Processo penale e rivoluzione digitale: da ossimoro a endiadi?*

Serena Quattrocolo

Abstract

Queste brevi riflessioni mettono a fuoco i principali aspetti problematici, fino ad ora emersi, dell'impatto della rivoluzione digitale sulla sfera del processo penale. In particolare, tre sono gli aspetti qui segnalati. Gli effetti della rivoluzione digitale sono intanto considerati sotto il profilo dell'attività investigativa di ricerca della prova, con particolare riguardo allo sfruttamento dell'enorme potenziale intrusivo di certi strumenti digitali. In secondo luogo, l'attenzione si porta sull'impiego processuale di dati generati automaticamente - attraverso algoritmi e, più in generale, modelli computazionali – il cui vaglio di attendibilità si scontra con il tradizionale diritto probatorio. Da ultimo, viene considerata l'ampia gamma di rischi insiti nell'uso di modelli computazionali di ausilio alla decisone giurisdizionale, in qualsiasi fase del procedimento essa sia adottata.

The paper focuses on the main issues related to the impact of the digital turn and the use of algorithms and computational models in the realm of criminal proceedings. The analysis encompasses three main topics: the impact of digital technologies as means to intrude individuals' privacy, in gathering evidence; the use of algorithm-generated data, used as evidence in trials; the compliance of computational models, as instruments supporting the decision-making process, with fundamental rights.

Sommario

Keywords

algoritmi - modelli computazionali - procedimento penale - equo processo - prova

* Su determinazione della direzione, in conformità all'art. 15 del regolamento della Rivista, l'articolo èstato sottoposto a referaggio anonimo

1. Qualche cenno introduttivo

Nel ricco ciclo di seminari organizzati per gli studenti dell'Università del Piemonte orientale ha trovato spazio, grazie alla sensibilità degli organizzatori, anche un profilo che, solo recentemente, sta conquistando spazio nella riflessione degli studiosi e nell'elaborazione della dottrina. Come sempre accade, i riflessi dei più profondi mutamenti della società si proiettano sulla giustizia penale con un significativo ritardo, ovvero quando quei mutamenti possono considerarsi ormai sedimentati nel sentire comune, tanto da arrivare ad incidere sulla sfera del diritto cui è demandato il presidio degli interessi più essenziali, attraverso la pena[1]. È certamente tempo di riconoscere, infatti, che la rivoluzione digitale in corso da alcuni decenni e fortemente accentuatasi proprio negli ultimi due lustri, pone la giustizia penale e, in primo luogo, il processo penale di fronte ad un cambiamento profondo, sia degli attori sociali, sia degli strumenti con i quali essi operano. È frequente – ed icastico – parlare di "società algoritmica", per riferirsi a quell'ampio fenomeno[2] che coinvolge individui e soluzioni tecnologiche nell'elemento che maggiormente distingue e contraddistingue l'odierna realtà: l'ipertrofica produzione di dati generati automaticamente – per lo più al di fuori del controllo di un agente umano - i quali possono essere impiegati con le più varie finalità, addirittura, appunto, all'interno del procedimento penale[3]. È impossibile ricostruire sinteticamente l'ampio dibattito filosofico sviluppatosi in questi ultimi anni, soprattutto attorno alla teoria di Luciano Floridi, che fotografa il percorso dell'evoluzione umana dalla preistoria alla c.d iper-storia[4]. Un cammino lungo e scandito da periodi storici assai eterogeni, che passa attraverso l'invenzione della stampa per arrivare a certificare l'affermazione - oltre le teorie classiche, westfaliane e poi montesquieuiane, del potere – dell'odierno "potere computazionale", con la sfera di ricadute, ancora per lo più inesplorate, che esso proietta sulla società, sul sapere e, ovviamente, sul diritto[5]. La locuzione "algoritmo", correlata al termine "società", può poi assumere una varietà di significati a seconda del contesto in cui viene utilizzato, con sfumature e variazioni anche considerevoli, spesso non condivise dagli stessi esperti del medesimo settore. Ai limitati fini di queste brevi riflessioni, si prenderanno le mosse dalla definizione offerta da Tarleton Gillespie, nel 2014, che è stata altresì assunta come paradigma dal prezioso studio già pubblicato dal Consiglio d'Europa, *Algorithms and Human Rights* nel dicembre 2017[6]. L'autore afferma: «*algorithms need not be software: in the broadest sense, they are encoded procedures for transforming input data into a desired output, based on specified calcula-*

[1] S. Quattrocolo, *Artificial Intelligence, Computational Modelling and Criminal Proceedings. A Framework for a European Legal Discussion*, New York, 2020, 3.

[2] Si veda L. Floridi et alii, *AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations*, in *Minds and Machines*, 2018, 689 ss.

[3] Interessante la lettura proposta da A. Garapon – J. Lassègue, *Justice digitale*, Parigi, 2018, 9 ss., che vede nel digitale una rivoluzione grafica la quale – sulla scia di quelle che in precedenza hanno segnato la storia, come ad esempio il comparire dell'alfabeto greco – sta producendo un impatto epocale sulla comunicazione e sui suoi riflessi.

[4] L. Floridi, *The Fourth Revolution*, Oxford, 2017, *passim*.

[5] M. Durante, *Potere computazionale*, Milano, 2019.

[6] Reperibile alla pagina https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5

tions. The procedures name both a problem and the steps by which it should be solved»[7]. Tali "*encoded procedures*" presuppongono la realizzazione e l'impiego di un modello computazionale, che riproduce un fenomeno, prendendo in considerazione tutte le variabili rilevanti e regolandone l'interazione.

Allo stesso modo, anche la locuzione "intelligenza artificiale", spesso usata in queste pagine, assume significati molto vari, a seconda del contesto in cui viene utilizzata. Per le finalità di questo lavoro, la più adatta pare quella contenuta nel EC JRC[8] report sull'intelligenza artificiale, ove si afferma che «"intelligenza artificiale" è un termine generico che si riferisce ad ogni macchina o algoritmo in grado di osservare l'ambiente, imparare e, sulla base dell'apprendimento e delle esperienze pregresse, assumere comportamenti intelligenti o proporre decisioni»[9].

Alla luce di queste due definizioni si può affermare che il fenomeno che stiamo vivendo e che stiamo cercando di analizzare non è, necessariamente, l'affermarsi di una società che delega la scelta alla macchina – come una certa visione distopica spesso suggerisce – ma che, a fronte della discrezionalità insita nell'intuito del singolo, ritiene utile circoscrivere i processi decisori, anche quello giudiziario, in una relazione prestabilita, rappresentata in un modello computazionale, affiancando i risultati di questa alla tradizionale attività umana[10].

In questo contesto, anche la sfera della giustizia penale deve misurarsi con l'effetto del vasto impiego di algoritmi e di intelligenza artificiale in software anche di uso quotidiano, non espressamente sviluppati per essere utilizzati in tale ambito. Infatti, la rivoluzione digitale ha fatto segnare, nell'ultimo decennio, un clamoroso balzo, alimentato da due fattori principali[11]: la diffusione globale di smartphones e altri strumenti di comunicazione telematica che generano quotidianamente, e in modo gratuito, quantità incommensurabili di dati; un aumento esponenziale della capacità computazionale, che consente di processare con tempi e costi assai ridotti rispetto ad alcuni anni fa, quella massa pressoché infinita di dati.

