Focus

Equo processo penale e sfide della società algoritmica

Serena Quattrocolo*

FAIR TRIAL AND THE CHALLENGES OF AN ALGORITHMIC SOCIETY

ABSTRACT: The lecture focuses on the main issues related to the use of algorithms and computational models in the realm of criminal proceedings. The analysis encompasses three main topics: the impact of software and models 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 decision-making instruments with fundamental rights.

KEYWORDS: algorithms; computational models; criminal proceeding; fair trail; evidence

SOMMARIO: 1. Introduzione – 2. Violazioni della riservatezza – 3. I rischi per la parità delle armi. – 4. Le interferenze nei processi decisori giurisdizionali.

1. Introduzione

Nell'ampio titolo scelto per questo intervento sono contenuti due concetti di portata generale. Il primo termine della relazione è "equo processo" che di per sé – più che un concetto – è uno slogan, come osservato da Mario Chiavario[1], con riferimento alla formula declamatoria impiegata dal primo comma del rinnovato testo dell'art. 111 Cost. Nella sua aspecificità, però, il termine si presta perfettamente a coprire diverse declinazioni delle garanzie processuali, sulle quali si concentrerà questa breve riflessione. Naturalmente, occorrerà fare riferimento sia al quadro costituzionale interno, sia al contesto dei Bills of Rights internazionali, nei quali gli specifici profili che concorrono a garantire la fondamentale e irrinunciabile equità del processo penale hanno ricevuto, ad oggi, la più compiuta elaborazione. In particolare, il riferimento principale, sarà alla Convenzione europea dei diritti dell'uomo.

Il secondo termine – società algoritmica – è evidentemente e volutamente generico, poiché si riferisce a un fenomeno ampio[2], che coinvolge soluzioni tecnologiche tra loro molto differenziate, accomunate dalla produzione di dati generati automaticamente, i quali possono essere impiegati, con varie finalità, all'interno del procedimento penale[3]. Si tratta di una locuzione più sociologica che tecni-

* *Professore ordinario di Diritto Processuale Penale, Università del Piemonte Orientale. Mail: serena.quattrocolo@uniupo.it.* *Contributo sottoposto al referaggio del Comitato Scientifico.*

[1] M. CHIAVARIO, voce "*Giusto processo*" (processo penale), in EGT, 2001, 1.

[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.

ca, i cui confini debbono essere di volta in volta specificati, in base all'approccio prescelto. La stessa locuzione "algoritmo", del resto, può assumere una varietà di significati a seconda del contesto in cui viene utilizzato, con sfumature e variazioni talora apprezzabili, spesso non condivise dagli stessi esperti del medesimo settore. Ai limitati fini di queste brevi riflessioni, ben si può adottare la definizione di algoritmo offerta da Tarleton Gillespie, nel 2014, che è stata altresì assunta come paradigma dal prezioso studio recentemente pubblicato dal Consiglio d'Europa, *Algorithms and Human Rights* (dicembre 2017)[4]. In particolare, "*algorithms need not be software: in the broadest sense, they are encoded procedures for transforming input data into a desired output, based on specified calculations. The procedures name both a problem and the steps by which it should be solved*."[5] In questo senso, la società algoritmica non è, necessariamente, quella che delega la scelta alla macchina – come una certa visione distopica spesso suggerisce – bensì quella che perde fiducia nella discrezionalità e nell'intuito del singolo, finendo per inquadrare i processi decisori, anche quello giudiziario, in una relazione prestabilita, un algoritmo, appunto.

Fatta questa doverosa premessa, si possono mettere a fuoco tre aree nelle quali l'impatto del trattamento (processing) automatizzato di dati, non necessariamente personali e non necessariamente sensibili, attraverso modelli computazionali 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.

