sup>4</sup> N. Scurich, *An Introduction to the Assessment of Violence Risk*, in *International Perspectives on Violence Risk Assessment*, 3, 10, J.P. Singh, S. Bjørkly, & S. Fazel eds., 2016, 3, p. 10.

<sup>5</sup> Sulla progressiva evoluzione che ha contraddistinto i *risk assessment tools* succedutisi nel tempo si rinvia a B.L. Garrett, J. Monahan, *Judging Risk*, cit., p. 451; C.D. Stimson, *The First Step Act's Risk and Needs Assessment Program: A Work in Progress*, pp. 4 ss.

<sup>6</sup> Cfr. B.L. Garrett, J. Monahan, *Judging Risk*, cit., p. 451.

<sup>7</sup> William D. Burrell, *Risk and Needs Assessment in Probation and Parole: The Persistent Gap Between Promise and Practice*, in *Handbook on Risk and Need Assessment: Theory and Practice*, Faye S. Taxman, ed., 2017, p. 26.

<sup>8</sup> A.S. Niederman et al., *The Institutional Life of Algorithmic Risk Assessment*, in *Berkeley Technology Law Journal*, vol. 34, No. 3, 2019, pp. 711-712.

<sup>9</sup> A.B. Cyphert, *Reprogramming Recidivism: The First Step Act and Algorithmic Prediction of Risk*, in *Seton Hall Law Review*, vol. 51, 51, 2020, p. 339; *Id.*, *Tinker-ing with Machine Learning: The Legality and Consequences of Online Surveillance of Students*, in *Nevada Law Journal*, vol. 20, No. 2, 2020, pp. 457 ss.; F. Basile, *Intelligenza artificiale e diritto penale: qualche aggiornamento e qualche nuova riflessione*, in F. Basile, M. Caterini, S. Romano (a cura di), *Il sistema penale ai confini delle hard sciences. Percorsi epistemologici tra neuro-scienze e intelligenza artificiale*, Pacini Giuridica, 2021, p. 14.

<sup>10</sup> Cfr. J.S. Wormith, *Automated Offender Risk Assessment: The Next Generation or a Black Hole?*, in *American Society of Criminology*, vol. 16, No.1, 2017, pp. 281, 288 ss.

<sup>11</sup> J. Villasenor, V. Foggo, *Artificial Intelligence. Due Process and Criminal Sentencing*, in *Michigan State Law Review*, vol. 2020, No. 2, 2020, p. 296.

riferimento a degli algoritmi, non è anche vero il contrario, perché non tutti gli algoritmi sono di AI<sup>12</sup>.

Gli strumenti di *risk assessment* contemporanei sono sistemi algoritmici basati sulla regressione logistica e altri metodi statistici di classificazione e vengono annoverati tra i “*simple machine learning*” tools<sup>13</sup>, distinguendosi dai “*real machine learning*” tools, rappresentati da sistemi algoritmici più complessi ed evoluti.

Nel caso di algoritmi di AI che “imparano dall’esperienza” e sono suscettibili di evoluzione anche indipendentemente dalla supervisione dell’uomo, si pongono così problematiche ulteriori rispetto a quelle già sollevate dagli strumenti correntemente in uso; la segretezza che ancora caratterizza le regole sul funzionamento di questi tools non permette peraltro di escludere che essi siano già utilizzati nel *criminal risk assessment*<sup>14</sup>.

Le ragioni per le quali il legislatore italiano potrebbe decidere di dare ingresso a sistemi algoritmici di *risk assessment* nel procedimento penale dovrebbero risiedere nella maggiore attendibilità<sup>15</sup> e imparzialità che *tali tools* garantirebbero nella valutazione della pericolosità criminale<sup>16</sup>, seguendo l’esempio degli USA, dove da quasi mezzo secolo si ricorre a metodi statistici di valutazione del rischio<sup>17</sup>, sul presupposto che essi assicurino valutazioni scevre da *bias* cognitivi che potrebbero invece caratterizzare quelle di un giudice in persona. Eppure, oggi i risultati prodotti dal ricorso a tali strumenti possono portarci ad affermare che tali promesse così allettanti non sembrano in concreto mantenute sul fronte delle garanzie di maggiore imparzialità che dovrebbero derivarne.

Fino ad ora gli algoritmi utilizzati per calcolare la pericolosità criminale di un imputato sembrano riprodurre infatti quegli stessi *bias* di natura discriminatoria, poiché essi si basano su calcoli statistici riferiti ad una pluralità di persone accomunate da medesime caratteristiche e raggruppate entro “classi di rischio”, e prendono quindi in considerazione anche elementi che non riguardano solo l’imputato e finiscono per essere spesso influenzati dall’etnia e dallo *status* socio-economico<sup>18</sup>.

Peraltro, anche quando si considerano fattori che riguardano solo l’imputato, come ad esempio precedenti arresti o condanne, invero non si tratta di una valutazione

---

<sup>12</sup> *Idem*, p. 301: «an algorithm is a sequence of steps to move towards a goal [...]. By contrast, an AI algorithm learns from experience, leading to changes over time in the nature of the algorithm itself» [«Un algoritmo è una sequenza di steps verso un obiettivo [...]. Al contrario, un algoritmo di AI impara dall’esperienza, determinando nel corso del tempo cambiamenti nella natura dell’algoritmo stesso»].

<sup>13</sup> A.S. Niederman et al., *The Institutional Life of Algorithmic Risk Assessment*, cit., p. 713.

<sup>14</sup> J. Villasenor, V. Foggo, *Artificial Intelligence*, cit., pp. 301-302.

<sup>15</sup> Secondo alcuni studi, un sistema algoritmico di calcolo della recidiva garantirebbe maggiore precisione rispetto a quella di un essere umano: A.M. Holsinger et al., *A Rejoinder to Dressel and Farid: New Study Finds Computer Algorithm is More Accurate than Humans at Predicting Arrest and as Good as a Group of 20 Lay Experts*, 2018, pp. 50 ss., disponibile a [questo indirizzo](#).

<sup>16</sup> Per un’analisi in chiave critica di tale aspetto, F. Basile, *Intelligenza artificiale e diritto penale: qualche aggiornamento e qualche nuova riflessione*, cit., p. 21; M. Amisano, *Profetica-mente: neuroscienze, intelligenza artificiale e previsione*, *ivi*, pp. 138 ss.

<sup>17</sup> William D. Burrell, *Risk and Needs Assessment in Probation and Parole*, cit., pp. 23-24.

<sup>18</sup> Si genera così il c.d. “*algorithmic bias*”, che riguarda situazioni nelle quali un individuo o un gruppo di individui è ingiustamente favorito o discriminato: per tale definizione, v. Sourdin, *Judges, Technology and Artificial Intelligence*, cit., p. 72.

“individualizzata”, ma che può riflettere delle vere e proprie “tendenze” della giustizia penale: è quanto avviene proprio nel sistema statunitense, caratterizzato da un accentramento delle funzioni di polizia verso alcune categorie di soggetti già ritenuti “più a rischio” di comportamenti antisociali, che si traduce in numeri sensibilmente più alti di condanne per imputati appartenenti a tali categorie, determinando così *outputs* distorti<sup>19</sup>. Ed anzi, i *risk assessment tools* non solo sembrano riprodurre gli stessi errori del passato<sup>20</sup> perché consegnano risultati “contaminati” da fattori anche indirettamente influenzati dalle condizioni economico-sociali e talvolta anche etniche<sup>21</sup>, ma determinano il rischio di moltiplicare i casi di *biased outputs* a causa di errori di calcolo o di codificazione, anche quando siano elaborati proprio allo scopo di annullare tale rischio, come i modelli di *risk assessment* più avanzati<sup>22</sup>.

In ordine all’obiettivo perseguito con il ricorso ai sistemi algoritmici di *risk assessment* durante le varie fasi del procedimento penale, si tratta di temperare una duplice esigenza: da un lato, la tutela della sicurezza e dell’ordine pubblico e, dall’altro, i diritti di libertà dell’imputato. In particolare, i *risk assessment tools* possono essere utilizzati al fine di:

- 1) valutare l’eventuale sussistenza di condizioni per mantenere la carcerazione preventiva quando prevista (*pretrial risk assessment instruments-PRAIs*);
- 2) valutare il rischio di recidiva o l’ammissibilità di misure alternative alla detenzione (*risk assessment instruments-RAIs*).

---

<sup>19</sup> Tale aspetto si interseca con le criticità sollevate dall’uso di *software* di c.d. polizia predittiva: in tema, F. Basile, *Intelligenza artificiale e diritto penale*, cit. p. 18: «se, ad esempio, un software predittivo individua una determinata “zona calda”, i controlli e i pattugliamenti della polizia in quella zona si intensificheranno, con inevitabile conseguente crescita del tasso dei reati rilevati dalla polizia in quella zona, che diventerà, quindi, ancora più “calda”, mentre altre zone, originariamente non ricondotte nelle “zone calde”, e quindi non presidiate dalla polizia, rischiano di rimanere, o di diventare, per anni zone “fredde”, ove la commissione di reati non viene adeguatamente monitorata».

