# Non esiste la "pallottola d'argento": l'*Artificial Face Recognition* al vaglio giudiziario per la prima volta[1]

*di Andrea Pin*

**Title**: There Is No "Silver Bullet:" The Artificial Face Recognition's First Judicial Scrutiny

**Keywords**: Artificial Face Recognition; Judicial Scrutiny; Privacy; EU law; Human Rights Act.

1. – Molto recentemente, la divisione gallese della *High Court of Justice* ([2019 EWHC 2341]) ha affrontato una questione che, per sua medesima ammissione, non aveva precedenti giudiziari in nessuna giurisdizione (par. 1): l'utilizzo dell'*Artificial Face Recognition* (AFR) nella previsione e nella individuazione dei reati. La sentenza *R* (*Bridges*) *v. The Chief Constable of South Wales Police et al.* fa dunque d'apripista a questioni giurisprudenziali che difficilmente mancheranno di sollecitare ancora le corti tanto interne quanto sovranazionali e internazionali, impegnate a equilibrare la tutela della privacy con quella della sicurezza e dell'ordine pubblico (*ex multis*, A. Bonfanti, *Big Data e polizia predittiva: riflessioni in tema di protezione del diritto della privacy e dei dati personali*, in *MediaLaws* 3/2018, p. 1). L'utilizzo dell'AFR, un sistema basato sull'intelligenza artificiale che effettua la ricognizione degli individui attraverso l'uso di telecamere, è da tempo oggetto di riflessione, sia per le speranze che essa consente di coltivare, sia per le perplessità che suscita (*ex multis*, P.W. Nutter, *Machine Learning Evidence: Admissibility and Weight*, 21 *J. Const. L.* 919, 920, 2019). Finalmente una Corte si è dunque occupata dell'"uso appropriato e non arbitrario dell'AFR in una società libera e civilizzata" (par. 1).



La decisione appare effettivamente significativa ben oltre i confini della giurisdizione, soprattutto in ragione dei criteri che il giudice adotta per decidere. Per i parametri normativi che utilizza, la sentenza esprime l'impostazione tipica del diritto inglese ed europeo, ma affronta il fenomeno anche da angolature che si prestano ad essere riprese in altri ordinamenti.

2. – La Corte immediatamente assume una posizione prospettica di lungo periodo, notando come l'uso della tecnologia da parte delle forze dell'ordine non sia mai stato messo seriamente in discussione. Come essa evidenzia, prevenire e rilevare i reati tramite metodologie scientificamente avanzate, bilanciando la protezione dei diritti

[1] L'autore ringrazia i revisori anonimi per le utili osservazioni volte al miglioramento del presente lavoro.

individuali con il pubblico interesse, ha infatti a lungo impegnato la giurisprudenza interna. Questo è avvenuto soprattutto a partire dall'impiego dei dati biometrici (ovvero il "riconoscimento delle persone basato sulla misurazione e l'analisi delle loro caratteristiche biologiche o dei loro dati comportamentali", secondo la definizione dell'*Home Office Biometrics Strategy 2018*, p. 5). Il corso è iniziato con le impronte digitali, cui si sono poi affiancati il DNA e il controllo facciale, il cui impiego ha favorito sia l'individuazione dei colpevoli sia l'esclusione dalle indagini di soggetti evidentemente estranei a un crimine (par. 5).

È all'interno di questo filone che la Corte inquadra il ricorso all'AFR. Si tratta, essa ammette, di una tecnologia dal grande potenziale per "la prevenzione e l'individuazione del crimine, la cattura dei sospetti o dei colpevoli e la protezione del pubblico", di gran lunga superiore alle tecniche che semplicemente riprendono i passanti (par. 25). Tuttavia, rileva la medesima Corte, l'AFR ha aspetti inquietanti: grazie alla sua capacità di riconoscere 40 individui al secondo (par. 35), è in grado di processare i dati biometrici della popolazione su larga scala e senza il coinvolgimento di quest'ultima (par. 43).