La rivoluzione digitale ha subito nell'ultimo decennio un clamoroso balzo in avanti, alimentato da due fattori principali[6]: la diffusione globale di *smartphones* e altri strumenti di comunicazione telematica che generano quotidianamente, e in modo gratuito, quintilioni di dati; un aumento esponenziale della capacità computazionale che consente di processare con tempi e costi irrisori quel numero pressoché infinito di dati. Ciò ha inevitabilmente estremizzato il divario, da sempre esistente, tra progresso tecnologico ed evoluzione normativa, quest'ultima rimasta confinata entro categorie che talvolta non rispondono più alla realtà.

2. Violazioni della riservatezza

Il primo profilo che merita di essere segnalato – seppur con approfondimento limitato, poiché già ampiamente trattato in altri contesti – è quello dell'impiego di software capaci di carpire segretamente conversazioni, flussi telematici e altri dati digitali. La gamma di azioni intrusive che possono essere realizzate attraverso *malwares*, inoculati da remoto nei dispositivi *hardware* è considerevole. Per *malwares* si intendono vari tipi di *malicious software*, tra i quali il maggiormente impiegato in ambito investigativo è il c.d. *trojan horse*, captatore informatico. Invisibile all'utente che ha in uso l'apparecchio infettato, esso consente varie forme di intrusione nella sfera digitale dell'interessato e,

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

[5] T. GILLESPIE, *The relevance of Algorithms*, in T. GILLESPIE, P. BOCZKOWSKI, K. FOOT (eds.), *Media Technologies*, Cambridge US, 2014, 167

[6] 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.

ISSN 2284-4503

in particolare: a) acquisizione di informazioni scambiate attraverso il mezzo infettato; b) attivazione da remoto di strumenti di geolocalizzazione, ripresa o registrazione audio; c) accesso e manipolazione dei *files* presenti nell'*hardware* infetto[7]. Il breve riassunto delle caratteristiche di questi software dimostra la loro evidente potenzialità investigativa, che è stata prontamente sfruttata dagli organi requirenti, a livello generalizzato, per lo più in assenza di un apposito quadro normativo. Per un verso, infatti, ci si è 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 – affermando che i *malwares* altro non siano che nuove modalità tecniche di svolgimento di tradizionali mezzi di ricerca della prova, appunto. Per altro verso, si è consapevolmente minimizzato l'evidente divario, in termini di intrusività che i captatori informatici presentano rispetto ai tradizionali strumenti intercettivi[8]. 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*[9], dove i profili di tale superiore insidiosità per la sfera di riservatezza degli individui sono dettagliatamente trattati[10], va sottolineata la necessità di una disciplina normativa apposita dei captatori telematici, che non possono essere regolati secondo il paradigma delle tradizionali intercettazioni (deve ricordarsi che la recente riforma delegata dalla "legge Orlando" al Governo è intervenuta sulle intercettazioni di comunicazioni, per il loro "riordino" e sull'impiego dei captatori informatici nel procedimento penale)[11]. In secondo luogo, non si può tacere 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 significato, di fronte alla possibilità di produrre, scambiare, conservare – ma anche carpire, intercettare, copiare – dati immateriali in uno spazio che non è più quello fisico. L'inarrestabile diffusione di apparecchi digitali che, per le loro ridotte dimensioni, seguono l'individuo ovunque e sempre, rende decisamente impossibile continuare ad applicare la tradizionale distinzione

[7] 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.

[8] 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»

[9] 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: http://www.europarl.europa.eu/RegData/etudes/STUD/2017/583137/IPOL_STU(2017)583137_EN.pdf)

[10] 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*»

[11] Complesso riassumere brevemente la vicenda italiana. A seguito di una nota decisione delle Sezioni Unite della Corte di cassazione (Cass. S.U. 1.7.2016, 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. www.penalecontemporaneo.it, 16.10.2016), avvenuto poi con il d.lgs. 29.12.2017 n. 216 (e d.m. 20.4.2018), che ha inserito nel codice di procedura penale una specifica disciplina del captatore informatico, ad oggi non ancora entrata in vigore.

tra luoghi pubblici e luoghi privati, come il domicilio, nel quale la captazione occulta tendenzialmente poteva avvenire soltanto a condizioni ancora più stringenti di quelle generali[12].
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, rappresenta ancora la cornice normativa di riferimento. Per un verso, nell'art. 8 Cedu, la Corte di Strasburgo ha individuato dei limiti bene precisi anche all'attività investigativa di analisi e profilazione dei dati [13]. Per altro verso, la Convenzione stessa lascia intravedere, sullo sfondo, un altro principio che può utilmente essere applicato a fronte dell'utilizzo processuale di mezzi di prova generati automaticamente.

