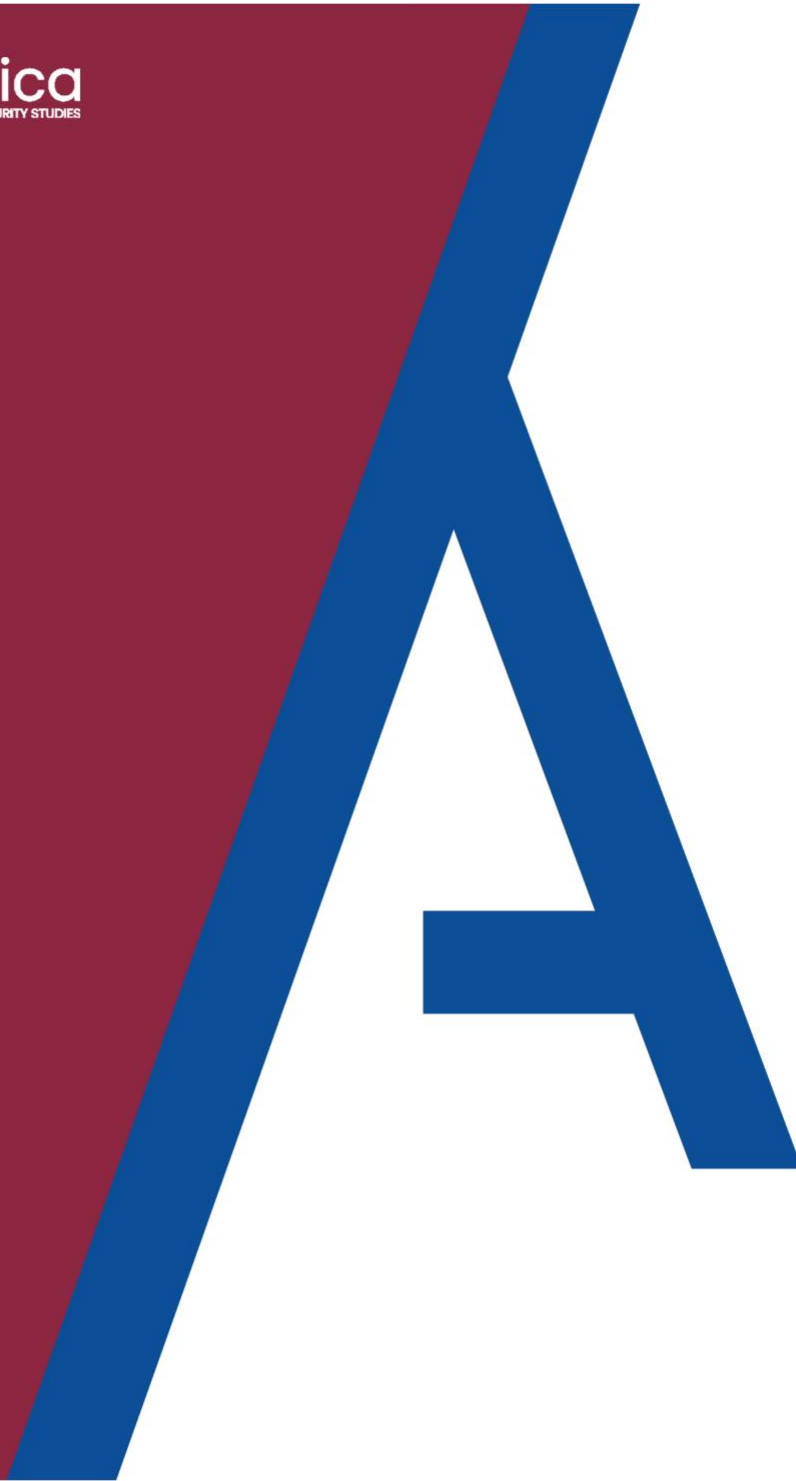


Analytica
FOR INTELLIGENCE AND SECURITY STUDIES



La sfida della discriminazione algoritmica nella
giustizia di Europa e Stati Uniti.

Giulia Molinari



Analytica for intelligence and security studies

Paper Cyber security

La sfida della discriminazione algoritmica nella giustizia di Europa e Stati Uniti

Giulia Molinari

(Tirocinante University of Groningen – Master of Arts International Relations and International Security)

Correzioni e revisioni a cura del Dott. Avv. Davide MANISCALCO e del Dott. Maurizio SPELTA



Introduzione

Nonostante molti strumenti dotati di intelligenza artificiale (IA) siano stati sviluppati da ormai diversi anni, è invece molto più recente la pubblicazione dei primi documenti ufficiali europei in merito allo sviluppo di tali sistemi e alle implicazioni socio-economiche, etiche e legali. Risalgono infatti ad aprile 2018 la Dichiarazione sulla cooperazione in materia di intelligenza artificiale¹ e la Comunicazione: L'intelligenza artificiale per l'Europa², i primi passi che l'Unione ha fatto per definire la propria posizione in un ordine globale sempre più antropocentrico dei sistemi di IA. Per parlare di intelligenza artificiale è importante capire di cosa si tratti. Secondo la sopraccitata Comunicazione (2018) 237, l'IA "indica sistemi che mostrano un comportamento intelligente analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere specifici obiettivi". Siamo dunque di fronte ad una tecnologia che, unita alla capacità e sempre maggiore potenza di calcolo, può aiutare l'essere umano nella risoluzione delle più grandi sfide, dalla medicina ai cambiamenti climatici, dall'istruzione all'economia. È importante però che lo sviluppo di tali sistemi adotti un approccio umano-centrico, rispettando i valori etici e i diritti universali quali la dignità umana, la protezione della privacy e la non-discriminazione. La Rome Call for AI ethics³ recentemente promossa dalla Pontificia Accademia per la Vita afferma appunto che l'IA debba essere a servizio dell'uomo (inteso come essere umano) e non mirata alla sua sostituzione. La necessità è dunque di disporre di linee guida e regolamentazioni il prima possibile poiché, citando la Comunicazione 237/18, "Il modo in cui ci relazioniamo all'IA determinerà il mondo in cui viviamo".