Il caso origina nel quadro di un progetto pilota, dal nome di *AFR Locate*, condotto dalla polizia del Galles meridionale a beneficio dell'intero Regno Unito. La sentenza dunque riguarda, come la stessa Corte ammette, la punta più avanzata dell'uso dell'AFR in Gran Bretagna (para. 8). Il progetto si basa sulla raccolta di immagini digitali di volti ripresi in tempo reale in un contesto dinamico, in cui le persone si muovono liberamente. Tali immagini vengono confrontate con un database in possesso delle forze dell'ordine. Il database concretamente impiegato in ciascuna occasione può mutare, ma complessivamente include destinatari di mandati di cattura, individui illegittimamente sottrattisi a misure detentive, sospettati di avere commesso reati, scomparsi o bisognosi di protezione, di interesse per i servizi di *intelligence* o la cui presenza in un particolare scenario desta una particolare preoccupazione (par. 30). Il software impiegato isola i volti e ne estrae dati biometrici al fine di compararli, distinguendo le persone per le quali c'è un riscontro nei database da quelle che invece non risultano inserite in tale lista. Ebbene, nei 50 utilizzi concretizzatisi tra il 2017 e il 2019, *AFR Locate* ha processato 500,000 volti (par. 36). Anche se naturalmente tale cifra non indica il numero di soggetti coinvolti, giacché il medesimo individuo può essere stato ripreso più volte, il dato è significativo, se si considera che l'intero Galles ha una popolazione di poco più di 3 milioni di individui. Di qui, la necessità per la Corte di considerare attentamente in che modo il progetto *AFR Locate* si sia effettivamente svolto, per valutare se esso abbia configurato una forma di monitoraggio di massa (par. 7).



Al fine di effettuare tale verifica, la sentenza effettua un approfondimento relativo alle modalità con le quali *AFR Locate* utilizza le immagini delle telecamere. Il software, essa nota, compara le immagini dei volti con i database esistenti, di volta in volta riscontrando o escludendo una corrispondenza. Il riscontro tra queste ultime e i volti contenuti nel database si basa sull'attribuzione di un "punteggio di somiglianza", il quale stabilisce la probabilità che due immagini siano riconducibili alla medesima persona. Come rileva la Corte, quest'operazione è particolarmente delicata perché esige di individuare una soglia appropriata, evitando un eccesso di "falsi positivi" (riscontri erronei) o di "falsi negativi" (mancati riscontri) (par. 24).

Nei due anni di operatività del progetto pilota, la polizia ha collocato delle postazioni mobili di telecamere in luoghi affollati, concentrando la sua attenzione su eventi socialmente significativi, come la finale della *Champions League* disputatasi a Cardiff nel giugno 2017 o altri eventi pubblici ai quali ci si attendeva una copiosa affluenza. Il software ha verificato l'esistenza di riscontri ("*matching*") tra gli individui ripresi dalle telecamere e alcuni dei volti di persone che, per diverse ragioni, comparivano nei database della polizia. Il database massimo con il quale il software

può operare ammonta a 2000 individui (par. 31); tuttavia, la polizia è in grado di ridurre ulteriormente il numero di individui rispetto ai quali effettuare la verifica. Ad esempio, complessivamente, il database specifico per la finale di *Champions League* conteneva 919 persone. Durante l'utilizzo del software e della telecamera in quella giornata ci sono stati 10 riscontri, di cui 8 risultati poi fondati (par. 11). In un'altra occasione c'è stato un solo riscontro, anch'esso poi verificato come effettivamente positivo (par. 14)

L'impiego di *AFR Locate* è condotto con modalità esplicite. La polizia anticipa sui social media che ne farà uso; esibisce dei cartelli in formato A2 nei luoghi in cui viene utilizzata la AFR e distribuisce volantini alle persone del pubblico in merito all'attività che sta svolgendo (par. 39). Infine, il protocollo prevede che, nel caso in cui il software individui un potenziale riscontro, le due immagini siano considerate da un operatore umano, ossia un agente di polizia: per la Corte, questa è una importante salvaguardia, su cui si tornerà più tardi (par. 33). Nel caso in cui l'operatore confermi la valutazione del software, le forze di polizia si attivano per identificare il soggetto per il quale si è verificato il *matching*. All'identificazione eventualmente segue, quando del caso, la cattura o la semplice segnalazione all'*intelligence* (par. 33).