Come anticipato, si possono mettere a fuoco tre aree nelle quali l'impiego di dati generati automaticamente (non necessariamente personali e non necessariamente sensibili), elaborati attraverso modelli computazionali, per finalità legate al processo penale, rischia di porsi in contrapposizione con i principi enunciati nella Costituzione italiana e nelle carte internazionali, a garanzia, tanto di diritti e libertà strettamente attinenti alla sfera personale - come la riservatezza – la cui violazione può essere perpetrata attraverso atti del procedimento penale, quanto, più specificamente, dell'equità del processo penale stesso. Il primo ambito è quello investigativo - dei mezzi di ricerca della prova - nel quale il proliferare di forme di comunicazione digitale ha aperto squarci sempre maggiori di vulnerabilità della riservatezza personale: più informazioni, più

---

[7] T. Gillespie, *The relevance of Algorithms*, in T. Gillespie - P. Boczkowski - K. Foot (eds.), *Media Technologies*, Cambridge US, 2014, 167.

[8] Joint Research Center, presso il servizio Scienza e Conoscenza della Commissione europea.

[9] M. Craglia, *Artificial Intelligence: a European Perspective*. EU Publication Office, Luxembourg, 2018.

[10] Cfr. M. Durante, *Potere computazionale,* cit., 231 ss.

[11] U. Pagallo – M. Durante, *The Philosophy of Law in an Information Society*, in L. Floridi (ed.), *The Routledge Handbook of Philosophy of Information*, New York, 2016, 396 ss.

strumenti di intrusione sono i fattori di un'operazione aritmetica dalle conseguenze impressionanti[12], come dimostrato anche dal recente attivarsi del legislatore processuale penale italiano. Il secondo ambito è quello più propriamente probatorio, investito dall'afflusso di dati generati in maniera automatizzata, fuori dal processo, con scarse se non inesistenti possibilità di verifica processuale della loro attendibilità. Il terzo è quello dell'impiego di modelli computazionali per assistere i diversi soggetti del processo penale nell'assunzione di scelte o nell'effettuazione di valutazioni, sulla base di ricchi o addirittura completi data-base, analizzati attraverso software capaci di stabilire, in tale massa di informazioni, correlazioni e rispondenze.

Per ragioni di coerenza, l'attenzione qui è focalizzata sul procedimento penale e, dunque, sul reato consumato (o tentato, naturalmente) e non sulle considerevoli applicazioni dei modelli computazionali in ambito di predizione e prevenzione del reato[13]. Tale ambito costituisce, ormai, un settore di studio autonomo che, pur ricorrendo alla modellizzazione matematica e all'efficienza dell'intelligenza artificiale, si fonda su considerazioni del tutto extra-giuridiche, legate allo studio dei fenomeni criminosi e dei contesti sociali. Esso rappresenta, dunque, uno scenario estraneo o, quantomeno, precedente a quello del procedimento penale, che si instaura a seguito della commissione del reato.

2. Rivoluzione digitale, investigazione penale e riservatezza

Il tema dell'impiego di software capaci di carpire segretamente informazioni e dati a fini investigativi è ampio e articolato. L'appena ricordata spinta della rivoluzione digitale ha inciso significativamente sulle modalità di investigazione, sempre più massicciamente fondate sull'*hacking*, l'accesso occulto a sistemi di produzione o elaborazione di dati digitali. Le brevi considerazioni che seguono sono basilari e non approfondiscono l'argomento, che meriterebbe un'ampia trattazione autonoma.

L'estrazione, occulta, di informazioni contenute in dati generati automaticamente è divenuta un irrinunciabile strumento investigativo. La gamma di azioni intrusive che possono essere realizzate, ad esempio, attraverso *malwares*, inoculati da remoto nei dispositivi *hardware,* è considerevole. Per *malwares* si intendono vari tipi di *malicious software*, un'ampia gamma di captatori informatici che possono accedere a molteplici funzioni degli apparati digitali in cui vengono inseriti, nascosti all'interno di files o di applicativi apparentemente innocui (c.d. *trojan horses*). Invisibile all'utente che ha in uso l'apparecchio infettato, il *malware* consente varie forme di intrusione nella sfera digitale dell'interessato e, in particolare: a) acquisizione di informazioni scambiate attraverso il mezzo infettato; b) attivazione da remoto di strumenti di geolocalizzazione, ripresa o

[12] Si veda, sul recente caso di cronaca legato al malware Exodus, F. Palmiotto, *Captatori informatici e diritto alla* difesa. Il caso Exodus, in *lalegislazionepenale.eu*, 16.10.2020.

[13] Per un'interessante sintesi, v. F. Basile, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Diritto Penale Uomo,* 2019, C. Costanzi, *Big data e garantismo digitale. Le nuove frontiere della giustizia penale nel XXI secolo*, in *lalegislazionepenale.eu,* 21 dicembre 2019.

registrazione audio; c) accesso e manipolazione dei *files* presenti nell'*hardware* infetto[14]. Si tratta di strumenti che vantano una evidente potenzialità investigativa che è stata prontamente sfruttata dagli organi inquirenti, in tutto il mondo, per lo più in assenza di un apposito quadro normativo. Per un verso, infatti, gli operatori e la prima giurisprudenza si sono mossi entro i margini della disciplina esistente, per lo più in tutti gli ordinamenti, per la regolamentazione delle intercettazioni di comunicazioni – ambientali, telefoniche e telematiche – partendo dal (discutibile) presupposto che i *malwares* altro non siano che nuove modalità tecniche di svolgimento di tradizionali mezzi di ricerca della prova, appunto. Per altro verso, le autorità giudiziarie hanno consapevolmente minimizzato l'evidente divario, in termini di intrusività, che i captatori informatici presentano rispetto ai tradizionali strumenti intercettivi[15]. Sul punto, nell'impossibilità di proporre in questa sede un approccio più approfondito, si devono sviluppare due riflessioni. In primo luogo, rinviando all'esaustivo rapporto commissionato dal LIBE Committee del Parlamento europeo in materia di *hacking by law enforcement*[16], dove i profili di tale superiore insidiosità per la sfera di riservatezza degli individui sono dettagliatamente trattati[17], va sottolineata la necessità di una disciplina normativa apposita dei captatori telematici, che non possono essere regolati secondo il paradigma delle tradizionali intercettazioni. Per quanto riguarda l'ordinamento italiano, la situazione normativa è decisamente fluida, poiché a seguito della riforma delegata dalla "legge Orlando" al Governo, sulla materia delle intercettazioni di comunicazioni e sull'impiego dei captatori informatici nel procedimento penale si è inserita una cospicua decretazione d'urgenza, prima con il d.l. 161/2019 (conv. con mod. in l. 7/2020) e, da ultimo, nel contesto dell'emergenza sanitaria, con il d.l. 28/2020, conv. con mod. in l. 70/2020),[18] che hanno segnato una progressiva estensione dell'impiego – pur appositamente regolamentato – del captatore informatico.