Documenti quali la Rome Call for AI ethics e il Libro Bianco sull'Intelligenza Artificiale⁴ ruotano proprio attorno a questo imperativo: la tecnologia deve essere al servizio dell'ambiente e degli uomini per migliorarne la qualità della vita e per risolvere sfide globali, garantendo allo stesso tempo il pieno rispetto dei diritti umani, dei valori e dei principi etici. La Rome Call for AI ethics, sottoscritta a febbraio 2020 dalla Pontificia Accademia per la vita, Microsoft, IBM, FAO e Governo Italiano, si concentra su 3 aree, ossia l'etica, l'istruzione e i diritti, sottolineando la necessità che l'IA venga sviluppata nel rispetto di libertà ed uguaglianza alla base dei diritti umani, che si occupi di trasformare il futuro partendo dall'istruzione delle giovani generazioni e che venga regolata affinché l'uomo e l'ambiente siano tutelati. Per raggiungere questi scopi la Rome Call for AI ethics identifica sei

¹ Bruxelles, 10 aprile 2018.

² COM (2018) 237 final.

³ Roma, 28 febbraio 2020.

⁴ COM (2020) 95 final.



principi da tenere in considerazione quando si tratta di intelligenza artificiale: trasparenza, inclusione, responsabilità, imparzialità, affidabilità e sicurezza, e privacy. Analogamente, il Libro Bianco sull'Intelligenza Artificiale si occupa di evidenziare l'imperativo che l'IA non sovrasti l'uomo ma anzi lo affianchi e gli sia di aiuto, rispettando i valori etici e i diritti umani. Centrale in questo documento della Commissione Europea è anche la definizione di un approccio comune a tutti gli Stati Membri per evitare la frammentazione del mercato unico. A tal proposito, la Commissione propone la creazione di un "ecosistema di eccellenza" e di un "ecosistema di fiducia". Mentre il primo si concentra, tra le altre cose, sulla necessità di collaborazione tra gli Stati Membri dell'Unione, sull'importanza di un allineamento dei centri di ricerca e innovazione sparsi all'interno dell'Unione e di un adeguato programma di finanziamento, e sulla garanzia di accesso ai dati, il secondo identifica sette requisiti fondamentali per l'elaborazione di un quadro normativo per l'IA, ossia intervento e sorveglianza umani, robustezza tecnica e sicurezza, riservatezza e governance dei dati, trasparenza, diversità, non discriminazione ed equità, benessere sociale e ambientale, e accountability.

L'obiettivo di questo paper è considerare i rischi conseguenti al mancato rispetto dei diritti umani fondamentali e dell'etica nello sviluppo e nell'applicazione di strumenti di intelligenza artificiale in particolare nell'ambito giudiziario. A livello europeo tale rischio è stato evidenziato nel Libro Bianco sull'Intelligenza Artificiale, in relazione alla vulnerabilità dei diritti "alle libertà di espressione e di riunione, la dignità umana, la non discriminazione fondata sul sesso, sulla razza, sull'origine etnica, sulla religione o sulle convinzioni personali, sulla disabilità, sull'età o sull'orientamento sessuale (ove applicabili in determinati settori), la protezione dei dati personali e della vita privata o il diritto a un ricorso giurisdizionale effettivo e a un giudice imparziale, nonché la tutela dei consumatori".

Le necessità evidenziate nel Libro Bianco vanno di pari passo con le linee guida illustrate precedentemente nella Comunicazione: L'intelligenza artificiale per l'Europa del 2018. In particolare, nella strategia europea di sviluppo dell'IA venivano delineati tre filoni: dare impulso alla capacità tecnologica e industriale dell'UE e all'adozione dell'IA in tutti i settori economici, prepararsi ai cambiamenti socio-economici, e assicurare un quadro etico e giuridico adeguato.

È proprio su quest'ultimo aspetto che si concentrerà l'analisi di questo elaborato; non solo il rispetto dell'etica ma anche la necessità di una regolamentazione a livello europeo appropriata ai futuri sviluppi dell'IA. Prima, verrà delineato l'approccio europeo all'IA al servizio delle forze dell'ordine e della magistratura. Successivamente, un confronto con il sistema statunitense porterà alla luce esempi reali dei rischi sopraccitati.