3. I rischi per la parità delle armi

Il secondo profilo di analisi riguarda, infatti, i mezzi di prova. Infatti, 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 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 quotidiani collegati alla rete internet. Questi rappresentano, evidentemente, un patrimonio conoscitivo talvolta fondamentale per le indagini e per il procedimento penale: si pensi, ad esempio, ad un frigorifero "intelligente" che si accenda per mantenere perfetta la conservazione dei cibi ogni volta che la temperatura della stanza aumenti, per la presenza fisica di persone. Il software che lo regola fornirà informazioni determinati agli investigatori che indaghino su un omicidio avvenuto proprio in quella 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 computazionale. 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?
L'assunto sul quale si basa l'interrogativo formulato qui sopra è l'impossibilità di falsificare il dato elaborato da un algoritmo se non è possibile accedere al codice sorgente che governa l'algoritmo stesso[14]. Tale scenario, che assumiamo al momento come valido, genera 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[15]. L'ingresso di saperi specialistici nel processo è difficilmente equilibrato, poiché una delle parti - per lo più quella pubblica -ha accesso alla scienza e alle tecnologie migliori, disponendo di mezzi eco-

[12] V. ampiamente, S. SIGNORATO, *Le indagini digitali*, cit., 49 ss.
[13] Volendo S. QUATTROCOLO, U. PAGALLO, *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.
[14] S. QUATTROCOLO, U. PAGALLO, *The impact of AI on criminal law*, cit., 392 ss.
[15] 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.

nomici 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 algoritmica, tuttavia, introduce la forma più 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[16]. Ciò, evidentemente, trascende in maniera netta il profilo della riservatezza, sopra a grandi linee enunciato, per porre uno scottante problema di parità delle armi. È infatti sotto il profilo della parità delle armi che la situazione segnalata si pone maggiormente in contrasto con il canone del giusto processo, tanto nel quadro costituzionale italiano, quanto, e soprattutto, nella cornice del dettato convenzionale. E' 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[17]. Notoriamente, *equality of arms* non implica una presunta, necessaria identità di facoltà o di posizioni 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[18]. 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 la svantaggino rispetto alle altre parti[19]. Insomma, il principio esprime essenzialmente, nella sua portata generale, un giusto equilibrio tra le parti processuali[20]. 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. Austri*a, 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*»[21] non soddisfa il parametro convenzionale.

Ancor più particolare, poi, è l'approccio che la Corte di Strasburgo adotta quando è chiamata ad accertare una violazione potenzialmente verificatasi nel procedimento probatorio, tradizionalmente devoluto alla disciplina nazionale. Infatti, in questo contesto, 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 tra le parti. È stato, infatti, ripetutamente sottolineato che «*it*

[16] E. VAN BUSKIRK-V.T. LIU, *Digital Evidence: Challenging the Presumption of reliability*, in *Journal of Digital Forensic Practice*, 2006 (1), 20

[17] 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, p. 192. V., in particolare, C. eur., 28.8.1991, *Brandstetter v. Austria*, § 66; C. eur., 23.6.1993, *Ruis-Mateos v. Spain*, § 63.

[18] Cfr. VAN DIJK – VAN HOOF, *Theory and Practice of the European Convention on Human Rights*, 3rd ed., Leiden, 1998, 430 ss.

[19] In questo senso, C. eur., 7.6.2001, *Kress c. Francia*, § 72.

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

[21] C. eur., *Brandstetter v. Austria*, cit., § 68.

