

## **Il divieto di perizia sulla personalità dell'imputato alla prova delle nuove tecnologie?**

di Veronica Tondi

### *1. Premessa*

Il divieto di perizia psicologica nei confronti dell'imputato rientra, ormai, negli ambiti dell'ordinamento processuale messi alla prova dalle nuove tecnologie, anche e soprattutto digitali.

Appare, quindi, più pressante il problema dell'eventuale superamento di limiti che, pur trovando alcune legittime giustificazioni nella salvaguardia del libero convincimento del giudice e, in ultima istanza, del diritto penale del fatto, se intesi in modo assoluto possono ritenersi tali da riflettere una sfiducia forse eccessiva nei confronti di determinate scienze, quali, appunto, quelle psicologiche.

Quest'ultima considerazione, e quindi un atteggiamento di scetticismo verso l'attendibilità delle scienze psicologiche e dei risultati delle analisi condotte in applicazione delle stesse, si pongono alla base del divieto di perizia sulle qualità psichiche indipendenti da cause patologiche, posto, come è noto, dall'art. 220 c.p.p. Altre cause della preclusione sono tradizionalmente individuate nell'intento di evitare forme di incidenza sul convincimento del giudice in senso sfavorevole all'imputato, in potenziale contrasto con il principio della presunzione di innocenza, nonché nella possibile alterazione degli esiti dell'analisi ad opera della stessa persona che vi è sottoposta, finanche per finalità utilitaristiche afferenti all'adozione di determinate decisioni nel processo.

Nondimeno, proprio il presupposto del condizionamento in senso pregiudizievole all'imputato dell'apprezzamento giudiziale potrebbe essere posto in discussione, in considerazione della possibilità di pervenire a decisioni più accurate e calibrate sulle caratteristiche dell'individuo, auspicabilmente *in bonam partem*, offerta dagli sviluppi non solo delle scienze psicologiche, ma di strumenti tecnologici idonei a supportare le valutazioni scientifiche tradizionali, elaborando un novero complesso e consistente di dati secondo regole di natura anche – ma non solo - psico-criminologica.

Infatti, si è avvertita negli ultimi anni, anche nell'ambito dello svolgimento di perizie tanto di natura psichiatrica – che, come è noto, sono ammesse nell'ordinamento italiano – quanto di natura psicologica, l'esigenza di colmare i difetti di concordanza delle valutazioni degli esperti, attraverso strumenti tali da supportare queste ultime, fornendo risultati caratterizzati da maggiore oggettività.

Tra le tecnologie maggiormente in grado di porsi in rapporto dialettico con il divieto in considerazione, si possono menzionare quelle legati alle neuroscienze, nonché gli algoritmi predittivi, nella misura in cui essi siano impiegati al fine di determinare il pericolo di commissione di reati futuri, o il rischio di recidiva, sulla base di un'indagine coinvolgente anche le caratteristiche psichiche dell'accusato.

Con riferimento alle neuroscienze e alle tecniche di analisi del cervello umano, delle sue componenti e della sua conformazione fisiologica, delle tipologie di reazioni che possono attivarsi al suo interno, si tratta di strumenti che già da anni hanno trovato impiego nel contesto del processo penale, specialmente per finalità di valutazione della capacità di intendere e di volere dell'imputato, oltre che di vaglio della credibilità delle dichiarazioni testimoniali.

Naturalmente, problemi legati all'osservanza del divieto di cui all'art. 220 c.p.p., si porrebbero laddove si ipotizzasse l'uscita dei nuovi strumenti di indagine sulla psiche dall'ambito della perizia psichiatrica – e dunque dall'indagine avente ad oggetto i profili patologici della mente dell'imputato – e li si impiegasse per cogliere aspetti legati, più ampiamente, alla personalità dell'individuo e alla sua propensione a compiere determinati atti, o a reagire, secondo determinate modalità, agli stimoli a cui è sottoposto. Indubbiamente, assai significativo è, in tal caso, il rischio di una concezione deterministica della persona umana, o addirittura il profilarsi di forme di diritto penale d'autore.

## 2. Risk assessment tools e procedimento penale

Con riferimento all'utilizzo di algoritmi predittivi per l'apprezzamento di rischi di natura processuale e non processuale, quando debbano essere adottate dal giudice decisioni caratterizzate da un'inevitabile proiezione verso il futuro, l'analisi si rivela più complessa, potendosi concordare con la dottrina che non riconduce, in ogni caso, tali strumenti entro l'ambito di applicazione del divieto di perizia psicologica e criminologica, non risolvendosi questi, sempre e comunque, in un'analisi di tal genere.