1. Contesto

Il concetto di Intelligenza Artificiale comprende un ampio spettro di sistemi di diverso tipo che si avvalgono di uno o più algoritmi per risolvere un problema o raggiungere un obiettivo. Il più importante e diffuso sottogruppo di modelli di IA è il cosiddetto machine learning: con questo termine si intendono quei modelli che funzionano con un algoritmo allenato su un set di dati che migliora quando riceve riscontri umani. I dati di cui necessita il machine learning vengono elaborati da una particolare disciplina, detta data science, che si occupa di clusterizzare⁵ i dati e di ordinarli prima del loro utilizzo per allenare l'algoritmo. Tali dati definiti inferred vengono creati sulla base dei dati forniti, ossia quelli ceduti in modo consapevole e attivo dall'interessato, e dei dati osservati, ossia quelli che l'interessato fornisce attraverso l'utilizzo di un dispositivo o la fruizione di un servizio.

Una particolare e complessa tipologia di machine learning è il deep learning: sistemi composti da più algoritmi che creano delle reti neurali artificiali imitando le sinapsi del cervello umano. Il deep learning dunque non è altro che l'aggregazione di più algoritmi che permette al modello di elaborare i dati in modo non lineare. Ciò che differenzia le due tipologie di IA è principalmente l'autonomia con cui operano: mentre il machine learning richiede dati strutturati per essere allenato e migliora di volta in volta con l'utilizzo e l'input umano (apprendimento supervisionato), il deep learning al contrario è alimentato con dati non strutturati e non richiede l'intervento umano per potersi migliorare (apprendimento automatico). Il comune denominatore tra queste due tipologie di intelligenza artificiale è l'importante quantità di dati di cui entrambi necessitano per poter essere allenati, i cosiddetti big data. Per poter processare le enormi quantità di dati è richiesta un'adeguata potenza di calcolo. L'Unione Europea in questo fronte è impegnata con diversi progetti come Leonardo, Horizon 2020, e Digital Europe 2021-2027 nel finanziamento di importanti programmi di ricerca e innovazione.

La diffusione sempre più ampia dei sistemi di intelligenza artificiale, dal settore industriale (robotica) al commercio, dagli smartphone alla difesa, è arrivata fino all'ambito giudiziario. Pertanto, è necessario chiarire come l'IA trova spazio nelle questioni giuridiche e con quale grado di autonomia può operare.

È dalla natura dei dati inseriti, dagli algoritmi sfruttati per la loro elaborazione e per l'apprendimento e dal risultato di tale processo che dipende la differenza tra i diversi strumenti di intelligenza artificiale. In particolare, il settore giuridico può avvalersi di 3 tipologie di IA: sistemi per l'analisi e la predisposizione automatica degli atti, sistemi capaci di prevedere l'esito di una causa e sistemi in

⁵ Con il termine "clusterizzare" si intende catalogare i dati raggruppandoli in base a caratteristiche simili.



grado di formulare giudizi⁶. Il primo modello si adatta molto bene ai meccanismi ripetitivi sui quali si basano il diritto civile e commerciale poiché consente l'analisi di importanti quantità di documenti. Di recente però si è pensato a come ampliare l'utilizzo di tali sistemi anche alla giustizia penale. In Italia, ad esempio, è stato sviluppato TOGA, un modello allenato su un database contenente le più di 5000 fattispecie di reato disciplinate dal Codice Penale e dalla legislazione speciale, che è in grado, tra le altre cose, di calcolare il tipo di pena ed eventuali pene alternative, verificare la procedibilità e valutare la durata delle misure cautelari. Un caso concreto di utilizzo di IA è in relazione alle indagini per il crollo del Ponte Morandi di agosto 2018, per le quali la Procura di Genova si è avvalsa di un sistema dell'FBI per analizzare ed incrociare la mole ingente di dati raccolti.

Il secondo modello predittivo si occupa di elaborare previsioni attraverso un calcolo probabilistico su base statistica o logica. Le sue applicazioni sono molteplici così come le implicazioni che ne conseguono: può essere utilizzato con uno scopo di prevenzione mediante l'individuazione di luoghi e orari in cui la probabilità che si verifichino dei reati è maggiore; trova spazio anche nell'interpretazione della legge per facilitare ed affiancare il lavoro del giurista; può essere utilizzata per prevedere l'esito di un giudizio. Quest'ultima applicazione si basa sul concetto di "diritto calcolabile" e di "certezza del diritto", due pilastri del sistema penale del nostro paese che stabiliscono la prevedibilità dell'esito e pertanto si ritiene che le macchine potrebbero efficacemente essere utilizzate per aiutare l'avvocato a non compiere mosse sbagliate durante il processo. Tuttavia, l'assoluta certezza dell'esito è impossibile da raggiungere poiché si dovrebbe riuscire a creare dei sistemi in grado di governare l'imprevedibilità, l'incertezza e il fattore umano che costituiscono il ragionamento giuridico che "deve essere rigoroso ma non può essere meccanico"⁷.

Il terzo modello di IA applicabile alla giustizia è quello capace di formulare un giudizio. Quest'ultimo tipo di sistema però solleva numerosi dubbi e problematiche legate ai diritti e all'etica: può il machine learning sostituire il giudice ed emettere una sentenza? Nel diritto civile e tributario questi sistemi si sono rivelati efficaci e capaci di risolvere i contenziosi con tempistiche ben ridotte rispetto al giudizio ordinario. Più controversa e problematica è invece la loro applicazione nell'ambito penale. In primis, una macchina difficilmente è capace di accertare la veridicità di una testimonianza, decisione che viene lasciata pertanto ad un giudice umano. Secondariamente, un algoritmo non può valutare la pertinenza e la validità degli indizi forniti dalle parti in causa e non può determinare che essi siano "gravi", "precisi" e "concordanti"⁸ per poter comprovare l'esistenza di un fatto. Inoltre, la sentenza

⁶ Traversi A., Intelligenza artificiale e giustizia, verso un giudice robot?, 2019, altalex.com

⁷ Pascuzzi G., Diritto e previsione, 2021.