In secondo luogo, appare evidente la perdita di significato che i concetti ai quali è ancorata, nel linguaggio costituzionale nazionale e nelle carte internazionali, la tutela della riservatezza, subiscono di fronte a mezzi di ricerca della prova così intrusivi. Domicilio e corrispondenza, architravi delle garanzie costituzionali contro le interferenze statuali, anche investigative, nella sfera personale degli individui, hanno perso signifi-

[14] In argomento, *ex multis*, M. Torre, *Il Captatore informatico*, Milano, 2017, spec. 12-17; M. Pittiruti, *Digital Evidence e processo penale*, Torino, 2017, 69 ss.; S. Signorato, *Le indagini digitali*, Torino, 2018, 237 ss.

[15] In questo senso, M. Daniele, *La prova digitale processo penale*, in *Riv. Dir. Proc.*, 2011, 288: «La loro capacità lesiva della *privacy* è addirittura superiore a quella delle intercettazioni».

[16] Studio commissionato dal Libe Committee del Parlamento europeo e realizzato dal Directorate-General for Internal Policies, *Legal Frameworks for hacking by Law Enforcement: Identification, Evaluation and Comparison of Practices* (reperibile alla pagina *europarl.europa.eu*).

[17] A p. 21 si afferma: «*although the use of hacking techniques will bring improvements in investigative effectiveness, the significant amount and sensitivity of data that can be accessed through these means acts as a stimulus for another key debate: ensuring the protection of the fundamental right to privacy*».

[18] Complesso riassumere brevemente la vicenda italiana. A seguito di una nota decisione delle Sezioni Unite della Corte di cassazione (Cass. pen., sez. un., 1° luglio 2016, n. 26889) un documento sottoscritto da quasi tutti i docenti italiani di diritto processuale penale aveva sollecitato la necessità di uno specifico intervento normativo (v. *penalecontemporaneo.it*, 16 ottobre 2016), avvenuto poi con il d.lgs. 216/2017 (e d.m. 20 aprile 2018), che ha inserito nel codice di procedura penale una specifica disciplina del captatore informatico, ad oggi non ancora entrata in vigore. I recenti interventi hanno poi sempre procrastinato l'entrata in vigore della nuova disciplina, avvenuta nel settembre 2020.

cato, di fronte alla possibilità di produrre, scambiare, conservare – ma anche carpire, intercettare, copiare – dati immateriali in uno spazio che non è più quello fisico. La più o meno consapevole accettazione, generalizzata, di apparecchi digitali che, per le loro ridotte dimensioni, seguono l'individuo ovunque e sempre, rende decisamente impossibile continuare ad applicare la tradizionale distinzione tra luoghi pubblici, luoghi aperti al pubblico e luoghi privati, come il domicilio, nel quale la captazione occulta è consentita solo in via eccezionale, a condizioni ancora più stringenti di quelle previste in via generale[19].

A fronte di un'interferenza, destabilizzante, tra strumenti digitali basati su modelli computazionali e valori essenziali del nostro patrimonio giuridico, il quadro delle garanzie fondamentali, sancite dalla Convenzione europea dei diritti dell'uomo e dalla Costituzione italiana, rappresenta ancora, certamente, la cornice normativa di riferimento. Per un verso, nell'art. 8 CEDU, la Corte di Strasburgo ha individuato dei limiti ben precisi anche all'attività investigativa di analisi e profilazione dei dati,[20] che possono rappresentare un utile parametro per gli ordinamenti nazionali. Per altro verso, la Convenzione stessa lascia intravedere, sullo sfondo, altri principi che possono rappresentare il criterio per stabilire (o ristabilire) i confini del concetto di *fairness* processuale anche nell'era della rivoluzione digitale. L'operazione, però, richiede, quantomeno, la disponibilità e la capacità di uscire dai paradigmi più familiari – i concetti di "domicilio", di "comunicazione", appunto, ma anche di "prova" e di "attendibilità" – per comprendere come essi siano stati riscritti nell'ultimo decennio, per riportarli poi, nell'alveo dei principi fondamentali della cultura giuridica europea.

3. La prova generata automaticamente e i rischi per la parità delle armi

Il secondo profilo di analisi riguarda i mezzi di prova. Se, poco sopra, l'attenzione si è soffermata sui nuovi mezzi di ricerca della prova, la realtà attuale dimostra come, anche senza l'impiego di strumenti di captazione occulta, da tutti i supporti digitali si possono estrarre informazioni di grande rilievo per il procedimento penale. Può trattarsi anche di metadati, che precisano condizioni oggettive riferite alla genesi del dato. Con la crescente rilevanza dell'IoT, può trattarsi di dati generati automaticamente, senza alcun intervento umano nella loro rilevazione, da oggetti di uso quotidiano collegati alla rete internet, come gli assistenti vocali di vario genere o gli elettrodomestici smart. Questi rappresentano, evidentemente, un patrimonio conoscitivo talvolta fondamentale per le indagini e per il procedimento penale: si pensi, ad esempio, al diffuso aspirapolvere Rumba, che immagazzina e conserva dati sui percorsi e sugli ingombri esistenti in ciascuna stanza della casa… L'insieme dei dati rilevati e conservati fornirà

[19] V. ampiamente, S. Signorato, *Le indagini digitali*, cit., 49 ss.

[20] Volendo, U. Pagallo - S. Quattrocolo, *The impact of AI on criminal law, and its twofold aspects*, in W. Barfield - U. Pagallo (eds.), *Research Handbook on the Law of Artificial Intelligence*, Cheltenham, 2018, 391. In generale, v. R. Sicurella - V. Scalia, *Data mining and profiling in the Area of Freedom, Security and Justice*, in *New Journal of European Criminal Law*, 2013, 409 ss.

informazioni determinati agli investigatori che indaghino, ad esempio, su un omicidio avvenuto in una specifica stanza…

La questione che qui si pone come oggetto centrale della riflessione riguarda la verifica dell'accuratezza del dato, generato e/o raccolto esclusivamente attraverso uno strumento digitale. È possibile contestarne l'attendibilità? Oppure la "prova digitale", per la sua natura e per la sua genesi, è oggettivamente impermeabile al confronto dialettico tra le parti nel processo? Tutti gli apparecchi digitali generano dati, attraverso un processo che, basato su algoritmi o, più in generale, modelli computazionali, difficilmente è trasparente: l'elemento conoscitivo che ne traiamo (perché no, a fini processuali?) è generato da un processo di cui possono non essere noti né gli input né i processi elaborativi… Una scatola nera, insomma, dalla quale si estrae qualcosa di utile, spesso, però, senza sapere come.