Specialmente con riguardo all'ambito cautelare, dunque, sembra che l'elaborazione automatica di dati, ad opera dei dispositivi di intelligenza artificiale, secondo regole di tipo essenzialmente statistico e di natura esperienziale, ma con un indubbio potenziamento delle capacità di analisi riferibili alla persona umana, non sia preclusa, anche allo stato della legislazione vigente. Per un verso, dunque, la tecnologia consente di raggiungere risultati indisponibili alla mente umana, in termini di tempi di elaborazione e di mole delle informazioni trattate, e sembra quindi ricondurre l'impiego del *tool* di cui si tratta all'ambito di operatività della perizia; per l'altro, essa si avvarrebbe, appunto, di regole anche e soprattutto esperienziali, non necessariamente implicanti un'indagine psichica, almeno in uno stadio precoce del procedimento.

Un simile potenziamento può ritenersi auspicabile, specialmente in un contesto in cui la decisione del giudice deve intervenire in tempi rapidi, a fronte di carichi di lavoro spesso ingenti e tali da non consentire, talvolta, un apprezzamento adeguato di tutte le circostanze rilevanti del caso. Si tratterebbe, dunque, di un supporto alla valutazione giurisdizionale, che naturalmente, in seguito, dovrebbe porre in essere il necessario ragionamento individualizzante, secondo il principio generale

della combinazione della valutazione di tipo statistico e probabilistico con l'applicazione della regola alla luce delle peculiarità del caso concreto.

Particolarmente rilevante sarebbe, in tal caso, la consapevolezza, in capo al giudice, dei limiti del dispositivo impiegato, proprio in ragione del suo fondarsi su regole essenzialmente statistiche ed esperienziali, specialmente al fine di evitare l'introduzione di una sorta di "presunzione tecnologica", invincibile proprio perché tale.

È noto come l'impiego dei *risk assessment tools*, per la valutazione dei rischi di fuga o di recidiva, e per l'apprezzamento della pericolosità sociale, sia ormai una realtà, tanto all'interno, quanto all'esterno dei confini europei, essendo questi impiegati, ad esempio, nell'adozione dei provvedimenti in materia di *bail* negli Stati Uniti, o ancora in Inghilterra, ai fini delle decisioni in materia di *parole* o *probation*. Si tratta di contesti in cui gli algoritmi hanno mostrato di essere idonei a fondare decisioni in senso favorevole al *defendant*, consentendo di limitare il ricorso al rilascio su cauzione monetaria, nonché di rendere le determinazioni in materia di *pretrial detention* più confacenti ai pericoli effettivamente riferibili allo stato di libertà della persona arrestata, o di adottare decisioni di *diversion* processuale.

Invero, i provvedimenti basati sulla valutazione dei pericoli di fuga e di commissione di reati futuri si affidano, tradizionalmente, all'intuito e sull'ampia libertà di apprezzamento del giudice, che potrebbero essere, dunque, quantomeno sostenuti dall'esito di un più completo riscontro fondato su dati empirici. Problema, quello del rafforzamento dell'attendibilità empirica delle valutazioni degli esperti in materia di pericolosità sociale, che, come si è detto, è stato avvertito anche nelle scienze psicologiche e nell'esecuzione di perizie psichiatriche.

Certamente, tuttavia, si tratta di una prospettiva che presenta profili problematici non trascurabili. Per un verso, infatti, si può menzionare la difficoltà di selezionare e di elaborare dati idonei a fondare, con carattere di attendibilità, la previsione di rischi quali quelli di fuga e di recidiva.

Si pongono, inoltre, significative sfide, legate al rispetto dei diritti fondamentali e all'esigenza di evitare inaccettabili, ma plausibili *bias* e forme di discriminazione, come è emerso in alcuni noti casi decisi negli Stati Uniti.

Nondimeno, se adeguatamente progettati e implementati, e a fronte dell'esplicazione di alcune garanzie, a cui si farà cenno a breve, questi strumenti potrebbero condurre a un esito più oggettivo e prevedibile delle valutazioni sulla sussistenza di esigenze cautelari da arginare attraverso restrizioni delle libertà individuali, senza incorrere, auspicabilmente, almeno in alcune delle diverse forme di *bias* cognitivo a cui la stessa mente umana è soggetta, e che rimangono suscettibili di alcuna dimostrazione.

Al limite, caratteri di opportuna trasparenza dell'algoritmo potrebbero rendere più accessibili e agevolmente dimostrabili forme di distorsione del ragionamento logico, più difficilmente contestabili quando la loro fonte sia la stessa mente umana. Indubbiamente, l'unica e rilevante espressione esterna del *bias* umano è la motivazione del provvedimento, pur restando pressoché imperscrutabili, e soprattutto soggettivamente variabili, in relazione al differente bagaglio di valori e di esperienze di ciascuno, le forme di pregiudizio e le distorsioni logiche idonee a influenzare la decisione.