⁸ Come prescritto dall'art. 192, comma 2, del codice di procedura penale.



emessa da una macchina è certa e non può ammettere alcun dubbio. Questo, però, va contro la necessità che il soggetto sia definito colpevole “al di là di ogni ragionevole dubbio” che un algoritmo non può tenere in considerazione. In ultimo, l’art. 22 del Regolamento UE 679/16 prevede che “l’interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato”.

2. Unione Europea, giustizia e IA

Come già discusso nella sezione introduttiva di questo paper, la questione dell’etica nello sviluppo di sistemi di intelligenza artificiale è di primaria importanza, tanto più quando si tratta di modelli utilizzati nell’ambito giudiziario. Proprio a questo scopo la Commissione Europea per l’Efficienza della Giustizia (CEPEJ) ha adottato la Carta etica europea sull’utilizzo dell’intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi⁹ all’interno della quale vengono identificati cinque principi cardine: rispetto dei diritti fondamentali, come il diritto all’autodeterminazione o ad un equo processo; non-discriminazione su base etnica, politica, religiosa o di genere; qualità e sicurezza dei dati utilizzati per allenare l’algoritmo; trasparenza, imparzialità ed equità in equilibrio con l’accessibilità dei dati; controllo da parte dell’utilizzatore, o *under user control*, per limitare l’eccessivo automatismo.

La Carta etica sottolinea come nel contesto europeo la diffusione dei sistemi di intelligenza artificiale nell’ambito della giustizia sia ancora molto limitata nel settore pubblico, mentre l’interesse nel privato sembra essere in crescita. Francia, Lettonia ed Estonia sono i paesi dell’Unione più avanzati in questo campo, dove vediamo l’impiego di sistemi di IA con scopi di analisi e previsione.

Tutta l’intelligenza artificiale ruota attorno ai dati necessari per far funzionare gli algoritmi. Il primo ostacolo, dunque, riguarda la loro accessibilità, detta anche apertura dei dati. La Carta etica specifica in primo luogo che i dati aperti concernono solamente i dati grezzi, o *raw data*, contenuti in banche dati strutturate. L’analisi e l’aggregazione di tali dati costituisce i *big data* di cui si nutrono i modelli di IA. È necessario, tuttavia, trovare un equilibrio tra l’accessibilità ai dati e la protezione di dati sensibili. In particolare, i dati personali delle parti e dei testimoni prima di diventare aperti subiscono un processo automatico di anonimizzazione e pseudonimizzazione per evitare rischi di discriminazione e profilazione, mentre la questione è più complessa quando si tratta dei nomi dei professionisti, poiché da un lato è bene che non siano oscurati per garantire la trasparenza delle decisioni ma dall’altro lato i giudici vedono venire meno la propria privacy. La risposta a tali

⁹ CEPEJ(2018)14, Strasburgo, 3 dicembre 2018.



problematiche però non può essere trovata a livello comunitario poiché i paesi membri dell'Unione Europea hanno in vigore diverse tipologie di sistemi giudiziari e pertanto spetta ad ogni singolo stato trovare un equilibrio tra trasparenza e privacy nell'ambito dei dati aperti giudiziari.

A tal proposito, la Carta Etica sottolinea la necessità di tenere a mente l'importanza della protezione dei dati, carburante di tali sistemi, sia in “fase di progettazione” che “per impostazione predefinita” degli algoritmi per evitare che il loro utilizzo leda la libertà e i diritti fondamentali dei soggetti. In particolare, ogni individuo ha diritto a non essere sottoposto a valutazioni basate solo su giudizi automatizzati, ad ottenere informazioni in merito al metodo di trattamento dei dati impiegato dall'algoritmo, ad opporsi a tale trattamento, e a richiedere un ricorso giurisdizionale. Pertanto, il rispetto della cyberetica non solo da parte degli utilizzatori dei sistemi di IA, ma anche da parte degli sviluppatori informatici e dei ricercatori è un imperativo se si vuole garantire la protezione dei big data e il rispetto di trasparenza, correttezza e neutralità di tali sistemi.

Uno dei principi cardine dell'approccio europeo alla giustizia è l'individualizzazione della pena, la quale, in contrasto con i sistemi giudiziari di common law, garantisce che a ogni soggetto venga inflitta una pena utile ed efficace, basata sullo studio della sua personalità e del grado di probabilità di recidiva/pericolosità sociale, che abbia come scopo ultimo non quello di punire ma piuttosto di rieducare con la prospettiva di reinserire in società il soggetto colpevole di reato. Il rischio è che venga attuato un approccio deterministico che impedisce l'analisi del soggetto per l'elaborazione di una pena individualizzata. Interessante è invece l'utilizzo di modelli di IA a supporto del lavoro degli esperti per gestire e raccogliere in modo più rapido e pratico le informazioni necessarie per la valutazione del soggetto.