La metafora, calata nel contesto processuale, esprime l'estrema difficoltà o l'impossibilità di falsificare il dato elaborato da una "black box" se non è possibile accedere al codice sorgente che governa l'algoritmo stesso[21] o se, nonostante la disponibilità del codice sorgente, il processo che ha generato l'*output* non sia verificabile *ex post*. Tale scenario, che assumiamo al momento come valido, rischia di determinare una situazione di squilibrio conoscitivo estremo tra le parti del processo.

Invero, lo squilibrio conoscitivo è fenomeno che si riscontra nel processo penale sin da quando, per la soluzione di casi complessi, si è iniziato a fare ricorso a competenze tecniche, scientifiche o artistiche[22]. Tuttavia, l'ingresso di saperi specialistici nel processo difficilmente è equilibrato, poiché una delle parti - quella pubblica - ha accesso alla scienza e alle tecnologie migliori, anche perché dispone di mezzi economici non limitati. Evidentemente, il fenomeno di *knowledge impairment* non è nuovo e ogni stagione del complicato rapporto tra scienza e processo penale ne ha riproposta una versione più o meno intensa (si pensi al debutto della profilazione del DNA nelle aule di giustizia, o al ricorso alla fMRI per l'accertamento di profili legati all'imputabilità). La prova generata automaticamente, tuttavia, rischia di introdurre una forma estrema di tale squilibrio, poiché il risultato probatorio può essere non criticabile laddove, appunto, l'inaccessibilità del codice sorgente o altre caratteristiche del software non consentano alla parte contro la quale la prova è introdotta nel processo di contestarne l'accuratezza e l'attendibilità[23].

Il tema si innesta, evidentemente, sul cuore del diritto probatorio, le cui regole rappresentano le chiavi del cancello che regola l'accesso delle conoscenze al processo. Per quanto appaia difficile formulare considerazioni generali in materia di prova – geloso appannaggio delle legislazioni nazionali[24], assai restie all'armonizzazione su tale profilo

---

[21] U. Pagallo - S. Quattrocolo, *The impact of AI on criminal law*, cit., 392 ss.

[22] V. A.J. Brimicombe – P. Mungroo, *Algortihms in the Dock: Should Machine Learning Be Used in British Courts?,* Proceedings of the fourth Winchester Conference on Trust, Risk, Information and the Law, 3 maggio 2017.

[23] E. Van Buskirk-V.T. Liu, *Digital Evidence: Challenging the Presumption of reliability*, in *Journal of Digital Forensic Practice*, 1, 2006, 20; C. Chessman, *A Source of Error: Computer Code, Criminal Defendants, and the Constitution*, *Calif. L. Rev.*, 2017, 179 ss.

[24] Si veda l'esplicita affermazione (spesso reiterata) della Corte europea dei diritti dell'uomo in GC, *Gäfgen v. Germany,* ric. 22978/05 (2010), § 162: *«while Article 6 guarantees the right to a fair hearing, it does not*

– l'importanza della questione che si pone spinge a verificare l'esistenza di principi generali che possano guidare la gestione processuale del fenomeno delle prove generate automaticamente.

È noto che né a livello convenzionale, né nel più recente quadro delle direttive europee processuali penali emanate sulla base dell'art. 82, §2 TFUE, si reperisce traccia di un sistema di invalidità[25] della prova ispirato alla tradizione romano-germanica, in cui, a fronte della predisposizione normativa di uno schema legale dell'atto probatorio, si ritrova una sanzione processuale, che "neutralizza" la prova non corrispondente allo schema[26]. Nemmeno si reperisce, nel contesto europeo *lato sensu* inteso, una regola generale di esclusione probatoria dei "frutti dell'albero avvelenato", elaborata dalla letteratura e dalla giurisprudenza nord-americana[27]. Piuttosto, la Corte europea tende a convogliare tutte le valutazioni sull'ammissibilità e sulla utilizzabilità della prova in un generale test di compatibilità con il processo equo[28], inteso come nozione onnicomprensiva che calibra e combina le singole garanzie di dettaglio[29].

Ed è proprio in forza del generale canone del giusto processo e, più in particolare, del principio della parità delle armi, che anche l'ammissione e la valutazione di prove generate automaticamente pare porsi potenzialmente in contrasto con garanzie fondamentali del dettato convenzionale. È noto che, innanzitutto e pur nell'assenza di una esplicita enunciazione nel testo dell'art. 6 CEDU, il principio della parità delle armi è stato modellato dalla giurisprudenza della Corte come architrave, insieme al connesso canone del contraddittorio, dell'equità processuale nel suo complesso[30]. Notoriamente, *equality of arms* non implica una presunta, necessaria identità di facoltà o di posizioni

lay down any rules on the admissibility of evidence as such, which is primarily a matter for regulation under national law».

[25] R. Kostoris (ed.), *Handbook of European Criminal Procedure*, New York, 2018, 58.

[26] L. Bachmaier Winter, *The EU Directive on the Right to Access to a Lawyer: A Critical Assessment*, in S. Ruggeri (ed.), *Human Rights in European criminal Law*, Springer, 2015, 114; A. Cabiale, *I limiti alla prova nella procedura penale europea*, Padova, 2019, 311 ss., M. Caianiello, *To Sanction (or not to Sanction) Procedural Flaws at EU Level? A Step forward in the Creation of an EU Criminal Process*, in *European Journal of Crime, Criminal Law and Criminal Justice*, 2014, 317 ss., S. Quattrocolo*, Artificial Intelligence*, cit., 74 ss.

[27] Si veda l'interessante raffronto di S. Thaman*, 'Fruits of the Poisonous Tree' in Comparative Law*, in *Southwestern Journal of International Law*, 2010, 333 ss.

[28] Impossibile riassumere qui l'ampio ventaglio delle posizioni adottate della Corte in un percorso pluridecennale, nel quale l'approccio dei giudici di Strasburgo alla "tainted evidence" è significativamente mutato e non nel senso di un ampliamento delle garanzie dell'imputato. Si rimanda, dunque, in generale ad A. Cabiale, *I limiti*, cit., 87 ss. e, con specifico riguardo al tema qui trattato a S. Quattrocolo*, Artificial Intelligence*, cit., 77 ss.

[29] V. Manes - M. Caianiello, *Introduzione al diritto penale europeo*, Torino, 2020, 217 ss.