3. – A rivolgersi alla Corte non è stato uno dei soggetti rientranti nei database, ma, al contrario, qualcuno che non vi compariva. Infatti, l'attore ha lamentato di essere stato presente in due occasioni di grande affluenza (par. 11), esponendo la propria figura alla telecamera senza essere stato avvertito della presenza di tale apparecchiatura (par. 13 e 16). Le doglianze hanno riguardato tre distinti profili. In primo luogo, la violazione del diritto alla privacy garantito dall'art. 8 della Convenzione europea dei diritti dell'uomo; in secondo luogo, il mancato rispetto delle norme interne di derivazione Ue a protezione dei dati personali; infine, la violazione dei doveri di uguaglianza imposti dalla normativa nazionale sul pubblico settore, in quanto il software cagionerebbe uno sproporzionato numero di riscontri erronei tra alcune fasce della popolazione (par. 18 e 21).



La sentenza affronta le doglianze glissando su di un paradosso iniziale. Non era infatti noto né all'attore, né alla polizia, né tantomeno alla Corte se questi fosse stato effettivamente ripreso dalla telecamera (par. 16). Questo perché, in mancanza di *matching*, *AFR Locate* elimina immediatamente e automaticamente i dati biometrici raccolti, privandoli alla disponibilità della polizia che gestisce il software (par. 37).

4. – In tema di privacy e Cedu, la Corte riconosce come l'"AFR permette che l'operazione relativamente comune consistente nell'osservare gli individui si svolga molto più velocemente, efficacemente ed estesamente". È la tecnologia di questo tipo che deve obbligare a "riflettere per il suo potenziale impatto sulla privacy" (par. 46).

La Corte ricapitola l'estesa protezione della "vita privata" assicurata dall'art. 8 Cedu, pacificamente ricomprendente la vita sociale delle persone e dunque le occasioni di aggregazione durante le quali gli individui sono stati ripresi. Poi essa passa a considerare in particolare gli aspetti tecnici dell'*AFR* alla luce della privacy. In primo luogo, nota la Corte, il fatto che *AFR Locate* smaltisca immediatamente le informazioni raccolte in mancanza di un riscontro non esclude l'applicazione della privacy, in quanto è la mera raccolta delle immagini, non la loro conservazione, ad attivarne la tutela (par. 52). Del resto, *AFR Locate* per il giudice fa ben più che raccogliere immagini, come si trattasse di un apparecchio fotografico installato in una pubblica piazza. "L'informazione digitale che contiene l'immagine viene analizzata e i dati facciali biometrici ne vengono estratti. Le informazioni vengono poi ulteriormente processate per confrontarle con la lista di individui" in possesso della

polizia (par. 55). *AFR Locate* dunque svolge operazioni ben oltre ciò che chiunque potrebbe attendersi come "prevedibile e non sorprendente" in un luogo pubblico (par. 55). Tramite il software la pubblica autorità assimila informazioni di carattere "intrinsecamente privato", non diversamente dalle impronte digitali o dal DNA (par. 57).

Tali considerazioni consentono alla Corte di affrontare il tema successivo, ossia se l'interferenza con l'art. 8 sia fondata su una base normativa come richiesto dalla Convenzione. Su questo essa opina per l'ipotesi positiva. Infatti, essa nota come i poteri di polizia radicati nel *common law* al fine di prevenire i reati e reprimerli ricomprendano tradizionalmente lo scatto di fotografie negli spazi pubblici (par. 70) e la creazione di database di individui che rivestono un interesse per gli scopi di tutela della sicurezza (par. 77). Non vi rientrano le operazioni sotto copertura o i "metodi intrusivi", che consistono in accessi in proprietà private o azioni che altrimenti si qualificherebbero come violenza. Tuttavia, *AFR Locate* non è ascrivibile a nessuna di queste due ipotesi: essa non richiede alcun "accesso fisico, contatto o uso della forza" ed esige che le forze dell'ordine informino del suo uso il pubblico (par. 82-84).