La Carta Etica identifica due macro tipologie di sistemi di IA: quelli impiegati per le attività investigative e quelli utilizzati durante il processo penale.

Per quanto concerne i primi, si tratta di modelli di IA di larga diffusione utili non solo per la prevenzione dei reati mediante l'individuazione di “punti caldi” ma anche per il perseguimento degli attori di reato. Entrambe le tipologie sono considerate molto efficaci e abbastanza diffuse. I sistemi volti alla prevenzione sono chiamati predictive policing tools o strumenti di polizia predittiva e si occupano di individuare zone e orari in cui è maggiore la probabilità di commissione di certi tipi di reati come furti e aggressioni. I dati con cui vengono allenati questi sistemi sono lo storico delle localizzazioni di tali reati poiché si verificano con certi elementi di regolarità che permettono l'individuazione dei cosiddetti hot spots o punti caldi dove aumentare la sorveglianza da parte delle forze dell'ordine. Tale metodo è definito predictive criminal mapping poiché elabora una mappa di zone a rischio prevedendo il reato prima che venga commesso. I sistemi X-LAW e KeyCrime sono



due esempi concreti di tale tipologia di sistemi di IA utilizzati dalle forze dell'ordine italiane per contrastare crimini che vengono commessi seguendo un certo pattern come le rapine. Per quanto concerne X-LAW, si tratta di un modello di algoritmo che opera su base probabilistica testato dal 2004 in alcune zone di Napoli, esteso poi a tutta la città nel 2011 e successivamente esportato a Prato, Salerno, Venezia e Parma. Questo sistema è in grado di elaborare dei modelli criminali che si configurano ciclicamente analizzando e incrociando le caratteristiche socio-ambientali con i crimini commessi quotidianamente. KeyCrime invece è una startup nata nel 2009 a Milano che utilizza il software di analisi dei crimini “delia”, il quale, attraverso l'apprendimento automatico, prevede zone e orari in cui la probabilità di commissione di rapine è maggiore. Le prospettive future sarebbero di ampliamento del software ad altre città e ad altre fattispecie di reato come furti, violenze sessuali e truffe.

Di altra natura sono invece i modelli di IA utilizzati per rendere più efficace l'individuazione di attori e vittime di reati. Quest'ultimi hanno lo scopo di analizzare grandi quantità di dati per trovare delle “anomalie” che possano costituire dei reati, ad esempio nel settore finanziario o del turismo sessuale. Gianos, ad esempio, è un software utilizzato da moltissime banche italiane fin dal 1994 per analizzare operazioni finanziarie sospette e individuare possibili casi di riciclaggio.

Un ulteriore potenziale utilizzo di sistemi di AI nelle fasi investigative potrebbe essere durante l'interrogatorio in ausilio all'interrogante ad esempio per l'identificazione e la clusterizzazione di elementi di comunicazione para-verbale riconducibili alla menzogna.

I principali limiti di questa prima tipologia di sistemi sono stati individuati nella previsione di reati senza track record, ossia quelli che non vengono commessi con regolarità in determinate zone a orari precisi, come attacchi terroristici o omicidi; nella creazione di un circolo vizioso o delle cosiddette “profezie che si auto-adempiono” secondo cui intensificando i controlli delle forze dell'ordine in alcune zone vengono registrati più reati, i quali costituiscono i data set con cui vengono allenati gli algoritmi che perciò classificheranno quelle zone come sempre più a rischio; in ultimo, gli strumenti di predictive policing si concentrano sulla necessità di contrastare il crimine individuando zone e orari a rischio maggiore piuttosto che sulla comprensione e analisi delle condizioni ambientali che favoriscono la commissione di reati.

La seconda tipologia di sistemi di IA riguarda quelli impiegati durante il processo penale. È importante sottolineare in primo luogo che l'approccio europeo in questo campo si limita ad adoperare modelli di machine learning per affiancare giudici e avvocati, risparmiando tempo e agevolando il processo decisionale, evitando invece di delegare a questi sistemi la possibilità di prendere decisioni in modo completamente autonomo. I modelli utilizzati con lo scopo di prevedere



la durata di un procedimento, di calcolare la probabilità di risoluzione o la stima dell'importo di un risarcimento, o ancora l'orientamento delle Corti, sono definiti nel gergo tecnico “giustizia predittiva”, “giustizia algoritmica” o “giustizia automatizzata”. All'interno della Carta Etica europea l'unico esempio citato di tali sistemi nell'UE è HART (Harm Assessment Risk Tool) impiegato nel Regno Unito per calcolare il rischio di recidiva, citando la CEPEJ(2018)14, con “un valore puramente consultivo”. Dati i recenti sviluppi in merito alla membership britannica nella UE, ritengo sia doveroso citare altri esempi di modelli predittivi attualmente in uso in Francia e Italia. Per quanto riguarda il sistema francese, Predictice¹⁰ è una startup creata nel 2016 da un team di professionisti del campo legale e digitale che permette di ottimizzare la ricerca di documenti legali da un database che viene costantemente aggiornato con le decisioni dei tribunali, di calcolare le probabilità di risoluzione di una controversia in un determinato tribunale e di stimare l'importo del risarcimento. Nel nostro Paese, invece, i progetti in corso si riferiscono ai Tribunali di Brescia e Genova, dove l'Università di Brescia e la Scuola Sant'Anna di Pisa hanno contribuito per l'elaborazione di sistemi di IA in grado di prevedere la durata di un procedimento e di valutare l'orientamento dei diversi uffici. È importante dunque sottolineare che l'approccio europeo all'utilizzo dell'intelligenza artificiale in materia di giustizia è da considerarsi con il solo scopo di analizzare, valutare e semplificare il lavoro del legale o del giudice e mai con l'obiettivo di sostituire la componente umana nel procedimento.