[30] Cfr. M. Chiavario, *Art. 6,* in S. Bartole - B. Conforti - G. Raimondi, *Commentario alla convenzione europea dei diritti dell'uomo e delle libertà fondamentali*, Padova, 2002, 192. Si tratta di un'acquisizione risalente nella giurisprudenza di Strasburgo, su cui v. già CEDU, *Neumeister v. Austria*, ric. 1936/63 (1968), § 22 delle motivazioni in diritto, che riconosce la parità delle armi come una caratteristica del dair trial, sulla base di numerose precedenti decisioni e opinioni della Commissione europea, allora incaricata di svolgere un filtro d'accesso alla Corte; CEDU, *Delcourt v. Belgium*, ric. 2689/65 (1970), § 28: *«The principle of equality of arms does not exhaust the contents of this paragraph; it is only one feature of the wider concept of fair trial by an independent and impartial tribunal»* e poi, successivamente, tra le tante, CEDU, *Brandstetter v. Austria*, ric. 11170/84 (1991), § 66; *Ruis-Mateos v. Spain*, ric. 12952/87 (1993), § 63; *Fitt v. UK*, ric. 29777/96 (2000), § 44; *Sabayev v. Russia*, ric. 11994/03 (2010), § 35; *J.M. e altri v. Austria*, ricc. 61503/14, 61673/14, and 64583/14 (2017), § 119.

di cui le parti essenziali del processo debbano sempre fruire, soprattutto laddove si tratti, appunto, di processo penale, il quale è caratterizzato – specialmente nelle sue fasi prodromiche – da un insuperabile squilibrio tra parte pubblica e difesa[31]. In questa connaturata differenza di ruoli istituzionali, il paradigma essenziale della parità delle armi è rappresentato dalla possibilità di presentare i propri argomenti in condizioni che non svantaggino una parte rispetto alle altre[32]. Insomma, il principio esprime nel suo nucleo essenziale e irrinunciabile, un giusto equilibrio tra le parti processuali[33]. Se indubbiamente tale affermazione può apparire per lo più declamatoria, essa va coniugata con più specifiche messe a punto della parità delle armi, come quella scolpita nel *leading case Brandstetter c. Austria*, in cui la Corte ha ribadito che è necessario che ciascuna parte abbia effettiva conoscenza delle allegazioni e delle argomentazioni della controparte e che fruisca della concreta possibilità di contestarle e falsificarle. «*An indirect and purely hypothetical possibility for an accused to comment on prosecution argument*»[34] non soddisfa il parametro convenzionale. All'interno del procedimento probatorio - ambito che la Corte riconosce, appunto, come devoluto alle discipline nazionali - è proprio la possibilità, per tutte le parti e, principalmente, per la difesa, di contestare l'accuratezza della prova a carico ad esprimere il senso proprio del suddetto giusto equilibrio. È stato, infatti, ripetutamente sottolineato che «*it must be examined in particular whether the applicant was given the opportunity of challenging the authenticity of the evidence and of opposing its use. In addition, the quality of the evidence must be taken into consideration, including whether the circumstances in which it was obtained cast doubt on its reliability or accuracy*»[35]. Ciò, invero, è funzionale a realizzare l'obiettivo intrinseco della parità delle armi, ossia consentire a tutte le parti le stesse *chances* di poter convincere il giudice della propria prospettazione dei fatti oggetto di prova[36].

Calando il tema delle prove raccolte e generate in via del tutto automatizzata all'interno del paradigma elaborato dalla Corte europea in lunghi anni di giurisprudenza, emerge un dato rilevante. L'eventuale impossibilità di accedere al codice sorgente o di poter effettivamente comprendere il funzionamento della *black box* che le ha generate, determina un rischio implicito per la parità delle armi, così come intesa dalla richia-

[31] Cfr. Van Dijk – Van Hoof, *Theory and Practice of the European Convention on Human Rights*, 3rd ed., Leiden, 1998, 430 ss.

[32] In questo senso, CEDU, *Kress* v. *France*, ric. 39594/98 (2001), § 72.

[33] J.F. Renucci, *Droit européen des Drotis de l'Homme. Droits aux libertés fondamentaux garantis par la CEDH*, 5a ed., Paris, 2013, 378.

[34] CEDU., *Brandstetter v. Austria*, cit., § 68.

[35] Così, CEDU, *Bykov v. Russia*, cit., § 90, da ultimo ripresa in *Svetina* v. *Slovenia*, cit., § 44, nella quale la questione denunciata dal ricorrente riguardava proprio l'impiego di prove raccolte sulla base di un iniziale, illegittimo (perché non espressamente autorizzato dal locale "giudice istruttore") accesso al telefono della vittima. Posta ancora una volta di fronte al problema dell'applicabilità della teoria dei frutti dell'albero avvelenato, la Corte ha rilevato che le giurisdizioni interne hanno fatto applicazione della contraria dottrina della "inevitable discovery"; tuttavia, poiché la questione della ammissibilità o meno delle susseguenti prove – che, appunto, secondo la Suprema Corte slovena sarebbero state scoperte comunque, a prescindere dall'illegittimo accesso – riguarda in definitiva l'interpretazione di norme interne, la Corte europea si limita ad osservare che le risultanze dell'accesso illegittimo non sono state poste alla base della decisione sulla colpevolezza dell'imputato, fondata, invece, su prove validamente raccolte, secondo la disciplina nazionale.

[36] CEDU, *Martinie* v. *France*, ric. 58675/00 (2006), § 46.

mata giurisprudenza europea. Se l'essenza dell'equità processuale risiede nel pieno diritto di poter provare a convincere, con strumenti efficaci, il giudice della propria ricostruzione dei fatti, anche contestando l'ammissibilità e l'accuratezza della prova, l'impossibilità di verificare *a posteriori* l'*output* di un algoritmo può rappresentare *in nuce* una violazione dell'art. 6, §1 CEDU (a prescindere dall'esistenza di una violazione, a monte, del diritto alla riservatezza).

Occorre dunque verificare se e quali rimedi possono essere utilizzati nel processo per contrastare l'intrinseca mancanza di trasparenza che circonda un dato generato automaticamente[37], difetto di trasparenza che può essere dettato, appunto, dall'impossibilità di rivelare il codice sorgente o dal funzionamento del modello computazionale utilizzato, non concepito per essere verificabile *ex post*.

L'ormai tradizionale risposta al problema dell'opacità[38] dei processi algoritmici e computazionali è la trasparenza[39]. Tuttavia, nell'ambito della trattazione automatizzata dei dati, la trasparenza pare essere divenuta l'unico e determinante parametro di legittimità del trattamento, sostituendosi subdolamente al canone della legalità. Se il software è concepito secondo parametri di trasparenza, la possibilità di validazione o di falsificazione dei suoi *outputs* è più elevata e a questo assunto sembrano ispirati il GDPR, recentemente entrato in vigore, e per certi versi anche la direttiva UE 2016/680, in materia di protezione delle persone fisiche con riguardo al trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali (art. 20), recentemente trasposta anche in Italia con il d.lgs. 51/2018[40].

Tuttavia, la trasparenza non è un concetto autosufficiente, ma si articola in relazione al risultato che si desidera ottenere[41]. Essa, ad esempio, si può raggiungere ottenendo l'accesso al *source code*, agli *inputs* e agli *outputs* del *software*[42]. In primo luogo, però, va precisato che tale accesso non garantisce una generale comprensione del processo che ha generato il risultato, perché soltanto gli esperti informatici possono essere in grado, e non sempre (vedi qui di seguito), di trarne degli elementi significativi e comprensibili. È stato osservato, quindi, che, in ogni caso, si tratta di una trasparenza "mediata"[43] dall'esperto. In secondo luogo, i codici sorgente possono essere sottoposti a segreti commerciali o industriali da parte dei proprietari del

---

37 F. Palmiotto, *The Impact of Algorithmic Opacity on Fair Trial Rights in Criminal Proceedings*, in M. Ebers - M. Canero Gamito (eds.), *Algorithmic Governance and Governance of Algorithms*, New York, 2020, 49 ss.