<sup>20</sup> Si veda C. O’Neil, *Weapons of Math Destruction. How Big Data Increases Inequality and threatens Democracy*, Penguin Books Ltd., New York, 2016, pp. 14, 40 e 162, secondo cui gli algoritmi riprodurrebbero gli errori dei giudici nel processo penale, in quanto elaborati da un essere umano e poiché codificano regole ed errori delle esperienze pregresse; S. Barocas, A.D. Selbst, *Big Data’s Disparate Impact*, in *California Law Review*, vol. 104, No. 3, 2016, pp. 671 ss.

<sup>21</sup> Ne costituisce un esempio l’*IOWA Risk Revised (IRR) tool*, che utilizza anche *inputs* che riguardano l’impiego dell’imputato, l’*housing status* e precedenti condanne, tutti fattori storicamente influenzati da disuguaglianze sociali e discriminazioni razziali. Negli Stati Uniti, una fascia di popolazione significativamente discriminata dall’utilizzo di tale algoritmo per il calcolo della recidiva è rappresentata dagli afro-americani, che risultano statisticamente più inclini alla recidiva perché si registra un numero elevato di condanne a carico di soggetti di tale origine etnica. Bisogna tuttavia considerare che alla base di una condanna può esservi anche l’impossibilità economica dell’imputato ad affrontare le spese processuali: i *Public Attorneys* spesso spingono infatti i loro assistiti al patteggiamento, con la conseguenza che la percentuale di persone condannate sarà inevitabilmente influenzata dalle risorse economiche a disposizione, come rilevano B. Kutateladze *et al.*, *Cumulative Disadvantage: Examining Racial and Ethnic Disparity in Prosecution and Sentencing*, in *Criminology: An Interdisciplinary Journal*, vol. 52, No.3, 2014, pp. 514 ss.

<sup>22</sup> Cfr. J.J. Avery, *An Uneasy Dance with Data: Racial Bias in Criminal Law*, in *Southern California Law Review Postscript*, vol. 93, No. 28, 2020, p. 32: «The fear is that, at best, algorithmic decisionmaking perpetuates historical bias; at worst, it exacerbates bias» («la paura è che, nel migliore dei casi, le decisioni algoritmiche riproducono i bias del passato; nel peggiore, li aggravano»).

Tali strumenti di “predizione del rischio”, utilizzati nel sistema penale statunitense come strumenti di attuazione della c.d. “giustizia preventiva”<sup>23</sup>, sembrano dunque già preludere a quella propensione del diritto penale – in Italia ad oggi solo paventata<sup>24</sup> – ad affidare a un algoritmo la tutela dei beni giuridici: in tal caso, dell’ordine pubblico.

Sebbene negli Stati Uniti il sistema penale sia caratterizzato dall’ampio ricorso a tali *tools*, anche sul presupposto che ciò dovrebbe contribuire a contrastare il fenomeno della “*mass incarceration*”, che oggi li rende il Paese con il numero più alto di detenuti su scala mondiale, la dottrina penale americana si interroga sull’effettiva imparzialità di tali sistemi. Le perplessità sollevate riguardano l’attendibilità degli *outputs*, suscettibili di errori di calcolo o di codificazione oppure basati su *inputs* discriminatori, nonché la stessa compatibilità di tali sistemi con i principi del giusto processo sancito nel 5° e nel 14° emendamento alla Costituzione americana<sup>25</sup>.

Dinanzi a queste nuove frontiere che ridefiniscono i tratti della pericolosità criminale di un imputato, la discrezionalità del giudice continua comunque a rivestire un ruolo centrale: lungi dal disporre che si sostituissero all’organo giudicante, infatti, le legislazioni dei vari Stati statunitensi hanno reso l’uso di tali sistemi non obbligatorio bensì raccomandato, lasciando margini di discrezionalità all’organo giurisdizionale circa l’*an* e il *quomodo* del loro utilizzo<sup>26</sup>.

## 2. La California sceglie di limitare il ricorso ai sistemi algoritmici di valutazione del rischio: l’eloquente esito della *Proposition 25* sul *replacement* del *cash bail* con i *pretrial risk assessment tools*.

Nonostante le prospettive di maggiore precisione e obiettività che i sistemi algoritmici di *risk assessment* promettono di raggiungere nella valutazione della pericolosità criminale, sembra opportuno soffermarsi sull’esito di un recente referendum tenutosi in California il 3 novembre 2020, che rappresenta l’epilogo di un *iter* legislativo intrapreso dal Senato e che offre un quadro chiaro del clima di generale diffidenza che ancora sembra caratterizzare il ricorso a tali strumenti nel sistema della giustizia penale.

Con la *Proposition 25* i californiani sono stati chiamati a esprimere la propria preferenza in ordine all’abrogazione della legge sul *cash bail*, che prevedeva il rilascio

---

<sup>23</sup> La “giustizia preventiva” ha come obiettivo la prevenzione dei reati e non la punizione *tout court* del colpevole, tipica della concezione retributiva: così, C. Slobogin, *Preventive Justice: A Paradigm in Need of Testing*, in *Behavioral Sciences & the Law*, vol. 36, No.4, 2018, pp. 1 ss.

<sup>24</sup> F. Basile, *Intelligenza artificiale e diritto penale*, cit., p. 12.

<sup>25</sup> XIV Amendment, Section 1, *US Constitution*: «[...] *nor shall any state deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws*» [« (...) né ciascuno Stato dovrà privare una persona della vita, della libertà o della proprietà, senza un giusto processo; e non dovrà negare ad alcuna persona entro la propria giurisdizione l’uguale protezione delle leggi»]. La disposizione, ratificata nel 1868, garantisce il giusto processo all’interno dei singoli Stati, richiamando il medesimo principio già sancito sul piano federale dal XV Amendment.

<sup>26</sup> Si interroga sull’ampiezza della discrezionalità da lasciare al giudice nell’uso dei *risk assessment tools*, rilevando come proprio tale aspetto sia invero determinante sull’impatto di tali strumenti nella giustizia penale, M. Stevenson, *Assessing Risk Assessment*, cit., pp. 333 ss. Nello stesso senso, rilevando altresì come il vasto spazio di discrezionalità di cui dispongono gli organi giurisdizionali sarebbe la conseguenza dell’assenza di una normativa che renda obbligatorio l’utilizzo dei *tools*, B.L. Garrett, J. Monahan, *Judging Risk*, cit., p. 450.

dell'imputato dietro pagamento di una cauzione in attesa della celebrazione del processo, e al suo posto l'entrata in vigore del *California Money Bail Reform Act*, anche conosciuto come *Senate Bill 10 (SB 10)*<sup>27</sup>, che avrebbe subordinato il mantenimento della carcerazione preventiva alla sussistenza di un pericolo di fuga calcolato da sistemi algoritmici di valutazione del rischio; in caso di esito positivo del referendum, la California sarebbe diventata il primo Stato d'America ad aver modificato tale legislazione mediante la sostituzione della cauzione in contanti con il ricorso a sistemi algoritmici di *risk assessment*<sup>28</sup>.

Le ragioni che avevano condotto all'approvazione della legge muovevano dall'esigenza di ripristinare condizioni di sostanziale uguaglianza: per i promotori della riforma, subordinare il mantenimento di una misura di carcerazione preventiva all'impossibilità economica dell'imputato di sostenere l'impegno economico della cauzione significava favorire la permanenza di situazioni discriminatorie in danno dei soggetti meno abbienti, scardinando ingiustificatamente la *ratio* della carcerazione preventiva dalla pericolosità dell'imputato o dal pericolo di fuga.

La novella si sviluppava parallelamente all'evoluzione giurisprudenziale che si manifestava progressivamente nello stesso senso nelle Corti penali californiane, inaugurata dalla sentenza della Corte d'Appello di San Francisco *Humphrey on Habeas Corpus*<sup>29</sup> che, richiamando l'indirizzo legislativo espresso già da diversi decenni sulla necessità di una riforma, accoglieva l'impugnazione dell'imputato affermando l'incostituzionalità della legge sul *bail* per contrasto con i corollari del *due process*.

Il *Senate Bill 10* prevedeva l'istituzione di organi di valutazione preliminare presso i Tribunali superiori ("*Pretrial Assessment Services*"), ai quali sarebbe stato demandato il compito di valutare i rischi di recidiva o il pericolo di fuga<sup>30</sup> e formulare raccomandazioni per le condizioni di rilascio<sup>31</sup>.

---

<sup>27</sup> *Senate Bill 10* del 28 agosto 2018, pubblicato a [questo indirizzo](#). In tema, A.S. Niederman et al., *The Institutional Life of Algorithmic Risk Assessment*, cit., pp. 705 ss.

<sup>28</sup> T. Fuller, *California Is the First State to Scrap Cash Bail*, in *New York Times*, 28 agosto 2018. È opportuno comunque notare che in molte giurisdizioni della California il sistema del *cash bail* è affiancato dall'utilizzo di strumenti di valutazione del rischio. Per tale aspetto e per un'analisi dei paventati pericoli della riforma, si rinvia a T.A. Merkl, L. Arzy, *California's Referendum to Eliminate Cash Bail, Explained*, 2 ottobre 2020, disponibile a [questo link](#). Il documento è antecedente al referendum e in particolare risalta la raccomandazione di non ricorrere a strumenti di *risk assessment* in alternativa alla cauzione in contanti in caso di esito contrario alla modifica della legge sul *bail* e quindi di riforme future, allo scopo di non favorire discriminazioni razziali che sarebbero sottese ai calcoli operati da tali sistemi.