3. Le sfide della “giustizia predittiva”

Diverse sono le implicazioni e le sfide riscontrate nell'utilizzo dell'IA nell'ambito della giustizia. La Carta Etica europea sottolinea quanto gli algoritmi siano efficaci nell'affiancare il lavoro dei professionisti legali e quanto in un futuro non troppo lontano possano velocizzare e facilitare il procedimento giudiziario, ma riscontra anche diverse difficoltà. Afferma infatti la necessità di un gruppo multidisciplinare di esperti per lo sviluppo di tali sistemi, i quali possano unire le competenze scientifiche dei programmatori con quelle accademiche dei professionisti del diritto, poiché solo con un approccio ampio e sfaccettato si potranno individuare e affrontare le differenti sfide poste dall'utilizzo della tecnologia nella giustizia per sviluppare dei sistemi efficaci ed etici. Proprio per quanto concerne l'etica di tali modelli di IA, la Carta Etica evidenzia delle criticità come la mancanza di inclusione del contesto sociale e ambientale che può influenzare in misura massiccia la commissione di un crimine o l'incapacità di stabilire una gerarchia di priorità nel calcolo della probabilità di recidiva.

¹⁰ <https://predictice.com/>



Come già discusso precedentemente in questo paper, un'altra sfida che l'IA pone è quella dell'accesso ai dati. In particolare, la Carta Etica porta alla luce un'importante differenza tra l'Unione Europea e gli Stati Uniti in merito al diritto di accesso agli algoritmi: mentre per l'UE si tratta del diritto imprescindibile di essere informati sul tipo di algoritmo utilizzato, negli USA non è pienamente riconosciuto questo diritto poiché andrebbe a scontrarsi con la protezione della proprietà individuale. L'approccio statunitense apre un'ulteriore questione, ossia quella della proprietà degli algoritmi da parte delle società private che li sviluppano. Questa proprietà però non dovrebbe far venire meno i principi di trasparenza e legittimità del trattamento dei dati personali degli interessati, i quali, in ogni momento, dovrebbero poter rivendicare i propri diritti di non essere sottoposti a decisioni automatizzate, di ottenere informazioni sull'algoritmo utilizzato per il trattamento dei dati, di opporsi a tale trattamento e di richiedere un ricorso.

Altre interessanti sfide sono portate alla luce da Claudio Castelli¹¹, Presidente della Corte d'Appello di Brescia, uno dei due Tribunali citati in precedenza per l'impegno in progetti di IA. Castelli sottolinea come gli algoritmi diano molta importanza ai precedenti giurisprudenziali, rischiando di avvicinare il nostro sistema di civil law a quello statunitense di common law, nel quale i precedenti sono alla base del sistema giudiziario. Tuttavia, afferma che l'introduzione un certo di grado di prevedibilità del diritto nel nostro ordinamento non sarebbe da considerarsi negativa in un sistema imprevedibile che spesso non tiene conto del precedente. Castelli cita tra i rischi il concetto di "autorealizzazione" o "effetto performativo": una sorta di circolo vizioso in cui il sistema produce degli stessi risultati in uscita che influenzano le decisioni dei giudici, le quali saranno poi a loro volta "carburante" per l'algoritmo.

A tal proposito è necessario evidenziare l'estrema importanza dei dati con cui vengono allenati i sistemi di IA, poiché è sulla base di questi che tali modelli produrranno dei risultati. Tuttavia, gli algoritmi sono creati dagli uomini per gli uomini e pertanto non sono immuni da discriminazioni, pregiudizi, errori. È proprio intorno al problema della discriminazione che ruoterà il confronto tra l'approccio europeo illustrato finora e la posizione statunitense in merito all'utilizzo dell'IA nella fase decisionale del processo. Gli algoritmi, infatti, producono risultati sulla base dei dati immessi per allenarli.

Ma cosa succede se tali dati riportano pattern¹² di discriminazione su base etnica, sessuale, o religiosa? In tal caso, l'algoritmo anziché elevarsi oltre pregiudizi e soggettivazioni che possono

¹¹ Castelli C., Giustizia predittiva: così a Brescia l'intelligenza artificiale aiuta a prevedere tempi e esiti delle cause, 20 Agosto 2019, [agendadigitale.eu](https://www.agendadigitale.eu)

¹² Con il termine "pattern" si intende uno schema o un modello.



verificarsi nella decisione presa da un giudice umano, garantendo una sentenza neutra e oggettiva, come i sostenitori dell'IA affermano, riproduce e addirittura legittima le disuguaglianze precedentemente esistenti nel sistema della giustizia.