38 J. Burrell, *How machines think: Understanding opacity in machine-learning algorithms*, in *Big Data and Society*, 1, 2016, 1 ss.

39 M. Hildebrandt, *Profile transparency by design? Re-enabling double contingency*, in M. Hildebrandt - de Vries, *Privacy, Due Process and the Computational Turn*, London, 2013, 239; J. Danaher, *Algorithmic Decision-making and the Problem of Opacity*, in *Computers and Law*, 8, 2016, 29 ss.

40 Pubblicato in G.U. 24 maggio 2018 ed entrato in vigore il 6 giugno 2018.

41 F. Palmiotto, *The Impact*, cit. 52.

42 J.A. Kroll - J. Huey - S. Barrocas - E.W. Felten - J.R. Reidenberg - D.G. Robinson - H. Yu, *Accountable Algorithms*, in *University of Pennsylvania Law Review*, 165(3), 2017, 675.

43 A. Koene - H. Webb - M. Patel, *First UnBias Stakeholders workshop*, 2017, in *unbias.wp.horizon.ac.uk*.

software[44]. In base al diverso atteggiamento dei singoli ordinamenti sul punto[45], si può verificare una ipotesi di impossibilità di accesso ai fini della verifica dell'attendibilità della prova.

Inoltre, nemmeno l'*open source code* – che parrebbe a prima vista la principale garanzia di trasparenza - può garantire la possibilità di un'effettiva giustificazione[46] a posteriori dei risultati prodotti dall'algoritmo, se questo non è stato concepito con criteri, più che di trasparenza, di responsabilità (*accountability*, intesa come possibilità, capacità di dar conto di come i risultati sono stati prodotti, partendo da determinati *inputs*)[47]. Per un verso, nell'ambito della ricerca e della raccolta della prova difficilmente è possibile utilizzare *software open source*, proprio perché l'efficacia dei captatori occulti sta nella segretezza, innanzitutto, del loro operare, ma anche delle loro modalità di funzionamento. Per altro verso, poi, quando il *software* faccia uso di forme anche non particolarmente ricercate di *machine learning*, la validazione *ex post* del risultato può diventare impossibile anche per lo stesso designer, in ragione dei processi di autoapprendimento appunto impiegati dal software.

Tuttavia, l'*explainable AI* è una sfida che vede oggi impegnati studiosi di vari settori, concentrati nel tentativo di arginare la tradizionale aura di opacità algoritmica. Se è vero che la "prova computazionale" esalta e mette in luce il rischio che, in una società basata sulla produzione, sulla comunicazione e sul trasferimento di dati, i soggetti processuali (*in primis*, le parti) vengano fortemente deprivati della loro rilevanza nel procedimento probatorio (dalla raccolta, ma anche dalla valutazione, dalla discussione e dalla valutazione)[48], i tempi sono maturi per esplicitare il rischio e neutralizzarlo. Riconosciuto che, nell'attuale realtà storica, i dati raccolti o elaborati digitalmente rischiano di vedersi garantita una patente di intrinseca attendibilità probatoria, semplicemente perché la verificazione dell'iter che li ha generati è troppo complessa o sfugge, almeno in parte, ad un controllo *ex post*[49] nelle scienze computazionali e nei consolidati principi del sistema processuale penale europeo possono rinvenirsi adeguate reazioni al descritto pericolo.

In primo luogo, come accennato, la conoscenza del problema dell'opacità algoritmica spinge gli esperti del settore a elaborare soluzioni che, pur senza disvelare i codici

---

[44] L'esistenza del segreto commerciale è stata considerata dirimente in un noto caso deciso dalla Corte suprema del Wisconsin, nel 2017; v., *infra*, nt. 56.

[45] Per quanto riguarda l'ordinamento italiano, si sono al momento registrare tre significative pronunce del Consiglio di Stato, le prime due decisamente orientate alla prevalenza dell'interesse pubblico alla trasparenza degli atti amministrativi (Cons. Stato, sez. VI, 8 aprile 2019, n. 2270; Cons. Stato, sez. VI, 13 dicembre 2019, n. 8472) e una terza che, tuttavia, riconosce il ruolo di controinteressato, nel processo amministrativo, al titolare di segreti commerciali del software, poiché la sua sfera giuridica potrebbe subire degli effetti negativi dal disvelamento dei codici sorgente (Cons. Stato, sez. VI, 2 gennaio2020, n. 30).

[46] M. Hildebrandt, *Algorithmic Regulation and the Rule of Law*, *Phil. Trans. R. Soc*, 2018, 1-11.

[47] Cfr. A. Kroll - J. Huey - S. Barrocas - E.W. Felten - J.R. Reidenberg - D.G. Robinson - H. Yu, *Accountable Algorithms*, cit., 662 ss.

[48] C. Chessman, *A Source of Error: Computer Code, Criminal Defendants, and the Constitution*, in *Calif. L. Rev.,* 2017, 179 ss.

[49] S. Quattrocolo, *Equità del processo penale e automated evidence alla luce della giurisprudenza della Corte europea dei diritti dell'uomo*, in *Rev. italo-española dir. proc.*, 2, 2019, 1 ss.

sorgente di un software, ne spieghino in maniera esaustiva, validandolo, il funzionamento[50]: il loro impiego nel processo potrebbe risultare sufficiente a garantire la parità delle armi. In secondo luogo, è possibile pensare che, nella insuperabile necessità di accedere ai codici sorgente, per la verifica della prova, e a fronte di istanze di tutela del segreto commerciale da parte del proprietario del software, si possano elaborare, all'interno delle discipline processuali nazionali, regole di *disclosure* "garantita", ovvero circoscritta all'interno del procedimento penale, pur compatibilmente con il principio di pubblicità del giudizio. In terzo luogo, ove nessuna delle soluzioni precedenti sia percorribile, in ragione delle specifiche qualità del software, e la verifica *ex post* dell'attendibilità della prova risulti impossibile, la via pare segnata verso l'esclusione dell'ammissione o dell'utilizzazione della medesima, sulla scorta della incompatibilità con l'essenza irrinunciabile del processo equo. Pur riconosciuta, infatti, la ritrosia della Corte europea a stabilire delle *exclusionary rules* probatorie[51], nell'evoluzione della giurisprudenza di Strasburgo sulla «*overall fairness of the proceeding*», sembra potersi leggere proprio la soluzione indicata[52].

4. Decisori giurisdizionali e... ausili digitali

Terzo e più complesso profilo di analisi riguarda la sfera di applicazione in talune articolazioni del procedimento penale, di software "predittivi" che possono asseritamente assistere l'autorità giudiziaria in operazioni decisorie. Tali strumenti sono, per ora, più diffusi negli ordinamenti di *common law*, per lo più nella fase dell'esecuzione della pena, anche se non sono pochi gli Stati che vi fanno ricorso per decisioni di *bail*[53] e/o *sentencing*, ovvero in materia di custodia cautelare.