La Corte passa poi a sottoporre l'uso dell'AFR al test di proporzionalità, nella sua formulazione quadripartita, consolidatosi con *Bank Mellat v Her Majesty's Treasury (No 2)* ([2014] AC 700). Nella sua declinazione inglese, tale test viene dunque ripercorso dalla Corte, la quale si interroga *i*) se la misura sia sufficientemente importante da giustificare la limitazione di un diritto fondamentale; *ii*) se sia razionalmente connessa all'obiettivo; *iii*) se fosse disponibile una soluzione meno intrusiva nei confronti dei diritti coinvolti ed ugualmente efficace; *iv*) se complessivamente la scelta compiuta bilanci in maniera equilibrata i diritti individuali e gli interessi della comunità (par 98; si veda *Lord Sumption* in *Bank Mellat*, p. 20). La sentenza riscontra i requisiti del test nella trasparenza dell'operazione, largamente preannunciata e pubblicizzata, nella mancanza di arresti di persone basati su falsi riscontri, nel fatto che non ci siano state lamentele da parte del pubblico e che il sistema elimini immediatamente le informazioni relative ai soggetti privi di riscontro nel database (par. 101).



La Corte si concentra infine soprattutto sul requisito della "proporzionalità in senso stretto", considerando i contesti in cui la tecnologia è stata utilizzata e gli esiti cui ha condotto. Essa giustifica l'uso di *AFR Locate* alla luce dei disordini accaduti nel corso degli eventi cui anche l'attore aveva partecipato e per il fatto che solo tramite l'impiego di tale software in quelle occasioni era stato possibile effettuare alcuni arresti (par. 102). Nota infine l'importanza di una lista di individui per i quali tentare il riscontro: questa infatti limita la potenziale sorveglianza generale del pubblico (par. 104). Complessivamente, tali elementi suffragano, nel ragionamento della Corte, l'uso di *AFR Locate* secondo i protocolli sopra descritti. L'unica ipotetica doglianza, secondo la Corte, potrebbe piuttosto riguardare la compilazione della lista per i potenziali riscontri: una ingiustificata inclusione di un individuo in un *database* potrebbe effettivamente dare luogo ad una violazione della privacy (par. 105). Al contrario, conclude la Corte, l'uso dell'AFR, in sé, non soffre di un chiaro o sistematico "deficit di proporzionalità" (par. 108).

5. – La Corte dedica meno spazio alle doglianze relative alla protezione dei dati personali, pur toccando sia la direttiva 96/46 quanto il GDPR. In particolare, la Corte non esprime alcuna posizione relativamente all'applicabilità dell'art. 22 del GDPR (transitato negli artt. 14 e 49-50 del DPA 2018), il quale, come noto, regola i diritti del soggetto che subisce una decisione basata su di un processo automatizzato. Tuttavia, anche nell'ambito della privacy la sentenza presenta alcune considerazioni di qualche rilievo. Innanzitutto, anche alla luce della giurisprudenza della Corte di

Giustizia, essa nota come i dati personali che vengono processati siano quelli di tutti coloro che vengono ripresi dalle telecamere, indipendentemente dal fatto che poi manchi un riscontro nel database, come invece sosteneva la polizia. È la "potenzialità ad essere identificati", derivante dal processo di riconoscimento del volto, a generare l'interesse alla tutela dei dati (par. 113 e 125), che pertanto riguarda i soggetti compresi nei *database* quanto la generalità della popolazione ripresa dalle telecamere di *AFR Locate* (par. 132). Del resto, reitera la sentenza, tale software, al fine di escludere un individuo ripreso da una telecamera, deve verificarne il riscontro nel database, il che configura, a tutti gli effetti, una forma di "*data process*" (par. 133).