ISSN 2284-4503

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»[22]. L'assunto richiama quanto già sottolineato in precedenza rispetto alle decisioni *Khan v. the UK* e *Schenk v. Switzerland* che rappresentano la "cerniera" tra queste considerazioni formulate, in generale, rispetto al vasto panorama delle prove penali e gli specifici aspetti che caratterizzano le prove ottenute attraverso mezzi occulti.

Date queste coordinate generali, dunque, i problemi posti dalle prove raccolte e generate in via del tutto automatizzata sono evidenti. L'impossibilità – variamente declinata – di accedere al codice sorgente o di poter effettivamente comprendere il funzionamento dell'algoritmo che le ha generate determina un rischio implicito per la parità delle armi, così come intesa dalla richiamata giurisprudenza europea. Se l'essenza dell'equità processuale risiede nella parità delle armi, che si sostanzia (anche) nel diritto della difesa di contestare 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 vita privata e familiare).

Rimane da accertare che l'assunto dal quale siamo partiti sia effettivamente corretto, ovvero che l'accesso al *source code* che regola l'algoritmo sia impedito alla difesa.

La prima e più immediata risposta al problema dell'opacità[23] dei processi algoritmici e computazionali è la trasparenza[24]. Si è già in parte accennato al fatto che, 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[25]. 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, appena 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. 18.5.2018 n. 51[26].

---

[22] Così, C. eur., *Bykov v. Russia*, cit., § 90, da ultimo ripresa in C. eur., *Svetina c. 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.

[23] J. BURRELL, *How machines think: Understanding opacity in machine-learning algorithms*, in *Big Data and Society*, 2016, vol 1, 1ss.

[24] J. DANAHER, *Algortithmic Decision-making and the Problem of Opacity*, in *Computers and Law*, 2016, fasc. 8, 29 ss.

[25] V. *supra* nt. 21.

[26] Pubblicato in G.U. 24 maggio 2018 ed entrato in vigore il 6 giugno 2018.

Tuttavia, la trasparenza non è un concetto autosufficiente, ma si articola in relazione al risultato che si desidera ottenere. Essa, ad esempio, si può raggiungere ottenendo l'accesso al *source code*, agli *inputs* e agli *outputs* del *software*[27]. Tuttavia, tale accesso può non essere utile perché soltanto gli esperti informatici sono in grado, di trarne degli elementi significativi e comprensibili. E' stato osservato, quindi, che in tal caso non può dirsi comunque garantita la trasparenza[28].

Inoltre, nemmeno l'*open source code* – che parrebbe a prima vista la principale garanzia di trasparenza - può garantire la possibilità di un'effettiva validazione a posteriori dei risultati prodotti dall'algoritmo, se questo non è stato concepito con criteri più che di trasparenza – appunto - di responsabilità (*accountability*, intesa come possibilità, capacità di dar conto di come i risultati sono stati prodotti, partendo da determinati *inputs*)[29]. Per un verso, come già sottolineato, nell'ambito della ricerca e della raccolta della prova difficilmente è possibile utilizzare *software open source*, proprio perché l'efficacia degli specifici mezzi intercettivi 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 molto semplici di *learning machine*, la validazione *ex post* del risultato può diventare impossibile anche per chi lo abbia messo a punto, stante il meccanismo di auto-apprendimento che lo alimenta.

La 'prova computazionale' esalta e mette in luce il rischio che, in una società in cui tutto ciò che è oggetto di conoscenza e di comunicazione è un dato – è cioè costituito da o è racchiuso in una 'espressione digitale' - i soggetti del processo vengano deprivati del loro ruolo nel procedimento probatorio (dalla raccolta, ma anche dalla valutazione, dalla discussione e dalla valutazione). Il dato, raccolto o elaborato digitalmente rischia di divenire di per sé attendibile perché la verificazione del processo che lo ha generato è troppo complessa o sfugge, almeno in parte, per via del ricorso a forme più o meno sofisticate di intelligenza artificiale, ad un controllo *ex post*[30]. In tale quadro, l'accusa ha accesso, per evidenti ragioni, alla migliore tecnologia, i cui risultati vengono trasferiti nel processo penale come prove. La difesa, per le ragioni sopra esposte, non ha la possibilità di mettere convincentemente in dubbio l'attendibilità di tale prova, poiché non ha gli elementi necessari alla falsificazione. Il giudice, per parte sua – soprattutto in quegli ordinamenti più nettamente ispirati al principio dispositivo della prova - può non avere motivo di dubitare di tale prova, in assenza di elementi concreti addotti dalla difesa, 'adagiandosi' sul convincimento che il dato digitale sia scevro da rischi di inaccuratezza[31].