Non è difficile imbattersi, nel nostro ordinamento, in decisioni stereotipate, o caratterizzate da un apprezzamento parziale degli elementi rilevanti, quando debba essere adottata una decisione in materia cautelare, sebbene ben diverso sia il tipo di apprezzamento imposto dalle fonti sovranazionali, e in primo luogo dall'art. 5 CEDU.

Basti pensare al ruolo esplicito dalla gravità del reato, a livello espresso o inespresso, come è progressivamente accaduto anche in altri sistemi, come quello statunitense. Un utilizzo della tecnologia coerente con il rispetto dei diritti fondamentali e delle garanzie potrebbe dunque essere di ausilio, nella prospettiva di evitare decisioni semplicistiche incidenti negativamente sui diritti e sulle libertà individuali.

## *2. Le fonti sovranazionali e la tutela dei diritti fondamentali*

Proprio con riguardo al tema delle garanzie, e all'esigenza di assicurare controllo umano, qualità, sicurezza, trasparenza, le fonti sovranazionali offrono significative indicazioni al riguardo. Basti pensare alla Carta etica europea sull'uso dell'intelligenza artificiale, o alla più recente proposta di regolamento dell'Ue, che stabilisce regole armonizzate in materia di intelligenza artificiale.

Ancora, viene in considerazione l'art. 11 della direttiva 2016/680, relativa alla 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, che stabilisce un divieto di decisioni basate unicamente su processi automatizzati, se ne possono derivare effetti giuridici sfavorevoli nei confronti della persona. Essenziale, dunque, è la mancata limitazione della decisione giudiziale all'*output* dell'algoritmo, specialmente quando ne derivino conseguenze sulla libertà personale, essendo necessario tenere in opportuna considerazione le diverse caratteristiche del caso concreto. La medesima disposizione stabilisce un importante divieto di profilazione risultante in discriminazioni basate su ragioni etniche e razziali, su dati genetici e biometrici, sulle convinzioni religiose e filosofiche.

L'importanza di rispettare i principi di trasparenza, responsabilizzazione, non discriminazione e accessibilità, quando gli strumenti di cui si tratta siano impiegati nel contesto del procedimento penale, si riflettono anche nella Risoluzione del Parlamento europeo del 6 ottobre 2021

sull'intelligenza artificiale nel diritto penale e il suo utilizzo da parte delle autorità di polizia e giudiziarie in ambito penale (2020/2016).

Se si ha riguardo, quindi, in modo più approfondito alla menzionata proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale, si possono individuare previsioni tese a garantire la protezione dei dati personali, istanze di non discriminazione ed eguaglianza, nonché il diritto all'equo processo e a rimedi effettivi, il diritto di difesa e la presunzione di innocenza.

Quando dall'applicazione dell'intelligenza artificiale possono derivare limitazioni della libertà personale, secondo la proposta e in termini generali, è fondamentale assicurare un elevato livello qualitativo dei dati che rappresentano l'*input* del sistema, e un'adeguata verifica dell'attendibilità dello strumento, al fine di evitare rischi di discriminazione e altri pregiudizi.

Inoltre, in stretta correlazione con la tutela dei diritti fondamentali si pongono esigenze di trasparenza e di accessibilità. Di conseguenza, alla luce della proposta, dovrebbero essere considerati sistemi ad alto rischio, tra gli altri, quelli impiegabili dalla pubblica autorità per finalità di *risk assessment*.

Compaiono, infatti, nel terzo allegato alla proposta, tra i sistemi ad alto rischio, i sistemi di IA destinati a essere utilizzati dalle autorità di contrasto per effettuare valutazioni individuali dei rischi delle persone fisiche al fine di determinare il rischio di reato o recidiva in relazione a una persona fisica, o il rischio per vittime potenziali di reati.

La qualificazione del sistema come ad alto rischio ha importanti implicazioni con riguardo ad aspetti come gestione dei dati, documentazione, trasparenza, controllo umano, accuratezza e sicurezza.

Sono, ad esempio, imposti requisiti in materia di set di dati di addestramento, convalida e prova (art. 10), risultanti dalla documentazione tecnica dello strumento.

Con riguardo alla trasparenza, essa è espressamente (art. 13) ricollegata alla possibilità di una congrua interpretazione e utilizzabilità dell'*output*: ad esempio, l'utilizzatore dovrebbe disporre di adeguata informazione sull'attendibilità dello strumento impiegato.

Nondimeno, solo ove opportuno, sono imposte le specifiche per i dati di *input* o qualsiasi altra informazione pertinente in termini di set di dati di addestramento, convalida e prova, sebbene si tratti di un profilo decisivo nei *risk assessment tools*.