<sup>29</sup> *Humphrey (Kenneth) on H.C. (A152056), First Appellate District of San Francisco City*, 25 gennaio 2018.

<sup>30</sup> Sulla definizione di "*risk*" si veda *Senate Bill 10 (SB 10), Article 1. Definitions*, § 1320.7 (h): «"*Risk*" refers to the likelihood that a person will not appear in court as required or the likelihood that a person will commit a new crime if the person is released before adjudication of his or her current criminal offense» [«Il "rischio" si riferisce alla probabilità che una persona non si presenti in tribunale quando previsto, o alla probabilità che una persona commetta un nuovo reato se viene rilasciata prima della pronuncia della sentenza»].

<sup>31</sup> *Idem*, Article 1, § 1320.7 (g).

Il testo legislativo faceva riferimento all'utilizzo di *risk assessment tools* corroborati e validati dalla ricerca scientifica<sup>32</sup>, che avrebbero qualificato i soggetti come a "basso", "medio" o "alto" rischio, e il dato che sembra più significativo riguarda il ruolo che avrebbero giocato tali sistemi nella determinazione giudiziale: lungi dal sostituire l'organo giurisdizionale, l'*output* generato da tali strumenti sarebbe comunque rimasto soggetto a revisione da parte del giudice<sup>33</sup>, il quale avrebbe avuto ampia discrezionalità nella valutazione sottesa alla decisione finale<sup>34</sup>.

Ciononostante, il referendum si è concluso con una sensibile prevalenza di voti contrari alla riforma, pari al 56,41% dei votanti contro il 43,59%, sebbene comunque il ricorso ai *pretrial risk assessment tools* sia ricorrente su larga scala nelle Corti penali della California, come risulta da un report di dicembre 2019 elaborato dal *Public Policy Institute of California*, secondo cui 49 contee su un totale di 58 già facevano uso dei *pretrial risk assessment tools*<sup>35</sup>.

Aver scelto dunque di mantenere in vigore una legge in sé discriminatoria, perché i suoi effetti sono condizionati dalla capacità economica dell'imputato, pur di non dare ingresso a sistemi algoritmici di valutazione del rischio in quanto ritenuti forieri di discriminazioni razziali, ha significato optare per quella che tra le due è stata ritenuta, a torto o a ragione, la disciplina normativa più "garantista" per l'imputato.

Sul caso *Humphrey* si è infine pronunciata la *Supreme Court of California* nel marzo 2021, confermando l'arresto della Corte d'Appello di San Francisco e dichiarando l'incostituzionalità del *cash bail*, in quanto misura che subordina il mantenimento della carcerazione preventiva unicamente all'incapacità economica dell'arrestato di impegnarsi al pagamento della cauzione e non invece alla sussistenza di ragioni di pubblica sicurezza o di obiettivo e concreto pericolo di fuga<sup>36</sup>.

---

<sup>32</sup> Sulla definizione di "pretrial risk assessment" si veda *Senate Bill 10 (SB 10)*, Article 1, § 1320.7 (f): «"Pretrial risk assessment" means an assessment conducted by Pretrial Assessment Services with the use of a validated risk assessment tool, designed to provide information about the risk of a person's failure to appear in court as required or the risk to public safety due to the commission of a new criminal offense if the person is released before adjudication of his or her current criminal offense» [«Per "pretrial risk assessment" si intende una valutazione condotta dai servizi di valutazione pre-processuale con l'uso di uno strumento di valutazione del rischio convalidato, progettato per fornire informazioni sul rischio che una persona non compaia in tribunale quando previsto, o sul rischio per la pubblica sicurezza dovuto alla commissione di un nuovo reato se la persona viene rilasciata prima della pronuncia della sentenza»].

<sup>33</sup> *Senate Bill 10 (SB 10)*, Article 5, § 1320.13.

<sup>34</sup> A.S. Niederman et al., *The Institutional Life of Algorithmic Risk Assessment*, cit., p. 719, che evidenziano però come la progettazione e l'uso di *risk assessment tools* comprimano comunque la discrezionalità dell'organo giudicante perché implicano scelte di natura tecnica e politica.

<sup>35</sup> H. Harris, J. Goss, A. Gumbs, *Pretrial Risk Assessment in California*, *Public Policy Institute of California*, report di dicembre 2019.

<sup>36</sup> *In re Humphrey (S247278)*, *Supreme Court of California*, 25 marzo 2021. La *California Supreme Court* ha ricordato che il pagamento della cauzione a cui è subordinato il rilascio del soggetto arrestato è finalizzato ad assicurare la presenza dell'imputato al processo e a tutelare la sicurezza collettiva. Purtroppo, ha precisato la Corte, in tale fase la colpevolezza non è ancora stata accertata, di talché la limitazione della libertà personale dovrebbe essere svincolata dalle capacità economiche dell'arrestato, perché solo così il *due process* può dirsi effettivamente rispettato, anche nel rispetto delle garanzie di uguaglianza: «An arrestee's release pending trial is often conditioned on whether the arrestee can make bail. To do so, an arrestee posts security – in the form of cash, property, or (more often) a commercial bail bond – which is forfeited if the arrestee later fails to appear

### 3. Il *First Step Act* e il nuovo *recidivism risk assessment system* del *Department of Justice*: PATTERN.

Restando in tema di riforme, stavolta sul piano federale, negli Stati Uniti grandi aspettative sono ora riposte sull'imponente riforma della giustizia penale attuata con il *First Step Act (Formerly Incarcerated Reenter Society Transformed Safely Transitioning Every Person Act)* del 21 dicembre 2018<sup>37</sup>, che mira a contrastare il fenomeno della *mass incarceration*<sup>38</sup> e a depurarla dalle distorsioni originate da discriminazioni razziali; il progetto si presenta ambizioso alla luce degli obiettivi prefissati, tra cui quello di limitare il ricorso a misure restrittive favorendo programmi di reinserimento elaborati dal *Department of Justice*.

Il *First Step Act* ha affidato ai sistemi algoritmici un ruolo centrale e anzi determinante per il conseguimento delle finalità perseguite: nel *Title I, Section 101*<sup>39</sup>, con il § 3632 dell'articolato, rubricato *Development of risk and needs assessment system*<sup>40</sup>, il Congresso ha autorizzato il Procuratore Generale degli Stati Uniti a sviluppare un nuovo *risk assessment system*<sup>41</sup> destinato all'utilizzo da parte del *Bureau of Prisons* (BOP), strumentale all'ammissione alla *early release* e a concedere riduzioni di pena ai soggetti detenuti negli istituti penitenziari federali<sup>42</sup> che presentino un rischio "basso" o "minimo" di tornare a delinquere.

---

*in court. Those who can't afford to satisfy the bail condition remain in jail until the end of the criminal proceedings. Underlying this arrangement is a major premise: that the state has a compelling interest in assuring the arrestee's appearance at trial and protecting the safety of the victim as well as the public. Yet those incarcerated pending trial – who have not yet been convicted of a charged crime – unquestionably suffer a "direct 'grievous loss'" of freedom in addition to other potential injuries (...). We find merit in Humphrey's claim. The common practice of conditioning freedom solely on whether an arrestee can afford bail is unconstitutional (...). Allowing the government to detain an arrestee without such procedural protections would violate state and federal principles of equal protection and due process that must be honored in practice, not just in principle».*

<sup>37</sup> *First Step Act (Public Law 115-391)*, pubblicato a [questo indirizzo](#).

<sup>38</sup> Nel disegno di legge [H.R. 5682-First Step Act del 23 maggio 2018](#) si indicava la riduzione del *recidivism* quale finalità principale del *First Step Act*, che veniva originariamente definito come «*an Act to provide for programs to help reduce the risk that prisoners will recidivate upon release from prison, and for other purposes*».

<sup>39</sup> Le cui disposizioni sono ora contenute nel nuovo *Subchapter D ("Risk and needs assessment system")* del *Chapter 229, Title 18, US Code*.

<sup>40</sup> *First Step Act, Sec. 101, §3632*.

<sup>41</sup> La definizione di *risk assessment tool* si trova nella *Sec. 101, § 3635 (6)* del *First Step Act*: «RISK AND NEEDS ASSESSMENT TOOL– The term "risk and needs assessment tool" means an objective and statistically validated method through which information is collected and evaluated to determine: "(A) the risk that a prisoner will recidivate upon release from prison; (B) the recidivism reduction programs that will best minimize the risk that the prisoner will recidivate upon release from prison; and"(C) the periodic reassessment of risk that a prisoner will recidivate upon release from prison, based on factors including indicators of progress and of regression, that are dynamic and that can reasonably be expected to change while in prison». » [«Il termine "*risk and needs assessment tool*" indica un metodo oggettivo e statisticamente convalidato attraverso il quale vengono raccolte e valutate le informazioni per determinare: "(A) il rischio che un detenuto sia recidivo al momento della liberazione; (B) i programmi che meglio ridurranno il rischio che il detenuto torni a delinquere dopo la liberazione; e (C) la rivalutazione periodica del rischio di recidiva al momento del rilascio dal carcere, sulla base di fattori tra cui gli indicatori di progresso e di regressione, che sono dinamici e che ci si può ragionevolmente aspettare che cambino durante il periodo di detenzione »].