In numerose giurisdizioni degli Stati del Nord America si utilizzano, ormai da tempo, software predittivi per sciogliere prognosi di pericolosità sociale e, in particolare, di rischio di recidivanza. Si tratta di strumenti di *risk assessment* strutturati sulla base di valutazioni psico-criminologiche[54], vietate, nel giudizio di cognizione italiano, dall'art. 220 c. 2 c.p.p. Per ragioni che qui non si possono approfondire, un simile divieto risulta assai raro, nel panorama mondiale, tanto più che in numerosi ordinamenti si è verificato, nel XX secolo, un fenomeno di forte apertura verso le scienze psico-criminologiche, che ne ha fatto, talvolta il centro del sistema sanzionatorio[55]. A prescindere dall'atteggiamento riservato alle scienze psicologiche, permangono a tutt'oggi, nel giudizio pe-

[50] Cfr. A. Kroll - J. Huey - S. Barrocas - E.W. Felten - J.R. Reidenberg - D.G. Robinson - H. Yu, *Accountable Algorithms,* cit., 676, segnalano, per esempio i c.d. sistemi "*zero-knowledge proof*", in grado di fornire risposte esaustive senza rivelare necessariamente i dati posti alla base del funzionamento di un modello.

[51] A. Cabiale, *I limiti*, cit., 89.

[52] S. Quattrocolo, *Artificial Intelligence*, cit., 96.

[53] Molto noto, poichè applicato in 39 giurisdizioni degli Stati Uniti, il *Public Safety Assessment:* strumento attuariale basato su 9 fattori, tra cui età, accusa e precedenti/carichi pendenti.

[54] Per una completa panoramica sull'evoluzione degli strumenti di risk assessment, G. Zara - D. P. Farrington, *Criminal Recidivism: explanation, prediction and prevention*, Oxon, 2016, 148 ss.

[55] S. Quattrocolo, *Artificial Intelligence*, cit., 144 ss.

nale, numerosi momenti in cui l'autorità giudiziaria è chiamata a svolgere valutazioni di tipo prognostico-predittivo, gravando il decisore di una prognosi estremamente complessa. Non solo il problema è rappresentato dalla indecifrabilità del comportamento umano, ma anche dalla tendenziale scarsità di informazioni disponibili, soprattutto laddove sia intervenuta una dichiarazione di colpevolezza da pare dell'imputato a porre fine all'attività istruttoria: soprattutto negli ordinamenti in cui il ricorso al *plea bargaining* sia massiccio, l'ausilio di strumenti algoritmici di *risk assessment* risulta particolarmente appetibile. Quindi, come accennato in precedenza, prima ancora che a una delegazione di scelte decisorie alla macchina, siamo di fronte alla riduzione – auspicata e cercata! - della discrezionalità del singolo, attraverso l'ausilio della "decisione algoritmica".
I lettori che hanno un minimo di familiarità con questi temi hanno già riconosciuto, in questi cenni, il richiamo ad un noto caso deciso dalla Corte Suprema del Wisconsin nel 2017[56], ove, da tempo, è stato adottato uno strumento attuariale di *risk assessment* chiamato COMPAS[57]. Tale strumento, ben noto agli studiosi di fenomeni di recidivanza, si basa sia su informazioni ottenute direttamente dall'imputato, in un'intervista, sia sul certificato del casellario e dei carichi pendenti, le quali vengono elaborate attraverso un modello computazionale in relazione a dati statistici di controllo, riferiti a un campione di popolazione non necessariamente corrispondente a quella dello Stato in cui si svolge il procedimento. Sul piano predittivo, quindi, lo strumento prevede il rischio di ricaduta violenta, in rapporto al dato statistico, senza tuttavia offrire una spiegazione di tale rischio.

---

[56] *State v. Loomis*, 881 NW 2d 749 (Wis 2016). Per un commento alla sentenza v. *Criminal Law – Sentencing Guidelines – Wisconsin Supreme Court Requires Warnings before Use of Algorithmic Risk Assessment in Sentencing – State v. Loomis*, in *Harward Law Review*, 2017, 1530 ss. Nella vicenda richiamata, sulla base del *risk assessment*, la corte locale aveva inflitto la pena della reclusione a sei anni (senza *parole*), e cinque anni di *extended supervision*, pena certamente elevata se rapportata ai fatti, marginali, per cui egli si era dichiarato colpevole, destando l'attenzione di tutti i media nazionali e di molti stranieri.
Nell'ambito di una istanza di *post-conviction release*, decisa dalla *circuit court* locale, l'imputato contestava diversi profili di violazione del principio del *due process*. Il consulente tecnico presentato dalla difesa evidenziava alcuni aspetti critici legati all'uso in fase deliberativa della pena dello strumento di *risk assessment*.

[57] *Correctional Offenders Managment Profilying for Alternative Sanctions*. Si tratta di uno strumento attuariale che valuta il rischio statico, non dinamico: infatti, gli strumenti attuariali non spiegano il recidivismo, si limitano a segnalarlo, valutando i fattori di rischio attraverso statistiche ufficiali e prospettive teoriche comprensive. Sul mercato, lo strumento è commercializzato in forma di software, da Northpointe inc. che ne detiene i diritti e le licenze commerciali. Gli strumenti di *risk assessment*, però, non sono necessariamente dei software. Nel panorama italiano l'applicazione del *risk assessment* non ha ancora trovato un riconoscimento ufficiale all'interno del sistema della giustizia penale (v. però nt. 43); tuttavia, esso è molto diffuso in altri ordinamenti e da tempo oggetto di studi anche da parte di autori italiani: G. Zara, F. Freilone, *Psychological assessment*, in B.A. Arrigo (a cura di), *The SAGE Encyclopedia of Surveillance, Security, and Privacy*, Thousand Oaks, 2018, 830 ss; G. Zara, *La validità incrementale della psico-criminologia e delle neuroscienze in ambito giuridico*, in *Sistemi intelligenti,* 2, 2013, 311. Le teorie psicologiche applicate dal COMPAS sono illustrate in v. T. Brennan – W. Dietrich – B. Ehret, *Evaluating the Predictive Validity of the Compas Risks and Needs Assessment System*, in *Criminal Jusice and Behaviour*, 2009, 21 ss. (Brennan risulta aver guidato anche gruppi di ricerca per conto del produttore del medesimo software). Esistono numerosi studi, di segno non univoco, sull'attendibilità del COMPAS e sui rischi di implicit bias ad esso connessi: T. L. Fass - K. Heilbrun - D. Dematteo; R. Fretz, *The LSI-R and the COMPAS: Validation Data on Two Risk-Needs Tools*, in *Crim. Just. & Behavior*, 35, 2008, 1095 ss., i quali concludevano per un evidente fattore di discriminazione su base razziale dei risultati del software COMPAS; J. Skeem - J. Eno Louden, *Assessment of Evidence on the Quality of COMPAS*, 2007.