4. Discriminazione algoritmica in UE

Nel 2020 la Commissione Europea ha pubblicato un report¹³ prodotto dal network europeo di esperti legali in uguaglianza di genere e non discriminazione in merito alla cosiddetta *algorithmic discrimination* o discriminazione algoritmica. Poiché il termine *bias*¹⁴ ha un significato di ampio spettro, gli autori dello studio prediligono la dicitura “discriminazione algoritmica” per intendere solo quei bias degli algoritmi che si scontrano con le leggi di uguaglianza di genere e non discriminazione. In primo luogo, viene fatta chiarezza sulle tre fasi di progettazione, sviluppo e utilizzo dell'algoritmo in cui possono comparire delle discriminazioni. Il primo livello è il *planning stage*, ossia l'identificazione di un obiettivo, del tipo di algoritmo che più si adatta a tale scopo e del data set che si vuole utilizzare. Successivamente, durante il *development stage*, i programmatori si occupano prima di creare la sequenza di codici che formerà l'algoritmo, per poi allenarlo sul data set selezionato e valutare la qualità dei risultati prodotti. Infine, nel *decision making and use stage* nuovi dati vengono elaborati dal modello. È importante che la supervisione e il monitoraggio siano costanti in tutte le tre fasi di nascita, sviluppo e utilizzo dell'algoritmo poiché, come illustrato a seguire, le discriminazioni possono avere luogo non solo nel *decision making* ma anche nelle due fasi precedenti.

È bene non pensare agli algoritmi come a strumenti totalmente scollegati da noi umani poiché, sebbene capaci di operare con un certo grado di autonomia, sono sempre un prodotto dell'intelligenza umana. Pertanto, è nel ragionamento umano da individuare il primo pericolo di discriminazione negli algoritmi. Questi modelli, infatti, riflettono la mente di chi li pensa prima e di chi li produce poi, perpetuando non solo nobili valori umani ma anche stereotipi, pregiudizi e discriminazioni.

Inoltre, gli autori del report sottolineano anche la presenza di bias cognitivi che si attivano quando l'uomo viene assistito dall'IA, come ad esempio il bias di automazione, per cui è stato rilevato che l'uomo tende a credere che l'algoritmo abbia una conoscenza molto più ampia e migliore della nostra,

¹³ Gerards J. (Utrecht University) and Xenidis R. (University of Edinburgh and University of Copenhagen), *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law*, European Commission, 2020

¹⁴ Con il termine “bias” si intende un “errore di interpretazione, giudizio arbitrario o distorto”.
<https://www.garzantilinguistica.it/ricerca/?q=bias>



tendendo a fidarsi di più dell'output di un'analisi di un algoritmo piuttosto che della stessa analisi prodotta da un uomo.

Successivamente, un altro elemento che può contenere importanti elementi di discriminazione è il data set, ovvero i dati selezionati nella fase di planning con cui l'algoritmo viene allenato in fase di sviluppo. La sfida dei dati riguarda ad esempio la veridicità delle informazioni, la rappresentatività dei dati che possono non essere comprensivi di tutta la popolazione escludendo delle minoranze e la qualità di tali dati. L'espressione inglese rubbish in, rubbish out (spazzatura in entrata, spazzatura in uscita) sottolinea quanto la qualità, la precisione e l'affidabilità dei dati in entrata siano imprescindibili per avere un output corretto.

Durante la fase di allenamento dei modelli, gli algoritmi vengono "nutriti" con grandi quantità di dati, i big data, nei quali individuano correlazioni e pattern che verranno applicati nell'analisi di nuovi dati. Questo aspetto si ricollega alla qualità dei training data poiché se essi contengono schemi ricorrenti di disuguaglianza e correlazioni discriminatorie, il modello non farà altro che rinforzarli e riprodurli. Perciò, per evitare che l'algoritmo riconosca delle correlazioni causali tra elementi discriminatori, si tende ad evitare caratteristiche quali sesso, etnia, o credo religioso. Tuttavia, i modelli di IA sono capaci di individuare le cosiddette variabili proxy¹⁵, ossia elementi apparentemente irrilevanti riconducibili però a caratteristiche come il genere o il colore della pelle che portano in conclusione ad un output tendenzialmente pregiudizievole.

Un ulteriore elemento importante per limitare bias e analisi non eque è la trasparenza: spesso il modo in cui operano gli algoritmi è considerato opaco e di difficile comprensione per coloro non esperti in materia. Si fa riferimento a questa mancanza di chiarezza in merito al funzionamento dei sistemi di IA con il termine black box o scatola nera, la quale fa sì che nella maggior parte dei casi un individuo sottoposto ad una decisione algoritmica non comprenda se il risultato dell'analisi sia "infettato" da bias di genere, etnia o qualunque altro tipo di discriminazione.

Le fasi di decision making e uso dei modelli di IA hanno un forte impatto anche sulla portata e sulla rapidità dell'analisi dei dati. L'enorme quantità di dati processati e l'elevata velocità di analisi sono i vantaggi più importanti dell'utilizzo di sistemi di IA, i quali permettono di elaborare molti dati e di prendere una decisione in tempi davvero brevi. Allo stesso tempo però, se nelle fasi di planning e sviluppo dovessero verificarsi alcune delle problematiche appena illustrate, tali discriminazioni

¹⁵ Con il termine "variabile proxy" si intende "una variabile che non risulta di per sé rilevante all'interno di un modello, ma sostituisce un'altra variabile non osservata o che non può essere misurata. La correlazione non deve essere necessariamente lineare, e può essere sia positiva che negativa." <https://amp.it.google-info.cn/8295532/1/proxy-statistica.html>



verrebbero applicate ad una mole massiccia di dati in pochi istanti. Velocità e portata di analisi dunque possono rivelarsi un'arma a doppio taglio che espande e accelera la diffusione di output non equi.