<sup>42</sup> Ad eccezione dei detenuti che abbiano commesso uno dei reati indicati nella *Sec. 101, § 3632(d)(4)(D)* del *First Step Act*.

Il nuovo sistema è stato annunciato dal *Department of Justice (DOJ)* il 19 luglio del 2019<sup>43</sup>, è denominato PATTERN (*Prisoner Assessment Tool Targeting Estimated Risk and Needs*) e rappresenta uno strumento algoritmico di valutazione del rischio-ricidiva di natura istituzionale, per tale ragione non coperto dal segreto industriale che caratterizza tutti gli altri *tools* simili che per questo non avevano trovato il consenso della dottrina penale statunitense<sup>44</sup>.

Il *DOJ* sembra aver puntato sulla trasparenza delle informazioni e sul coinvolgimento dei vari portatori di interessi privati e pubblici e ciò risulta evidente sin dalla fase "genetica" di PATTERN: durante lo sviluppo del *tool* sono state infatti tenute tre *listening sessions* con l'obiettivo di dare voce ad esperti, *stakeholders* tra cui organizzazioni rappresentative delle vittime di reati e di interessi pubblici.

Nel *report* pubblicato il 15 gennaio 2020, il *DOJ* annunciava così che gli aggiornamenti e le migliorie apportate a PATTERN hanno rappresentato il risultato diretto dei suggerimenti e delle sollecitazioni provenienti da tali organismi coinvolti nel procedimento di elaborazione del nuovo sistema<sup>45</sup>.

PATTERN rappresenta quindi un significativo passo in avanti verso il superamento di una delle maggiori criticità sollevate dai sistemi utilizzati nelle varie fasi del procedimento penale, ovvero l'assenza di trasparenza.

Purtuttavia, tra le critiche sollevate, vi è quella relativa all'utilizzo di alcuni dati elaborati da un altro sistema, BRAVO (*Bureau Risk and Verification Observation*)<sup>46</sup>, del quale sono poche le informazioni disponibili al pubblico; di talché, si afferma, PATTERN è ancora lontano dal garantire il soddisfacimento delle aspettative in esso riposte<sup>47</sup>.

---

<sup>43</sup> *DOJ*, Office of the Attorney General, *The First Step Act of 2018: Risk and Needs Assessment System – Update*, gennaio 2020, p. 1.

<sup>44</sup> Tra gli altri, R. Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, in *Stanford Law Review*, vol. 70, No. 5, 2018, pp. 1343 ss; A. Rizer, C. Watney, *Artificial Intelligence Can Make Our Jail System More Efficient, Equitable, and Just*, in *Texas Review of Law & Politics*, vol. 23, No. 1, 2019, pp. 214-215. Si rimanda altresì allo *statement* di D.E. Patton, *Executive Director del Federal Defenders of New York*, *Oversight Hearing on "The Federal Bureau of Prisons and Implementation of The First Step Act"*, 2019, pp. 3,8, il quale osservava che la segretezza dei modelli black box genera dubbi sulla ragionevolezza dei risultati prodotti dall'algoritmo: «*Across risk assessments in criminal justice, the secrecy that permeates black box instruments causes significant concerns about how reasonable they are in practice*».

<sup>45</sup> Si vedano il *report* del *DOJ*, *The First Step Act of 2018: Risk and Needs Assessment System*, 2019, e il comunicato stampa del 15 gennaio 2020, *Department of Justice Announces Enhancements to the Risk Assessment System and Updates on First Step Act Implementation*.

<sup>46</sup> BRAVO è un sistema di classificazione afferente al BOP e realizzato negli anni '70: A.B. Cyphert, *Reprogramming Recidivism*, cit., p. 346; C.D. Stimson, *The First Step Act's Risk and Needs Assessment Program*, cit., p. 13. Il *DOJ* ha precisato che gli *items* e gli *scores* di PATTERN derivano da quelli di BRAVO e non dalla sua versione aggiornata, BRAVO-R ("*Bureau Risk Assessment Verification and Observation Revised*"), come erroneamente era stato comunicato nel precedente *report*: v. *The First Step Act of 2018: Risk and Needs Assessment System – Update*, cit., p.35. La mancanza di trasparenza caratterizza peraltro anche il successivo BRAVO-R, come rilevava A.P. Albert nel corso di una delle *listening sessions*: *Stakeholder Statements Submitted to NIJ's September Listening Sessions: Comments on the Department of Justice's Report, The First Step Act of 2018: Risk and Needs Assessment System*, 2019, p.28.

<sup>47</sup> A.P. Albert, *Stakeholder Statements Submitted to NIJ's September Listening Sessions*, cit., p. 356 ss., che evidenziava come fossero state disattese le indicazioni del Congresso, che al § 3632(a) del *First Step Act*

È comunque evidente l'affidabilità che il Congresso ritiene caratterizzi le decisioni algoritmiche assunte sulla base dei *risk assessment systems* per determinare il rischio di recidiva, definiti strumenti "obiettivi e statisticamente convalidati"<sup>48</sup>. Tra i profili di maggiore interesse della riforma risulta quello per cui, oltre ad assegnare uno *score* che determina l'inserimento in una fascia più o meno alta di rischio, il sistema è in grado di permettere al detenuto di diminuire il proprio *risk score* (*recte*: i c.d. fattori "dinamici" di valutazione, che si contrappongono a quelli "statici"<sup>49</sup>) nel caso in cui partecipi attivamente e con profitto ai programmi specifici elaborati dal *Department of Justice* volti al reinserimento dei soggetti detenuti<sup>50</sup>, sebbene qualche perplessità sia stata comunque sollevata in ordine alle modalità con cui gli organi preposti all'espletamento di tale procedura negli istituti penitenziari federali eserciteranno la discrezionalità per effettuare tali valutazioni<sup>51</sup>.

PATTERN sfrutta l'evoluzione informatica già sviluppata per mettere a punto i precedenti sistemi algoritmici come COMPAS, e mette in campo altresì le più recenti tecnologie che fanno di tale *tool* il più evoluto strumento di *risk assessment*, in grado di dare informazioni più accurate degli altri sistemi in una misura stimata pari al 15%<sup>52</sup>.

Il nuovo sistema algoritmico, ultimato nel gennaio 2019, è programmato per prevedere il rischio di recidiva entro tre anni dalla liberazione<sup>53</sup>, distingue il "*violent recidivism*" dal "*general recidivism*" – nonostante sia stato evidenziato un *deficit* di determinatezza sotto il profilo di tali definizioni<sup>54</sup> – e assegna diversamente gli *scores* per ciascuna delle due tipologie, differenziando altresì il rischio-recidiva delle donne da quello degli uomini e prendendo in considerazione diciassette *items*.

---

aveva incaricato il Procuratore Generale di rendere pubbliche le informazioni sul funzionamento del sistema non oltre 180 giorni dall'entrata in vigore del *First Step Act*. Si veda anche la *Comment Letter to the Department of Justice on PATTERN First Step Act* della *Leadership Conference on Civil and Human Rights*, 3 settembre 2019.

<sup>48</sup> *First Step Act*, Sec. 101 (a), § 3635 (3).

<sup>49</sup> DOJ, *Office of the Attorney General, The First Step Act of 2018: Risk and Needs Assessment System – Update*, cit., pp. 7 ss. I fattori "statici" di rischio condizionano comunque negativamente la valutazione poiché si basano su elementi quali l'età e i precedenti penali: il risultato è che più della metà dei soggetti risulta ad "alto" o "medio" rischio. Sul punto, si vedano le critiche mosse nel corso di una delle *listening sessions*: K. Gotsch, *Stakeholder Statements Submitted to NIJ's September Listening Sessions*, cit., p. 5.

<sup>50</sup> Qualche perplessità è stata sollevata in ordine all'adeguatezza dei programmi del *Bureau of Prison* (BOP) a favorire la rieducazione e il reinserimento del condannato: su tale aspetto si rinvia a R. Anspach, *How the First Step Act Got People Out of Prison and Back With Their Families*, dicembre 2019.

<sup>51</sup> B. L. Garrett, J. Monahan, *Judging Risk*, cit., pp. 439 ss.

<sup>52</sup> Tale è la stima riportata dal DOJ, *Office of the Attorney General, in The First Step Act of 2018: Risk and Needs Assessment System*, 2019, p. 57, fig.3, disponibile a [questo indirizzo](#). Si veda altresì C.D. Stimson, *The First Step Act's Risk and Needs Assessment Program*, cit., 14-15, che riporta i risultati di studi secondo i quali PATTERN ha registrato una capacità predittiva pari all'80% per gli uomini e al 70% per le donne, a fronte del 67% di quella di COMPAS.

<sup>53</sup> DOJ, *Office of the Attorney General*, cit., p. 50; *id.*, *2020 Review and Revalidation of the First Step Act Risk Assessment Tool*, gennaio 2021, p.2

<sup>54</sup> Manca, ad esempio, una definizione che consenta di individuare i "*violent arrests*": sul punto, A.B. Cyphert, *Reprogramming Recidivism*, cit., p. 348. La bipartizione operata tra *general* e *violent recidivism* rappresenta peraltro una recente evoluzione dei *risk assessment tools*, come riportato dal DOJ nel *The First Step Act of 2018: Risk and Needs Assessment System*, cit., p. 44.