4. Le interferenze nei processi decisori giurisdizionali

Terzo e più complesso profilo di analisi riguarda la sfera di applicazione, in talune articolazioni del procedimento penale, di software 'predittivi' che possono anche assistere il giudice in operazioni de-

[27] 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*, 2017 (Vol. 165, Issue 3), 675.

[28] A. KOENE, H. WEBB, M. PATEL, *First UnBias Stakeholders workshop*, 2017, in https://unbias.wp.horizon.ac.uk

[29] Cfr. A. KROLL, J. HUEY, S. BARROCAS, E.W. FELTEN, J.R. REIDENBERG, D.G. ROBINSON, H. YU, *Accountable Algorithms*, cit., 662 ss.

[30] 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.*, vol. 2/2019, 1 ss.

[31] E. VAN BUSKIRK-V.T. LIU, *Digital Evidence*, cit., 20

cisorie. Si tratta di strumenti per lo più diffusi, per il momento, in alcuni ordinamenti di *common law*, principalmente con riguardo alla fase dell'esecuzione della pena, ma talvolta estesi a decisioni di *bail* e/o *sentencing*, ovvero in materia di custodia cautelare e di quantificazione della pena. Altrettanto diffuso è l'impiego di tali sistemi nell'ambito dell'attività di polizia amministrativa di sicurezza, ovvero con finalità di prevenzione del reato. Tuttavia, l'area della prevenzione del reato – scarsamente regolamentata in tutti gli ordinamenti – sfugge ad un'analisi sistematica, cadendo al di fuori di questa riflessione.

Sono numerosi gli Stati del Nord America e dell'Australia che 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 che, ad esempio, nel nostro ordinamento sono vietate nel giudizio di cognizione dall'art. 220 co. 2 c.p.p. Tuttavia, fuori da tale divieto e stante l'estrema complessità della prognosi che grava sul giudice e la scarsità di informazioni disponibili, soprattutto laddove, come nella quasi totalità dei casi, sia intervenuta una dichiarazione di colpevolezza da pare dell'imputato, il ricorso a strumenti algoritmici di *risk assessment* rappresenta un ausilio appetibile. Quindi, come accennato in precedenza, prima ancora di delegazione di scelte decisorie alla macchina, siamo di fronte alla riduzione della discrezionalità del singolo, attraverso l'ausilio della decisione algoritmica.

Caso che ha richiamato l'attenzione internazionale è quello deciso dalla Corte Suprema del Wisconsin nel 2016[32]. Qui, come in numerose altre giurisdizioni, è stato adottato uno strumento attuariale di *risk assessment* chiamato COMPAS[33] che si basa sia su informazioni ottenute direttamente

[32] 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*.

[33] 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*, Sage, Thousand Oaks, 2018, 830 ss; G. ZARA, *La validità incrementale della psico-criminologia e delle neuroscienze in ambito giuridico*, in *Sistemi intelligenti,* 2013, fasc. 2, p. 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*, 2008, vol. 35. 1095 ss., i quali concludevano per un evidente fattore di discriminazione su base razziale dei risultati del software COMPAS; J.

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. Sul piano predittivo, quindi, lo strumento prevede il rischio di ricaduta violenta, senza tuttavia offrire una spiegazione di tale rischio, ma in rapporto al dato statistico.