Ma, soprattutto, l'art. 14 della proposta concerne il problema fondamentale del controllo umano, ponendo in luce la necessità, per la persona chiamata ad adottare una decisione basata sull'*output* del sistema di IA ad alto rischio, di essere consapevole dei limiti propri della tecnologia impiegata e di essere in grado finanche di decidere in difformità rispetto a tali esiti, se il caso concreto lo richiede.

Assai opportunamente, la proposta esprime l'esigenza di una piena consapevolezza del cd. *automation bias*, inteso quale possibile tendenza a fare acritico affidamento sull'*output* del sistema di IA ad alto rischio, o a sopravvalutarlo.

Tali previsioni confermano la possibilità di fare riferimento ai *tools* di cui si tratta solo quale supporto, e non in sostituzione, delle decisioni umane implicanti l'apprezzamento di determinati rischi.

Tra i sistemi di IA vietati, l'art. 5, inoltre, colloca «l'uso di sistemi di IA da parte delle autorità pubbliche o per loro conto ai fini della valutazione o della classificazione dell'affidabilità delle persone fisiche per un determinato periodo di tempo sulla base del loro comportamento sociale o di caratteristiche personali o della personalità note o previste», quando ne derivino conseguenze pregiudizievoli in un contesto diverso rispetto a quello nel cui ambito i dati sono raccolti, o quando vengano a determinarsi effetti pregiudizievoli ingiustificati o sproporzionati rispetto all'entità del comportamento sociale tenuto. Si tratta di previsioni significative, nella misura in cui sembrano escludere la possibilità di una ricerca e raccolta massiva di informazioni riguardanti gli individui, in violazione dei diritti fondamentali e della disciplina in materia di protezione dei dati personali, con l'obiettivo di predisporre una sorta di profilo sociale di rischio della persona, utilizzabile dalle pubbliche autorità in diversi contesti.

### 3. *Considerazioni conclusive*

Probabilmente, sono auspicabili più incisive previsioni, anche a livello europeo, sulla trasparenza dei *risk assessment tools*.

Il coinvolgimento del settore pubblico nella predisposizione di strumenti accessibili e affidabili, con un ampio contributo di giuristi e di esperti di diverse discipline, potrebbe condurre a modalità nuove e auspicabilmente più accurate di adozione delle decisioni coinvolgenti ragionamenti di tipo predittivo, anche nella fase delle indagini. E, tanto, senza, naturalmente, sopprimere il momento intuitivo della decisione, che, come dimostrato anche da studi recenti, incide in senso positivo sul processo deliberativo.

Solo il giudice può, in definitiva, nell'adattare le regole statistiche ed esperienziali alle peculiarità del caso concreto, assolvere adeguatamente all'adozione di decisioni fortemente incidenti sui diritti fondamentali. Peraltro, qualora venga in considerazione un apprezzamento di natura psico-criminologica, oppure ove si tratti di vagliare le regole di programmazione di un certo algoritmo, il vaglio giudiziale secondo i criteri propri della prova tecnico-scientifica, nonché l'esplicazione di un adeguato contraddittorio, anche mediante gli apporti dei consulenti tecnici di parte, appaiono non rinunciabili.

Peraltro, qualora l'impiego dello strumento tecnologico si risolva in un'indagine propriamente diretta alle caratteristiche psichiche non patologiche dell'imputato – come non necessariamente accade in tale ambito, secondo quanto si è anticipato – esso non sembrerebbe utilizzabile se non a fronte di un superamento, o di una revisione, del divieto di cui all'art. 220 c.p.p. A tale ultimo proposito, si è già riflettuto sulle complesse implicazioni di una disposizione che, per un verso, risponde a rilevanti istanze di garanzia, ma per l'altro limita l'ingresso, nel processo, di contributi scientifici, e tecnici, che potrebbero rappresentare un valido ausilio nell'adozione di decisioni implicanti la valutazione di profili della personalità dell'individuo, lasciate al libero apprezzamento e all'intuito del giudice.

In definitiva, l'intrinseca limitazione delle capacità di elaborazione della mente umana, nonché dei tempi e delle risorse disponibili, e l'auspicio di una maggiore accessibilità delle distorsioni cognitive interferenti sulle valutazioni di cui si tratta – mediante l'imposizione di requisiti di trasparenza nella selezione dei dati e nella programmazione dell'algorithmo – sembrano rendere la prospettiva di ausilio degli strumenti tecnologici una sfida da affrontare con orientamento positivo, purché si proceda nella previsione e nel rafforzamento delle garanzie, secondo gli spunti già offerti dalle fonti sovranazionali.