Il *tool* è stato già utilizzato su larga scala nel 2020, permettendo di individuare circa 7000 detenuti presso gli istituti penitenziari federali ritenuti idonei all'isolamento domiciliare durante la pandemia da Covid-19<sup>55</sup>. "Inaugurato" durante la pandemia, PATTERN ha contribuito così alla salvaguardia della salute dei soggetti valutati positivamente, grazie alla intuibile rapidità con cui tali valutazioni sono state effettuate.

Nonostante gli sforzi compiuti dal *DOJ* per garantire massima trasparenza al nuovo sistema per lo meno nella fase della sua implementazione, al momento rimangono irrisolti alcuni interrogativi riguardo al raggiungimento dell'auspicata imparzialità di giudizio che dovrebbe contraddistinguere uno strumento algoritmico di *recidivism risk assessment*: secondo uno studio preliminare effettuato dal *Brennan Center for Justice* infatti, sul campione utilizzato per lo studio PATTERN ha individuato oltre la metà di tutti i *black men* (53%) ad alto rischio di recidiva, rispetto al 29% di *white men*; il 30% di *white men* è stato classificato "a basso rischio", mentre solo il 7% dei *black men* ha avuto la stessa classificazione<sup>56</sup>.

Dunque, si conclude nello studio preliminare, anche se PATTERN è stato progettato per essere "neutrale" rispetto alle origini etniche dei soggetti sottoposti a valutazione<sup>57</sup>, nei suoi *outputs* rischia di produrre risultati discriminatori<sup>58</sup> e ciò, si afferma, merita un attento monitoraggio: ciò perché a differenza di quanto avviene nell'utilizzo di altri sistemi algoritmici di *risk assessment* come ad esempio COMPAS, il cui risultato rappresenta uno degli elementi su cui basare la decisione, per le finalità del *First Step Act* l'*output* generato da PATTERN è l'unico elemento da cui dipende la valutazione<sup>59</sup>.

Alla luce della rilevanza determinante degli *output* generati da PATTERN, quindi, il *First Step Act* ha opportunamente previsto l'istituzione di una *Independent Review*

---

<sup>55</sup> Si veda il *Memorandum from the Office of the Attorney General for Director of Bureau Prisons* del 26 marzo 2020, nel quale si indica lo score fornito dal sistema PATTERN tra i requisiti da considerare per determinare i detenuti idonei all'isolamento domiciliare, pubblicato a [questo indirizzo](#). L'elevato numero di detenuti appartenenti alle minoranze etniche ha determinato che su di essi si sono abbattuti in misura preponderante gli effetti della pandemia da Covid-19: cfr. il Rapporto del CDC (*Centers of Disease Control and Prevention*), pubblicato a [questo indirizzo](#), in cui si indica la giustizia penale come uno degli ambiti di discriminazione delle minoranze che ha generato un'esposizione maggiore di questi soggetti al rischio malattia o morte da Covid-19.

<sup>56</sup> Fonte: *Brennan Center for Justice*, disponibile a [questo indirizzo](#).

<sup>57</sup> DOJ, *2020 Review and Revalidation of the First Step Act Risk Assessment Tool*, cit., p. 7.

<sup>58</sup> La difficoltà nell'epurare dai *bias* originati da diseguaglianze sociali gli *outputs* di PATTERN derivano dal fatto che essi si basano su *dataset* del BOP, che riproducono tutte le discriminazioni razziali che caratterizzano il sistema della giustizia penale statunitense. Rileva che ogni valutazione prognostica basata su dati riferiti al passato produrrà *outputs* discriminatori S.G. Mayson, *Bias In, Bias Out*, in *The Yale Law Journal*, vol. 128, No. 8, 2019, p. 2251. Con l'obiettivo del raggiungimento di una maggiore imparzialità del nuovo *risk assessment tool*, nello *statement* contenuto nella *Brennan Center for Justice letter*, cit., pp. 1 ss., erano state avanzate proposte volte ad eliminare il numero di arresti a carico del soggetto tra i dati da calcolare, mantenendo invece il numero di condanne inflitte e di *incarcerations*. In senso critico sull'inserimento del numero di arresti tra i parametri da valutare, anche A.P. Albert, *Stakeholder Statements Submitted to NIJ's September Listening Sessions*, cit., p. 29: «PATTERN assumes that an arrest itself is proof of a crime without any Due Process» [«PATTERN presume che un arresto sia di per sé prova della commissione di un reato senza alcun giusto processo»]. In accoglimento di tali istanze, il DOJ ha espunto dagli *inputs* l'età del primo arresto e della prima condanna, ritenuti fattori discriminatori, sebbene si sia dimostrato che in seguito a tale modifica l'accuratezza predittiva di PATTERN si fosse ridotta di circa l'1%: in tema, D. Stimson, *The First Step Act's Risk and Needs Assessment Program*, cit., p. 18.

<sup>59</sup> A.B. Cyphert, *Reprogramming Recidivism*, cit., p. 342.

*Committee* che annualmente revisioni il sistema al fine di garantirne l'affidabilità e proprio nel corso della prima verifica sono stati rilevati errori di codificazione o di calcolo nell'assegnazione degli *scores*<sup>60</sup>, essendo stata ad esempio calcolata nella prima versione la variabile "età al momento della valutazione" del rischio come "età al momento della liberazione", così determinando un errore nell'*output*<sup>61</sup> che ha richiesto una tempestiva correzione e rivalutazione del rischio nei casi coinvolti dagli errori.

Nonostante i passi in avanti compiuti verso il raggiungimento di una maggiore trasparenza e l'apprezzabile coinvolgimento anche di figure giuridiche e di organismi rappresentativi dei vari interessi privati e pubblici nella fase di implementazione del sistema, parte della dottrina statunitense rimarca come ad oggi non si sia, ancora una volta, centrato l'obiettivo.

Più in generale, la segretezza di alcuni dati utilizzati dai sistemi di *risk assessment* correntemente in uso – compreso comunque PATTERN, perché utilizza alcuni *items* del poco esplorato BRAVO – non permette di sapere se qualcuno di questi *tools* sfrutti già le potenzialità dell'AI per l'autoapprendimento<sup>62</sup>.

Ciò varrebbe dunque anche per PATTERN che, si afferma, se alla luce delle sue caratteristiche rappresenta l'ultima evoluzione dei modelli di quarta generazione<sup>63</sup>, dall'altra parte si caratterizza per l'insufficienza di informazioni relative al suo funzionamento<sup>64</sup>, di talché non può escludersi che esso sfrutti le potenzialità del *machine learning* per il calcolo del rischio-ricidiva<sup>65</sup>: nemmeno sotto il profilo della trasparenza, dunque, ad oggi PATTERN sembra segnare l'auspicato iato rispetto ai suoi predecessori.

#### **4. Le insidie nascoste dei *risk assessment tools* sotto la lente dei principi costituzionali del giusto processo: l'antesignana prospettiva statunitense.**

I significativi sforzi compiuti dal Congresso nella fase di implementazione di PATTERN non sono stati supportati da una più ampia riforma del sistema penale che regolamentasse l'utilizzo dei sistemi algoritmici per il *recidivism risk assessment* nel rispetto dei principi costituzionali del *due process*. La dottrina statunitense più sensibile a questi aspetti ha così evidenziato l'impatto dirompente di tali sistemi sulle garanzie costituzionali dell'imputato e i connessi rischi di compromissione del diritto di difesa.

Il pericolo di disattendere gli enunciati della Costituzione si fa poi ancora più insidioso nel caso in cui il *risk assessment* si basi sull'AI, ovvero quando l'algoritmo, per le sue caratteristiche, sia in grado di "apprendere" autonomamente ed evolversi anche indipendentemente dall'intervento, e quindi dal controllo, dell'uomo: in tal caso, si pongono

---

<sup>60</sup> Come riportato dal DOJ in *2020 Review and Revalidation of the First Step Act Risk Assessment Tool*, cit., p. 7.

<sup>61</sup> *Idem*, p. 5.

<sup>62</sup> J. Villasenor, V. Foggo, *Artificial Intelligence*, cit., pp. 301, 302.

<sup>63</sup> C.D. Stimson, *The First Step Act's Risk and Needs Assessment Program*, cit., p. 13.

<sup>64</sup> C.N. Jones, *A broken PATTERN: A Look at the Flawed Risk and Needs Assessment Tool of the First Step Act*, in *Howard Human & Civil Rights Law Review*, vol. 5, No. 2, 2021, pp. 202 ss.

<sup>65</sup> In senso critico sul punto, A.B. Cyphert, *Reprogramming Recidivism*, cit., pp. 360-361.

dunque problematicità ulteriori rispetto a quelle già rilevate sotto il profilo della trasparenza per i sistemi algoritmici tradizionali di *risk assessment*.

Mentre, infatti, un algoritmo tradizionale è statico, poiché non si evolve nel corso del tempo e quindi, per lo meno in assenza della tutela del segreto industriale, sarebbe più facile dimostrare un errore dell'*output* che ha prodotto, ciò non è possibile se l'algoritmo è di *AI* ed è dinamico, ovvero progettato per mutare in forza del *machine learning*, perché è in continua evoluzione e di conseguenza sarebbe arduo per l'imputato dimostrare l'inesattezza degli *inputs* sui quali l'*output* si basa<sup>66</sup>.