Nel caso richiamato, presa visione della valutazione dell'imputato fornita dal COMPAS, la corte locale lo aveva condannato alla reclusione a sei anni (senza *parole*), e cinque anni di *extended supervision*, pena certamente elevata se rapportata ai fatti, marginali, per cui l'imputato si era dichiarato colpevole. La difesa aveva presentato una *post conviction motion*, al rigetto della quale veniva proposto ricorso innanzi alla Corte suprema statale. I motivi della doglianza consistevano in tre punti. Innanzitutto, si denunciava la violazione del diritto dell'imputato ad essere valutato sulla base di informazioni accurate. Poi, si lamentavano la violazione del diritto ad una sentenza individualizzata, nonché l'impiego, erroneo, da parte dello strumento, del sesso fra i parametri presi in considerazione nel giudizio di pericolosità. Inoltre, essendo lo strumento tutelato da segreto commerciale, la difesa riteneva che le parti e il giudice non avessero avuto sufficienti spiegazioni sui criteri con cui erano stati determinati i punteggi di rischio, e i singoli fattori pesati, introducendo così nel *sentencing* elementi decisori sottratti alla *discovery* della difesa.

La Corte suprema statale, tuttavia, confermava la decisione di primo grado, senza apparentemente cogliere tutti gli spunti formulati dalle argomentazioni difensive, ma limitandosi ad escludere la violazione del *due process*, data la possibilità per l'imputato di confrontare i dati individuali di partenza (*input*) e le valutazioni di rischio finali (*output*) sulla base del manuale d'uso dello strumento, potendo così adeguatamente confutarne l'attendibilità[58]. A prescindere dalla criticabilità di tale assunto, un passaggio importante della sentenza è sfuggito all'attenzione dei molti media che si sono occupati della vicenda, proposta come un esempio di "pena stabilita dalla macchina"[59]. Nel testo della decisione, infatti, si ritrova una sorta di "decalogo cautelativo" che i giudici devono impiegare nell'utilizzo di tali strumenti "predittivi", articolato in cinque avvertimenti che devono sempre essere inseriti nel *pre-sentencing report*, ovvero: l'eventuale esistenza di un segreto commerciale che copre il software; l'incapacità del software di effettuare una valutazione altamente individualizzata, essendo basato su un set di dati riferiti a gruppi sociali, non normalizzata rispetto alla popolazione di ciascuno Stato; la creazione dello strumento per finalità specificamente collegate a scelte proprie della fase esecutiva, successiva al *sentencing*, nonchè l'esistenza di dubbi, nella comunità scientifica, circa l'attendibilità del modello computazionale - pur segreto – che lo regola.

Alle precauzioni suggerite dalla Corte Suprema del Wisconsin fa da eco una successiva sentenza, pronunciata dalla Corte Suprema del District of Columbia, sezione minorile, del 15 maggio 2018. La vicenda era molto simile alla precedente, ma aveva ad oggetto un diverso strumento di *risk assessment*, il SAVRY, non digitalizzato, somministrabile solamente attraverso un professionista[60], ed elaborato per la valutazione di minorenni.

---

[58] Uno studio di Angwin et alii (V. J. Angwin et alii, *Machine bias*, in 23 maggio 2016), pubblicato dalla ONG americana ProPublica ha mostrato la scarsa rilevanza criminogenica di alcuni fattori utilizzati nel COMPAS. È bene tuttavia segnalare che le conclusioni dello studio diffuso da ProPublica sono state fortemente criticate (v. A.W. Flores – K. Bechtel – C.T Lowenkamp, *False Positives, False Negatives and False Analysis: A Rejoinder to «Machine Bias: There is Software used across the Country to Predict Future Criminals. And it is biased against the Blacks»*, in *Federal Probation,* 2016).

[59] A. Liptak, *Sent to Prison by a Software Program's Secret Algorithms*, in *The New York Times*, 1 maggio 2017.

[60] Si tratta del SAVRY, impiegato in almeno nove Stati dell'Unione, su cui cfr. G.M.Vincent - J. Chapman - N. E. Cook, *Risk-Needs Assessment in Juvenile Justice: Predictive Validity of the SAVRY, Racial Differences, and*

In questo caso, la difesa formulava una istanza rivolta alla Corte di esclusione della prova fornita dal *risk assessment*, nonché di tutta la relazione predisposta dai servizi sociali, anche sulla base del medesimo e di qualsiasi testimonianza o altra prova ad esso collegata, denunciandone la inutilizzabilità sulla base della *rule* 702 delle *Federal Rules of Evidence*, così come interpretato dalla Corte Suprema federale nel caso *Daubert v. Merrel Dow Pharmaceutics*[61]. Tale decisione, infatti, rappresenta, a tutt'oggi, lo statuto di ammissibilità e utilizzabilità della prova tecnico-scientifica nel procedimento penale e non solo negli Stati Uniti, ma anche in numerosissimi altri ordinamenti che hanno seguito tale pronuncia.

La Corte, in parziale accoglimento delle richieste della difesa dell'imputato, minore al tempo del fatto contestato, ha fatto divieto di utilizzare per la decisione del caso specifico, la valutazione generale di *violence risk* predisposta dalla *Child Guidance Clinic* sulla base del *risk assessment*: pur senza pronunciarsi, in generale, sulla validità della teoria scientifica che sorregge il SAVRY, la Corte ha infatti ritenuto che, nell'applicazione al caso specifico, i risultati del test non fossero scientificamente attendibili.

Il richiamo a queste due recenti decisioni nordamericane ha una duplice ricaduta. Per un verso, ci ricorda come anche la tradizione europea e, in particolare, quella italiana, continuino a gravare il giudice penale di valutazioni di carattere predittivo: dalla decisione cautelare, fino alla quantificazione della pena in sentenza (si pensi, oltre che al generale paradigma dell'art. 133, 2 c.p., al perdono giudiziale), alla concessione di benefici, anche poi penitenziari, l'andamento del processo, il suo esito, la sanzione, il successivo trattamento, spesso si basano su predizioni che il giudice è chiamato a svolgere, senza dettagliati parametri, né specifiche informazioni. Per altro verso ci dimostra la necessità di avviare, senza ritardo, una seria riflessione giuridica che si sovrapponga a quella squisitamente computazionale che si concentra soltanto sulla efficacia, in termini di attendibilità, dello strumento algoritmico o computazionale, come effettivamente suggerito dalla Carta etica europea per l'impiego dell'intelligenza artificiale nei sistemi giudiziari, pubblicata dalla CEPEJ, nel dicembre del 2018[62].

the Contribution of Needs Factors, in *Crim. Just. &Behavior*, 2011, 47 ss. V., più recentemente e più in generale, J. Skeem - N. Scurich - J. Mohanan, *Impact of the Risk Assessment on Judges' Fairness in Sentencing Relatively Poor Defendants*, in *University of Virginia School of Law SSRN Papers series*, 15 gennaio 2019.

[61] 509 U.S. 579 (1993).

[62] Disponibile in *coe.int*. In tema, volendo, S. Quattrocolo, *Intelligenza artificiale e giustizia: nella cornice della carta etica europea, gli spunti per un'urgente discussione tra scienze penali e informatiche*, in *lalegislazionepenale.it*, 18 dicembre 2018.