La Corte poi passa a valutare l'impatto della raccolta dei dati sulla protezione della privacy. La normativa nazionale prescrive che tale valutazione debba considerare il complesso di operazioni effettuate con i dati, i rischi potenziali per le libertà e i diritti degli individui coinvolti e infine le misure volte ad evitarli, considerando anche altri interessi potenziali (Sezione 64, *Data Protection Act* 2018, derivante dall'art. 35 GDPR).

Sul punto la Corte conferma implicitamente la tradizionale deferenza del giudiziario inglese nei confronti degli altri poteri. Essa attesta che la polizia gallese ha effettuato una valutazione d'impatto, contenente una dettagliata descrizione del trattamento dei dati e delle sue potenziali implicazioni per la privacy (par. 148). Tuttavia, una volta assicuratasi che tale valutazione si sia svolta tenendo conto dei fattori in gioco, la Corte si arresta. Evita infatti radicalmente di poter "sostituire la propria opinione a quella del responsabile della tenuta dei dati"; ricava per sé lo spazio per eventualmente riscontrare una violazione solo nel caso in cui il responsabile del trattamento dei dati "abbia svolto il suo compito secondo una metodologia rivelatasi falsa, o in una maniera chiaramente fallace" (146). Il che, nota la Corte, non risulta avvenuto nel quadro dell'utilizzo dell'AFR da parte della polizia. Anche sul punto, la Corte pertanto conclude per la legittimità dell'operato della polizia e dei protocolli di *AFR Locate*.



6. – Infine la Corte affronta il dovere di eguaglianza imposto sul pubblico settore, in base all'*Equality Act 2010*. Questo prescrive che "le autorità pubbliche devono, nell'esercizio delle loro funzioni, avere riguardo a tre aspetti: (a) la necessità di eliminare la discriminazione…; (b) la necessità di promuovere l'uguaglianza di opportunità tra persone che condividono e coloro che sono prive di una caratteristica protetta; (c) la necessità di promuovere le buone relazioni" tra queste due categorie (Sezione 149). L'attore sosteneva che *AFR Locate* restituisse risultati indirettamente discriminatori in base al sesso e alla razza; esso infatti, a suo dire, eccedeva nei "falsi positivi" soprattutto quando venivano analizzati i dati biometrici di donne, persone di colore e minoranze etniche (par. 152).

Su quest'ultima doglianza la Corte manifesta una forte irritazione, ritenendo che essa abbia "un'aria di irrealtà" (par. 153), poiché il software non offrirebbe alcuna evidenza discriminatoria. L'unica pezza d'appoggio per la doglianza consisterebbe nella testimonianza di un esperto, che aveva evidenziato come lo sviluppo della capacità del software dipenda da un *training*, ossia una fase di apprendimento in cui alla macchina vengono somministrati esempi dai quali essa impara. Se gli esempi riguardano in misura inferiore alcune categorie, rispetto a queste *AFR Locate* sviluppa dunque minore familiarità e può errare più facilmente.

In realtà, la doglianza non era totalmente destituita di fondamento. Come la stessa Corte deve ammettere, i test avevano dimostrato un eccesso di falsi positivi soprattutto tra le donne, rispetto a quanto accadeva tra gli uomini. Grazie ad alcuni periti, essa aveva tuttavia ricondotto quest'anomalia alla presenza nel database di "due volti femminili … dai connotati molto diffusi". L'eccesso di falsi positivi non sarebbe

stato dunque causato da un pregiudizio di genere, ma da una certa tipologia di volto contenuta nel database: fenomeno noto come "*lambs*" (par. 154-155).

A chiudere definitivamente la questione, a detta della Corte, sarebbe stata la presenza dell'uomo. Infatti, con *AFR Locate* "nessuna azione viene intrapresa nei confronti di un membro del pubblico a meno che un agente (l'operatore del sistema) non abbia verificato il riscontro potenziale generato dal software, giungendo all'opinione che ci sia un riscontro tra il membro del pubblico e la lista dei volti" (par. 156). In quest'enfasi sulla verifica dell'operatore, la sentenza pare aderire alla dottrina prevalente sul tema, che ritiene che l'AFR debba affiancare, senza sostituire, l'identificazione di un soggetto effettuata da un essere umano (J. Nawara, *Machine Learning: Face Recognition Technology Evidence in Criminal Trials*, 49 *U. Louisville L. Rev.* 601, 603, 2011).