Ma anche quando i *risk assessment tools* non siano basati sul *machine learning*, la segretezza che contraddistingue le regole sul loro funzionamento impedisce all'imputato l'accesso alle informazioni su cui si è basata la valutazione e si traduce anche in questo caso in una violazione del giusto processo, precludendo una verifica di eventuali *inputs* distorti che potrebbero aver determinato una significativa sopravvalutazione del rischio.

Un processo "giusto" richiede, infatti, nel quadro delle garanzie costituzionali statunitensi, come anche di quelle italiane, il rispetto dell'imperativo secondo cui non si può precludere all'imputato l'accesso alle informazioni su cui si è basata una data previsione o decisione.

L'enunciato espresso nel 5° e nel 14° emendamento alla Costituzione americana, conosciuto come *Due Process Clause*, rimanda a un insieme di corollari posti a tutela del diritto al giusto processo e a quello di difesa, tra i quali il principio di *accuracy*, secondo cui è diritto dell'imputato che il provvedimento giurisdizionale si basi su fatti puntuali e precisi. Trasposto tale principio nel "giudizio algoritmico", il "fatto" non è altro che l'*input* sul quale la valutazione, e quindi la decisione, si basa: nel caso di *inputs* distorti, un "giusto processo" deve garantire all'imputato la possibilità di impugnare la decisione e dunque di conoscere l'erroneità alla base del calcolo prodotto dal *tool*, dal momento che l'onere della prova nel sistema delle impugnazioni anche nel sistema statunitense grava sul *defendant*.

Dal corollario dell'*accuracy* emerge in tutta la sua evidenza la centralità del rispetto di un altro dei corollari del *Due Process*, quello di *transparency*: si tratta del principio sul cui rispetto probabilmente la dottrina penale, anche italiana<sup>67</sup>, si è soffermata maggiormente quando ha posto in luce i punti di frizione più evidenti dei sistemi algoritmici con il diritto di difesa dell'imputato, ovvero proprio l'assenza di trasparenza degli *outputs* prodotti e posti alla base di provvedimenti giurisdizionali.

Sul piano delle garanzie costituzionali poste a presidio del diritto dell'imputato al giusto processo, l'assenza di trasparenza si traduce sostanzialmente in una *probatio diabolica* che ricade sull'imputato: la segretezza dei metodi di funzionamento e di

---

<sup>66</sup> Su tale aspetto, sulla distinzione tra algoritmi statici e dinamici e in relazione alla diversa funzionalità dell'*AI* nell'uno e nell'altro caso si rinvia a J. Villasenor, V. Foggo, *Artificial Intelligence*, cit., pp. 311 ss.

<sup>67</sup> Nella dottrina penalistica italiana si vedano F. Basile, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *questa rivista*, 29 settembre 2019, pp. 14 e 22; M. Caterini, *Il giudice penale robot*, in *La legislazione penale*, 2020, pp. 11-12; M. Gialuz, *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti e Europa*, in *Dir.Pen.Cont.*, 29 maggio 2019 pp. 4 ss.; S. Quattrocolo, *Equo processo penale e sfide della società algoritmica*, in *BioLaw Journal*, n. 1, 2019, pp. 142 ss..

processamento dei *dataset*, in quanto coperta dalla tutela del segreto industriale, preclude infatti l'esercizio del diritto di difesa e, in definitiva, determina perfino una legittimazione del *deficit* di trasparenza.

Sotto tale aspetto, l'implementazione di PATTERN ha segnato un importante cambiamento di rotta verso il rigoroso rispetto della *transparency*, non solo perché il nuovo sistema algoritmico non è di proprietà di società private e quindi è svincolato dalle regole sul segreto industriale<sup>68</sup>, nonché per le revisioni previste dal *First Step Act*, effettuate da organismi indipendenti nominati dal *DOJ* e i cui primi risultati sono stati già divulgati al pubblico, ma anche per il coinvolgimento, nella fase "genetica" del nuovo *tool*, di enti e associazioni rappresentative degli interessi in gioco, che inaugura un nuovo *modus operandi*.

PATTERN, infine, si rivela il *recidivism prediction tool* di cui ad oggi sono state divulgate più informazioni di quelle rese disponibili dalle società private produttrici di sistemi algoritmici di *risk assessment*<sup>69</sup>. A fronte di questi dati incoraggianti, come si è visto, tuttavia le aspettative riposte sembrano essere state per il momento disattese e da più parti si sollevano istanze volte ad una maggiore divulgazione di dati da parte del *DOJ*<sup>70</sup>.

Sempre sotto il profilo dell'*accuracy*, è stato osservato come i *risk assessment tools* generino potenziali valutazioni distorte, essendo basati su calcoli statistici e non sulla pericolosità specifica del singolo individuo soggetto a valutazione: l'accuratezza della valutazione quindi sconta l'assenza di "individualizzazione" nel giudizio prognostico, di talché un individuo classificato tra i soggetti ad "alto rischio" potrebbe invero non esserlo, proprio perché il sistema prende in considerazione gruppi di individui in base all'alea di rischio.

L'impatto che tali sistemi possono allora produrre sul principio di *accuracy* nel *risk assessment* riguarda anche il tasso di errore che si può verificare nel calcolo, proprio perché di natura statistica e non basato su una valutazione "individualizzata", allorché l'*output* indichi un "falso positivo", cioè colui che presenti un alto rischio – ma erroneo – di commettere crimini, e un "falso negativo" che rientri erroneamente in un'alea di rischio basso<sup>71</sup>.

---

<sup>68</sup> Proprio la maggiore trasparenza che caratterizzerebbe i *risk assessment tools* afferenti ad organi istituzionali è stata alla base dell'auspicio dell'implementazione di sistemi di questa natura da parte della dottrina: in tema, D. Kehl, P. Guo, S. Kessler, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessment in Sentencing*, 2017, 28; A. Roth, *Machine Testimony*, in *Yale Law Review*, vol. 126, No. 7, 2017, pp. 2025-2026. In senso favorevole all'implementazione nel sistema giudiziario italiano di un sistema algoritmico di natura istituzionale, purché le regole sul suo funzionamento siano condivise da esponenti della magistratura e dell'ambito forense, A. Santosuosso, *Intelligenza artificiale e diritto*, Mondadori, 2020, pp. 91-93.

<sup>69</sup> Tale aspetto è sottolineato da A.B. Cyphert, *Reprogramming recidivism*, cit., p. 335, che tuttavia osserva che in termini di trasparenza maggiore sarebbe dovuto essere l'impegno del *DOJ* nel rilasciare informazioni circa il dataset utilizzato per programmare l'algoritmo.

<sup>70</sup> *Idem*, pp. 359 ss.

<sup>71</sup> Cfr. B.L. Garrett, J. Monahan, *Judging Risk*, cit., p. 450.

## 5. Gli orientamenti della giurisprudenza statunitense e le soluzioni prospettate dalla dottrina per un uso “costituzionalmente orientato” degli *AI-based risk assessment tools*.

Guardando alla giurisprudenza statunitense, secondo l’orientamento maggiormente consolidato è da escludersi una violazione dei corollari del *due process* determinata dal ricorso a sistemi algoritmici di *risk assessment*.

La celebre sentenza sul caso *State v. Loomis*<sup>72</sup> emessa nel 2016 dalla *Wisconsin Supreme Court*, chiamata a pronunciarsi in ordine all’asserita violazione dei principi del giusto processo qualora alla base del *sentencing* vi sia un *risk assessment* operato da COMPAS, esprime il tendenziale orientamento dei giudici d’Oltreoceano sull’utilizzo dei *risk assessment tools*: pur se le valutazioni di rischio vengono effettuate su base collettiva e non individuale, e sebbene il funzionamento di tali *software* sia coperto dal segreto industriale, non vi sarebbe alcuna violazione dei principi del giusto processo, poiché gli *outputs* prodotti dall’algoritmo rappresentano solo uno degli indici di pericolosità dell’imputato.

Seguendo tale impostazione, dunque, non vi sarebbe alcuna violazione del giusto processo in quanto il risultato prodotto dall’algoritmo “si limita” a corroborare l’attendibilità di altri elementi sui quali si sia basata la decisione del giudice, lungi dall’assumere efficacia determinante nella valutazione.

Tra gli arresti più recenti, sempre in relazione all’utilizzo di COMPAS, vi è quello sul caso *People v. Younglove*<sup>73</sup>: nel rigettare l’impugnazione dell’imputato, la Corte d’appello del Michigan ha dichiarato che l’utilizzo del sistema COMPAS per il calcolo della recidiva non si poneva in contrasto con il principio del *due process*.

Anche in questo caso, il motivo di doglianza dell’appellante riguardava l’asserita incompatibilità di COMPAS con il diritto ad una sentenza “individualizzata”, basata sulle caratteristiche specifiche del *defendant*<sup>74</sup>; si adduceva inoltre la discriminazione generata da *outputs* basati su dati che prendevano in considerazione la razza e il genere<sup>75</sup> e l’inaccessibilità dei dati processati che avevano determinato l’*output* di rischio.

---

<sup>72</sup> *State v Loomis*, disponibile a [questo indirizzo](#).

<sup>73</sup> *People v Younglove*, disponibile a [questo indirizzo](#).