Anche con l'ausilio di tale strumento, 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 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 venivano articolati in tre punti. Il primo denunciava la violazione del diritto dell'imputato ad essere valutato sulla base di informazioni accurate. Il secondo lamentava la violazione del diritto ad una sentenza individualizzata, mentre il terzo riguardava il controverso uso, da parte dello strumento, del sesso fra i parametri presi in considerazione dal *risk assessment*. Infatti, 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. Negli argomenti difensivi che criticano i risultati del *risk assessment* sembrano mescolarsi profili incentrati sulla natura digitale della valutazione e sulla sua opacità, e questioni di attendibilità scientifica della teoria psico-criminologica applicata allo strumento valutativo, il cui effetto complessivo sarebbe stato quello di impedire la fisiologica interazione della difesa.

La Corte rigettava le questioni formulate e confermava la decisione della corte inferiore, senza apparentemente cogliere la loro duplice ispirazione, 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 confutarne l'attendibilità[34]. Tuttavia, un aspetto importante è sfuggito all'attenzione dei molti media che si sono occupati della vicenda, come di un esempio di 'pena stabilita dalla macchina'[35]. L'estensore di quella decisione ha voluto stilare un 'decalogo cautelativo' che i giudici devono impiegare nell'utilizzo di tali strumenti 'predittivi', indicando cinque avvertimenti che devono sempre essere inseriti nel *presentencing report*, ovvero: l'eventuale esistenza di un segreto commerciale che copre il software;

SKEEM, J. ENO LOUDEN, *Assessment of Evidence on the Quality of COMPAS*, 2007, in http://ucicorrections.seweb.uci.edu/files/2013/06/CDCR-Skeem-EnoLouden-COMPASeval-SECONDREVISION-final-Dec-28-07.pdf

[34] Uno studio di Angwin et Alii (V. J. ANGWIN ET ALII, *Machine bias*, in https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing 23.5.2016), 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).

[35] A. LIPTAK, *Sent to Prison by a Software Program's Secret Algorithms*, in *The New York Times*, 1.5.2017, in https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html?hp&action=click&pgtype=Homepage&clickSource=story-heading&module=first-column-region&region=top-news&WT.nav=top-news&_r=0

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*, nonché l'esistenza di dubbi, nella comunità scientifica circa l'attendibilità del modello computazionale - pur segreto – che lo regola.

Tali preoccupazioni sembrano essere state raccolte in una successiva sentenza, pronunciata dalla Corte Suprema del District of Columbia, sezione minorile, del 15 maggio 2018. La vicenda era del tutto analoga alla precedente, ma aveva od oggetto un diverso strumento di *risk assessment*, non digitalizzato, somministrabile solamente attraverso un professionista[36].

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* 720 delle *Federal Rules of Evidence*, così come interpretato dalla Corte Suprema federale nel caso *Daubert v. Merrel Dow Pharmaceutics*[37]. 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*.

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, il nostro processo, siano inclini a incorporare nel processo penale valutazioni di carattere predittivo, che, a partire dal contesto cautelare, fino a quello della quantificazione della pena (si pensi alla diminuente per la minore età), alla concessione di benefici, anche poi penitenziari, incidono in modo significativo sull'esito del processo, sulla sanzione e sul trattamento. 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.

A tal fine occorre, innanzitutto, imparare a distinguere i rischi che si presentano. In primo luogo, l'"algoritmicità" – che toglie spazio all'intuito dei singoli per delegare le scelte alla formula prestabilita - può essere contrastata con lo strumento "classico" del *Daubert test*: se la teoria scientifica tradotta nell'algoritmo non è convincente, quest'ultimo non deve trovare spazio nel processo penale, né nelle scelte decisionali del giudice. In secondo luogo, la "computazionalità" e la conseguente opacità della decisione debbono essere misurate secondo i parametri che le carte dei diritti fondamentali già ci offrono: giusto processo, parità delle armi, presunzione di innocenza devono guidarci anche nell'affrontare questo nuovo tipo di sfida.

[36] 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 the Contribution of Needs Factors*, in *Crim. Just. &Behavior*, 2011, 47 s. 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*, 15.1.2019.

[37] 509 U.S. 579 (1993).