7. – Le osservazioni offerte dalla Corte gallese hanno effettivamente una portata ampia, viste anche le promesse che paiono derivare dall'AFR. Nonostante, come ammette la sentenza, questa non sia una "pallottola d'argento" senza difetti e assolutamente efficace (par. 107), è senz'altro un potente strumento per la tutela dell'ordine e della sicurezza pubblica. Intervenendo in un campo giurisprudenziale che attualmente pare essere una *tabula rasa*, la sentenza rileva le potenzialità di un sistema di controllo penetrante come l'*AFR* e reputa alcune tutele come fondamentali per consentirne l'utilizzo, alcune delle quali echeggiano chiaramente le preoccupazioni della Corte di Giustizia dell'Unione almeno da *Digital Rights Ireland* (C-293/12 e C-594/12; *ex multis*, A. Vedaschi, *I programmi di sorveglianza di massa nello Stato di diritto. La "data retention" al test di legittimità*, in *DPCE*, 2014, p. 1224 ss.). L'immediata eliminazione dei dati raccolti non rilevanti, la trasparenza sull'utilizzo di tale tecnologia e la supervisione umana presentano altrettante garanzie essenziali nella riflessione della Corte; del resto vi pongono l'accento anche gli *Orientamenti etici per un'IA affidabile* resi dal Gruppo indipendente di esperti ad alto livello sull'intelligenza artificiale (Commissione Europea, 2019).



Il ruolo della supervisione umana è giustamente posto in particolare evidenza dalla Corte, sebbene questa taccia sull'applicazione dell'art. 22 GDPR e della relativa disciplina nazionale. Sul punto, il GDPR prevede che, in caso che il processo automatizzato produca effetti giuridici o di genere analogo, il soggetto che li subisce abbia il diritto di esigere un intervento umano; il fatto che la Corte non si sia espressa in alcun modo potrebbe forse far pensare che non associ effetti giuridici al monitoraggio facciale Tuttavia la sentenza non offre alcun elemento che faccia propendere per questa ipotesi.

La Corte getta invece luce su altri aspetti interessanti, rispetto ai quali la supervisione umana ha un ruolo importante da svolgere a prescindere dall'applicabilità dell'art. 22. I sistemi di riconoscimento facciale automatico, infatti, normalmente si sviluppano attraverso *training* il cui risultato non raggiunge mai la certezza. In effetti, si può parlare di gradi più o meno elevati di risultati "probabilmente approssimativamente corretti" (S. Russell-P. Norvig, *Artificial Intelligence. A modern approach.* 3rd ed, Harlow, 2016, p. 714). La supervisione umana dunque effettua una verifica davvero rilevante, consentendo di innalzare ulteriormente il livello di approssimazione verso l'esattezza.

Rimangono però nell'argomentazione alcuni punti d'ombra. In primo luogo, la Corte adotta un livello di scrutinio piuttosto blando relativamente alla valutazione dell'impatto dell'AFR sulla privacy. Come sopra visto, essa si ritiene soddisfatta nel rilevare che la polizia gallese ha tenuto conto degli elementi giuridicamente rilevanti quando ha sviluppato il protocollo per l'utilizzo del software; ritiene invece al di là delle sue competenze una valutazione più approfondita relativamente al modo in cui

protocollo bilancia concretamente gli interessi in gioco. È difficile predire se il vaglio giudiziario prenderà anche in altri ordinamenti la stessa strada, anche alla luce delle tipicità del diritto inglese relativamente all'intensità dello scrutinio nei confronti dell'apparato amministrativo.