<sup>74</sup> «Because COMPAS statistically analyzes data from a general population in making its determinations (...) its use is inappropriate in an individualized sentencing decision (...) [COMPAS] has discriminatory impacts as concerns race and gender inputs, and that the scores it produces lack transparency»[«Poiché COMPAS analizza statisticamente dati che sono tratti da una generalità di persone (...) il suo utilizzo non è appropriato per una sentenza di condanna individualizzata (...) [COMPAS] ha effetti discriminatori per ciò che riguarda gli inputs di razza e di genere, e gli scores che genera sono carenti di trasparenza»].

<sup>75</sup> A supporto della sua impugnazione, l’appellante citava pubblicazioni scientifiche nelle quali la dottrina aveva espresso dubbi sull’attendibilità degli *outputs* prodotti da un *risk assessment tool*: J. Monahan, *A Jurisprudence of Risk Assessment: Forecasting Harm Among Prisoners, Predators, and Patients*, in *Virginia Law Review*, vol. 92, No. 3, 2006, p. 391; S.B. Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, in *Stanford Law Review*, vol. 66, No.4, 2014, pp. 804 ss. Sulle preoccupazioni sollevate dalla dottrina in ordine all’affidamento che i giudici ripongono in COMPAS, ritenendolo particolarmente attendibile ed accurato nei suoi risultati, si veda E. Yong, *A Popular Algorithm is No Better at Predicting Crimes than Random People*, in *The Atlantic*, 2018, p. 17, anch’esso citato nella sentenza della Corte d’Appello. L’Autore osserva come il sistema sia considerato particolarmente accurato nei risultati ma che, secondo recenti studi

Rilevando che l'imputato non aveva assolto all'onere della prova di dimostrare l'erroneità della valutazione e non aveva specificato quali fossero i diritti fondamentali violati dal ricorso al COMPAS, onere che grava sull'appellante, la Corte d'Appello ha concluso dichiarando l'insussistenza di violazioni del *due process* nel caso di specie. Data la segretezza delle regole che governano il funzionamento dei sistemi di *risk assessment*, *People v. Younglove* rappresenta dunque l'emblema della *probatio diabolica* che ricade sull'imputato che venga "giudicato" da un algoritmo.

Sull'opportunità di ricorrere ai *risk assessment tools* per la dosimetria della pena, poi, secondo un orientamento giurisprudenziale i risultati prodotti dai sistemi algoritmici possono invero migliorare la valutazione individualizzata a fondamento della condanna da parte del giudice e favorire l'individuazione del trattamento sanzionatorio più appropriato per il reo: è quanto affermato in *Malenchik v. State*<sup>76</sup> dall'*Indiana Supreme Court*.

Di diverso avviso era stata, solo due anni prima, l'*Indiana Court of Appeal* in *Rhodes v. State*<sup>77</sup>, che aveva accolto l'impugnazione dell'appellante circa l'asserito illegittimo ricorso al *Level of Service Inventory-Revised (LSI-R) risk assessment* nel *sentencing* per la commisurazione della pena. Nel caso di specie, lo *score* risultante dall'elaborazione dell'LSI-R aveva determinato l'applicazione di una circostanza aggravante a carico dell'imputato, il quale rivendicava il diritto ad un trattamento sanzionatorio "individualizzato" e non basato su modelli di punteggio standardizzato. Dopo aver rilevato che le sentenze di condanna emesse in primo grado sono soggette a revisione solo se, pur irrogando una pena ricompresa entro i limiti edittali, sia ravvisabile un abuso di discrezionalità da parte del tribunale, in *Rhodes v. State* la Corte d'appello dell'Indiana riconosceva che «l'uso di un modello di punteggio standardizzato, come il LSI-R, sminuisce la responsabilità del tribunale del processo di emanare un'appropriata sentenza individualizzata» e che aveva costituito «un abuso di discrezionalità affidarsi a modelli di punteggio per determinare una condanna»<sup>78</sup>.

A fronte dunque dell'impostazione seguita da parte di quella giurisprudenza statunitense riluttante nel riconoscere quella che sembra invece essere un'ineludibile compromissione delle garanzie costituzionali dell'imputato "giudicato" pericoloso o potenzialmente recidivo da un algoritmo tradizionale, la dottrina più sensibile al rispetto dei corollari del giusto processo è già andata oltre, proiettandosi sulle future ulteriori

---

di ricercatori del Dartmouth College, tale accuratezza nella previsione del rischio risulta invero solo maggiore di quella che assicurerebbe quella effettuata da un gruppo di persone casualmente interpellate su internet: «COMPAS is no better at predicting an individual's risk of recidivism than random volunteers recruited from the internet». Per approfondimenti su tale indagine e sui risultati emersi, si vedano J. Dressel, H. Farid, *The accuracy, fairness, and limits of predicting recidivism*, in *Science Advances*, 2018, p. 1 ss., disponibile a [questo indirizzo](#). In senso diverso, A.M. Holsinger et al., *A Rejoinder to Dressel and Farid*, cit.

<sup>76</sup> *Anthony Malenchik v. State of Indiana*, disponibile a [questo indirizzo](#).

<sup>77</sup> *Rhodes v. State*, disponibile a [questo indirizzo](#).

<sup>78</sup> «The use of a standardized scoring model, such as the LSI-R, undercuts the trial court's responsibility to craft an appropriate, individualized sentence [...]. Therefore, it is an abuse of discretion to rely on scoring models to determine a sentence». Il riferimento all'abuso di discrezionalità è contenuto anche nella citata sentenza della *Wisconsin Supreme Court* sul caso *Loomis*, ma in quella circostanza i giudici hanno ritenuto che l'utilizzo del sistema COMPAS nel *sentencing* non aveva integrato gli estremi di un abuso di discrezionalità nell'esercizio del potere giurisdizionale.

problematiche che saranno generate dai *risk assessment tools* di ultima generazione che sfrutteranno le potenzialità dell'AI<sup>79</sup>.

Nel tentativo di non rinunciare aprioristicamente a questi sistemi, sono così stati tracciati alcuni principi-guida che dovrebbero ispirare il ricorso a tali strumenti nel rispetto dei diritti costituzionalmente sanciti dalla *Due Process Clause*.

I criteri elaborati muovono dall'esigenza di bilanciare la tutela del segreto industriale con il principio di trasparenza (*transparency*), e si estendono alle garanzie di verificabilità (*auditability*) e di coerenza (*consistency*) dei dati utilizzati e degli *outputs* prodotti.

Sotto l'aspetto della verificabilità della decisione algoritmica – obiettivo che nel sistema penale italiano potrebbe essere raggiungibile con l'ostensione della motivazione nel provvedimento giurisdizionale, onere che dovrebbe essere "rafforzato" nel caso di scelta dell'opzione ermeneutica più sfavorevole per il *defendant*<sup>80</sup> – nella dottrina statunitense il *focus* è posto su soluzioni che permettano di "cristallizzare" il risultato fornito dall'algoritmo e di raccogliere e conservare questi dati a tutela del diritto di difesa dell'imputato che potrà così accedervi in un momento successivo. Si evidenzia, tuttavia, l'ardua scelta sull'individuazione di "chi" dovrà espletare tale compito e di "come" verranno raccolte le informazioni, snodi che saranno cruciali per un'effettiva "comprensione" della decisione algoritmica da parte di chi dovrà impugnarla.

Il principio della coerenza della decisione punta poi ad assicurare l'identità dell'*output* nel corso del tempo: l'obiettivo è quello di garantire che il sistema che ripeta l'operazione in un momento cronologicamente successivo utilizzando i medesimi *inputs*, generi sempre lo stesso *output*, sebbene nel frattempo l'algoritmo si sia evoluto grazie al *machine learning*. Proprio la capacità di autoapprendimento dell'algoritmo, infatti, determina il rischio che uno stesso *dataset* utilizzato più volte dal *tool* per il calcolo del rischio generi una varietà di *risk scores* dissimili. La diversità delle valutazioni generate nel corso del tempo, pur se basate sempre sugli stessi *inputs*, dipenderebbe così unicamente dal "caso", ovvero dal momento nel quale il calcolo viene effettuato. Trasposto nel sistema penale italiano, un algoritmo di AI di calcolo della pericolosità o della recidiva si porrebbe in evidente contrasto anche con il principio di uguaglianza di cui all'art. 3 Cost.

Forse nell'ordinamento giuridico italiano non sono ancora maturi i tempi per dare ingresso ai *predictive algorithms* nella valutazione della pericolosità criminale di un imputato in funzione di supporto della decisione del giudice umano, ma l'esperienza statunitense sollecita riflessioni sui futuri scenari che potranno profilarsi all'orizzonte del legislatore, di fronte ai quali quest'ultimo non potrà farsi trovare impreparato. Solo con un

---

<sup>79</sup> J. Villasenor, V. Foggo, *Artificial Intelligence*, cit., pp. 339 ss., ai quali si rinvia per una disamina delle criticità degli *AI-based risk assessment tools* e delle possibili soluzioni volte a rendere tali sistemi compatibili con i principi costituzionali del *due process*.