Un secondo aspetto problematico riguarda le doglianze sulla discriminazione potenzialmente causata dal software. Infatti, la Corte liquida piuttosto superficialmente l'osservazione di un perito che aveva esplicitamente affermato che il "*bias* è stato scoperto essere un aspetto dei comuni sistemi di AFR". Sul punto essa replica semplicemente che il perito non ha offerto una opinione sulla possibilità o sul grado in cui "tale *bias* può essere affrontato" da contromisure come "assicurare che un operatore umano controlli se effettivamente sussiste un riscontro" (par. 157).

La considerazione della Corte non appare così facilmente condivisibile. Il tema delle potenziali disuguaglianze è profondamente legato alle caratteristiche del software, che, dietro una parvenza di oggettività e certezza, può sviluppare atteggiamenti discriminatori. Questa possibilità si concretizza soprattutto quando entra in gioco la componente probabilistica (F. Faini, *Big data, algoritmi e diritto*, in questa *Rivista* 3/2019, p. 1870), che origina nella fase nella quale un software viene sottoposto ad un *training* (A. Roth, *Trial By Machine*, 104 *Geo. L.J.* 1245, 1269, 2016). Non a caso, molta enfasi viene normalmente posta sul cosiddetto "*dataset*", ossia sul materiale con il quale il software viene alimentato, giacché un *dataset* squilibrato può riverberarsi sui risultati (S. Olhede-P. Wolfe, *When algorithms go wrong, who is liable?*, in *Significance*, 8/2017, p. 8). Per tornare al caso di specie, se il *training* imperfetto produce un eccesso di falsi positivi nei casi di donne e minoranze, all'operatore umano viene sottoposto un maggior numero di verifiche su di una porzione della popolazione piuttosto che un'altra. Il fatto che l'operatore escluda i falsi positivi non toglie il fatto che la sua attenzione sia maggiormente sollecitata nei confronti di una certa fascia del pubblico esposto alla telecamera.



Il tema dei *biases* negli algoritmi è particolarmente scottante negli Stati Uniti, ove le corti hanno preso ad utilizzare con larghezza software per calibrare le condanne penali alla luce di una vasta mole di dati sul potenziale di recidiva. Nonostante l'impiego del software sia ancora dibattuto e anzi abbia ricevuto un cauto avallo da alcuni giudici (si veda *State v. Loomis*, 881 N.W.2d 749, 2016), molte voci si sono levate sul "*disparate impact*" dell'AI nei confronti di alcune minoranze etniche (*ex multis*, M. Hamilton, *The Biased Algorithm: Evidence of Disparate Impact on Hispanics*, 56 *Am. Crim. L. Rev.* 1553, 2019). Ed è particolarmente degno di nota come alcuni studi ammettano che il software non può contemporaneamente garantire un trattamento uguale tra gli individui quanto tra i gruppi etnici, ma che dev'essere operata una scelta tra garantire il principio di uguaglianza tra i singoli o un trattamento equanime tra tutti i gruppi etnici e di genere che compongono la società (J.A. Kroll-S. Barocas-E.W. Felten-J.R. Reidenberg-D.G. Robinson-H.Yu, *Accountable Algorithms*, 165 *U. Pa. L. Rev.* 633, 685, 2017). Alcuni propongono persino che, per evitare *bias* nel *dataset*, debbano strutturarsi forme di sorveglianza di massa, tramite le quali i limitati dati a disposizione su cui si costruiscono inconsapevoli discriminazioni verrebbero integrati dall'immissione di una mole ingente di informazioni relativa all'intera società (si veda Roth, *cit.*, p. 1302).

Per tornare alla sentenza gallese, i limiti del software sembrano dunque difficilmente eliminabili tramite la presenza di un operatore umano, anche se quest'ultimo può senz'altro svolgere una funzione di vaglio importante. La Corte sul punto ha argomentato in maniera forse troppo semplicistica, ponendo poca attenzione alle implicazioni discriminatorie dell'AFR. Probabilmente non se n'è troppo curata perché l'AFR è qui utilizzato nella fase di sorveglianza e non sanzionatoria. Un altro contesto potrebbe costringere a una maggiore ponderazione.

Note e commenti – *DPCE on line*, 2019/4
ISSN: 2037-6677