<sup>80</sup> Sull'onere motivazionale della sentenza emessa sulla base dell'*output* prodotto dall'algoritmo e, nel caso di norme dalla dubbia portata, sulla motivazione "confutativa" con cui i giudici che optassero per l'opzione interpretativa più sfavorevole all'imputato dovrebbero esplicitare le ragioni per le quali risulta implausibile l'interpretazione a quest'ultimo più favorevole, si veda M. Caterini, *Il giudice penale robot*, cit., 15 ss. Sulla motivazione quale strumento per garantire la "controllabilità" della decisione algoritmica, v. anche V. Manes, *L'oracolo algoritmico e la giustizia penale*, in *DisCrimen*, 2020, pp. 20-22.

utilizzo dei sistemi algoritmici di *risk assessment* normativamente disciplinato e soprattutto “costituzionalmente orientato”, si potrà affermare che i diritti di libertà e di difesa celebrati dalla Costituzione non hanno ceduto dinanzi all’inesorabile progresso informatico.

## Bibliografia

### Saggi, articoli e monografie

M. Amisano, *Profetica-mente: neuroscienze, intelligenza artificiale e previsione*, in F. Basile,

R. Anspach, *How the First Step Act Got People Out of Prison and Back With Their Families*, dicembre 2019.

J.J. Avery, *An Uneasy Dance with Data: Racial Bias in Criminal Law*, in *Southern California Law Review Postscript*, vol.93, No. 28, 2020, pp. 28 ss.

S. Barocas, A.D. Selbst, *Big Data’s Disparate Impact*, in *California Law Review*, vol. 104, No. 3, 2016, pp. 671 ss.

F. Basile, *Intelligenza artificiale e diritto penale: qualche aggiornamento e qualche nuova riflessione*, in F. Basile, M. Caterini, S. Romano (a cura di), *Il sistema penale ai confini delle hard sciences. Percorsi epistemologici tra neuro-scienze e intelligenza artificiale*, Pacini Giuridica, 2021, pp. 11 ss.

F. Basile, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in questa rivista, 29 settembre 2019, pp. 1 ss.

William D. Burrell, *Risk and Needs Assessment in Probation and Parole: The Persistent Gap Between Promise and Practice*, in *Handbook on Risk and Need Assessment: Theory and Practice*, Faye S. Taxman, ed., 2017.

M. Caterini, *Il giudice penale robot*, in *La legislazione penale*, 2020, pp. 1 ss.

M. Caterini, S. Romano (a cura di), *Il sistema penale ai confini delle hard sciences. Percorsi epistemologici tra neuro-scienze e intelligenza artificiale*, Pacini Giuridica, 2021, pp. 131 ss.

A.B. Cyphert, *Reprogramming Recidivism: The First Step Act and Algorithmic Prediction of Risk*, in *Seton Hall Law Review*, vol. 51, 2020, pp. 331 ss.

A. B. Cyphert, *Tinker-ing with Machine Learning: The Legality and Consequences of Online Surveillance of Students*, in *Nevada Law Journal*, vol. 20, No. 2, 2020, pp.457 ss.

J. Dressel, H. Farid, *The accuracy, fairness, and limits of predicting recidivism*, in *Science Advances*, 2018.

B.L. Garrett, J. Monahan, *Judging Risk*, in *California Law Review*, vol. 108, 2020, pp. 439 ss.

M. Gialuz, *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa*, in *Dir. pen. cont.*, 29 maggio 2019, pp. 1 ss.

A.M. Holsinger et al., *A Rejoinder to Dressel and Farid: New Study Finds Computer Algorithm is More Accurate than Humans at Predicting Arrest and as Good as a Group of 20 Lay Experts*, in *Federal Probation Journal*, vol. 82, No. 2, 2018, pp. 50 ss.

C.N. Jones, *A broken PATTERN: A Look at the Flawed Risk and Needs Assessment Tool of the First Step Act*, in *Howard Human & Civil Rights Law Review*, vol. 5, No. 2, 2021, pp. 185 ss.

D. Kehl, P. Guo, S. Kessler, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessment in Sentencing*, in *Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School*, 2017, pp. 1 ss.

B. Kutateladze et al., *Cumulative Disadvantage: Examining Racial and Ethnic Disparity in Prosecution and Sentencing*, in *Criminology: An Interdisciplinary Journal*, vol. 52, No.3, 2014, pp. 514 ss.

V. Manes, *L'oracolo algoritmico e la giustizia penale*, in *DisCrimen*, 2020, pp. 1 ss.

S.G. Mayson, *Bias In, Bias Out*, in *The Yale Law Journal*, vol. 128, No. 8, 2019, pp. 2218 ss.

T.A. Merkl, L. Arzy, *California's Referendum to Eliminate Cash Bail, Explained*, *Brennan Center for Justice*, 2 ottobre 2020.

J. Monahan, *A Jurisprudence of Risk Assessment: Forecasting Harm Among Prisoners, Predators, and Patients*, in *Virginia Law Review*, vol. 92, No.3, 2006, pp. 391 ss.

A.S. Niederman et al., *The Institutional Life of Algorithmic Risk Assessment*, in *Berkeley Technology Law Journal*, vol. 34, No. 3, 2019, pp. 705 ss.

C. O'Neil, *Weapons of Math Destruction. How Big Data Increases Inequality and threatens Democracy*, Penguin Books Ltd., New York, 2016.

S. Quattrocolo, *Equo processo penale e sfide della società algoritmica*, in *BioLaw Journal*, n. 1, 2019, pp. 135 ss.

A. Rizer, C. Watney, *Artificial Intelligence Can Make Our Jail System More Efficient, Equitable, and Just*, in *Texas Review of Law & Politics*, vol. 23, No. 1, 2019, pp. 181 ss.

A. Roth, *Machine Testimony*, in *Yale Law Review*, vol. 126, No.7, 2017, pp. 1972 ss.

A. Santosuosso, *Intelligenza artificiale e diritto*, Mondadori, 2020.

N. Scurich, *An Introduction to the Assessment of Violence Risk*, in *International Perspectives on Violence Risk Assessment*, 3, 10, J P. Singh, S. Bjørkly, & S. Fazel eds., 2016.

A. Slobogin, *Preventive Justice: A Paradigm in Need of Testing*, in *Behavioral Sciences & the Law*, Vol. 36, No. 4, 2018, pp. 391 ss.

T. Sourdin, *Judges, Technology and Artificial Intelligence. The Artificial Judge*, Elgar, 2021.

S.B. Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, in *Stanford Law Review*, vol. 66, No. 4, pp. 803 ss.

M. Stevenson, *Assessing Risk Assessment in Action*, *Minnesota Law Review*, vol. 103, No. 58, 2018, pp. 303 ss.

C.D. Stimson, *The First Step Act's Risk and Needs Assessment Program: A Work in Progress*, in *Heritage Foundation Legal Memorandum, No.265, 2020*.

J. Villasenor, V. Foggo, *Artificial Intelligence, Due Process and Criminal Sentencing*, in *Michigan State Law Review*, vol. 2020, No. 2, 2020, pp. 295 ss.

R. Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, in *Stanford Law Review*, vol. 70, No. 5, 2018, pp. 1343 ss.

J.S. Wormith, *Automated Offender Risk Assessment: The Next Generation or a Black Hole?*, in *American Society of Criminology*, vol. 16, No. 1, 2017, pp. 281 ss.

E. Yong, *A Popular Algorithm is No Better at Predicting Crimes than Random People*, in *The Atlantic*, 2018.

M. Yu, G. Du, *Why Are Chinese Courts Turning to AI?*, in *The Diplomat*, 2019.

## Documenti istituzionali

DOJ, *2020 Review and Revalidation of the First Step Act Risk Assessment Tool*, 2021.

DOJ, Office of the Attorney Gen., *The First Step Act of 2018: Risk and Needs Assessment System– Update*, 2020.

DOJ, *Department of Justice Announces Enhancements to the Risk Assessment System and Updates on First Step Act Implementation*, 15 gennaio 2020.

DOJ, *Office of the Attorney General, Memorandum for Director of Bureau Prisons*, 2020.

DOJ, Office of the Attorney Gen., *The First Step Act of 2018: Risk and Needs Assessment System*, 2019.

## Altri documenti

Brennan Center's Public Comment on the First Step Act's Risk and Needs Assessment Tool, 2019.

*Comment Letter to the Department of Justice on PATTERN First Step Act*, 2019.

*Oversight Hearing on "The Federal Bureau of Prisons and Implementation of The First Step Act"*, 2019.

*Stakeholder Statements Submitted to NIJ's September Listening Sessions: Comments on the Department of Justice's Report, The First Step Act of 2018: Risk and Needs Assessment System*, 2019.

H. Harris, J. Goss, A. Gumbs, *Pretrial risk assessment in California*, *Public Policy Institute of California*, 2019.

*Stakeholder Statements Submitted to NIJ's September Listening Sessions: Comments on the Department of Justice's Report, The First Step Act of 2018: Risk and Needs Assessment System*, 2019.

## Fonti normative

*California Money Bail Reform Act (SB 10)*, 2018.

*H.R. 5682-First Step Act*, 2018

*First Step Act (Public Law 115-391)*, 2018.

## Riferimenti giurisprudenziali

*In re Humphrey (S247278), Supreme Court of California (2021)*.

*People v Younglove* (2019).

*Humphrey (Kenneth) on Habeas Corpus (A152056), First Appellate District of San Francisco City* (2018).

*State v Loomis* (2016).

*Anthony Malenchik v. State of Indiana* (2010).

*Rhodes v. State* (2008).