In ultimo, è bene ricordare anche la sfida della responsabilità di una decisione discriminatoria prodotta dall'algoritmo che si trova al centro di un importante dibattito in merito all'accountability dell'output di un modello di IA.

Questo ampio quadro delle principali sfide che riguardano la discriminazione e l'intelligenza artificiale proposto dagli esperti legali per la Commissione Europea, si conclude con una promettente soluzione. Gli autori del report propongono un quadro normativo integrato poiché l'attuale regolamentazione europea per la gender equality e la non-discrimination si concentra limitatamente alle discriminazioni per genere ed etnia, non considerando molte altre caratteristiche che possono portare ad un diverso trattamento di un individuo, come età, disabilità, orientamento sessuale e religione. È proprio per sopperire a tale limite che gli esperti propongono "PROTECT", ossia un approccio multidisciplinare che combina soluzioni legali, tecnologiche e knowledge-based. L'acronimo è composto da sette raccomandazioni e linee guida da adottare in tutte le tre fasi di planning, sviluppo e utilizzo di sistemi di IA. PROTECT sta dunque per prevent, prevenire discriminazioni fin dalla fase di design attraverso la creazione di gruppi di lavoro diversificati e qualificati; redress, combinare diversi strumenti legislativi in materia di non discriminazione per promuovere norme eque e rimedi giudiziari chiari; open, incoraggiare la trasparenza e l'apertura dei dati al fine di poter monitorare l'operato dell'algoritmo; train, educare e formare i programmatori e gli specialisti informatici in merito alle questioni di uguaglianza e non discriminazione, nonché sensibilizzare giudici, funzionari, regolatori e la società tutta; explain, istituire requisiti di trasparenza, responsabilità e informazione; control, coinvolgimento umano attivo (come ad esempio nei sistemi HITL o human in the loop) per evitare un'eccessiva automazione accanto a una costante supervisione; test, monitoraggio continuo degli output e creazione di meccanismi di ispezione, classificazione e certificazione degli algoritmi.

5. Discriminazione algoritmica negli USA

Per offrire un confronto costruttivo, ritengo sia importante portare alla luce l'attuale situazione in merito all'utilizzo dei sistemi di IA da parte della polizia e del sistema penale degli Stati Uniti.



In primis, è bene chiarire che l’approccio statunitense differisce da quello europeo per quanto riguarda la giustizia: negli USA è infatti in vigore un sistema di common law, incentrato sul precedente giurisprudenziale piuttosto che su un codice scritto dal legislatore, facendo sì che il diritto venga “scritto” dai giudici. Di fatto, il sistema di common law e lo studio della giurisprudenza basato sull’analisi di casi si presta molto bene all’utilizzo di algoritmi per calcolare la probabilità di recidiva, fissare la cauzione, e stabilire colpevolezza o innocenza. Così come precedentemente illustrato per l’UE, anche negli Stati Uniti gli algoritmi non affiancano solo i giudici e gli avvocati ma anche le forze di polizia con strumenti di predictive policing mediante hot spot (“zone calde” a rischio criminalità) e heat list (lista di individui pericolosi).

Per quanto riguarda la regolamentazione, tuttavia, negli USA si presenta un approccio frammentato: a livello federale tutti gli stati condividono la stessa costituzione, ma ognuno ha un diritto penale “indipendente”. Per questo motivo è difficile parlare degli Stati Uniti in generale poiché ogni stato ha le proprie regolamentazioni anche in merito all’utilizzo dell’IA nei tribunali. Nel 2018, lo stesso anno della pubblicazione della Carta Etica europea trattata in precedenza in questo paper, negli USA è stata promulgata dal New York City Council la prima legge in materia di trasparenza delle decisioni prese dagli algoritmi¹⁶. È stata avvertita dunque anche negli Stati Uniti la necessità di legiferare in merito a questi relativamente recenti sviluppi della tecnologia nella giustizia nonostante la legge sopraccitata sia in vigore solamente nello stato di New York.

Come già largamente mostrato nei capitoli precedenti, sia gli algoritmi utilizzati dalle forze di polizia che quelli impiegati nei tribunali richiedono i big data per poter essere allenati all’auto-apprendimento. L’affidabilità di un sistema di IA dipende quindi dalla qualità di tali dati. Purtroppo, però, gli USA non sono noti per l’assoluta equità del trattamento riservato da forze di polizia e giudici ai cittadini afroamericani. Prima di criticare e condannare gli algoritmi per essere discriminatori, perciò, sarebbe bene spostare l’attenzione sui dati con cui vengono allenati. Secondo il report pubblicato da The Sentencing Project nel 2018¹⁷, per i cittadini americani di origine afroamericana sono maggiori le probabilità di arresto, condanna e di ricevere pene più lunghe rispetto ai bianchi. In particolare, secondo i dati del Bureau of Justice Statistics relativi al 2016¹⁸, gli adulti afroamericani vengono incarcerati 5,9 volte di più dei bianchi e nel 2010 il 33% degli adulti neri aveva almeno una

¹⁶ <https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=3137815&GUID=437A6A6D-62E1-47E2-9C42-461253F9C6D0>

¹⁷ Report of The Sentencing Project to the United Nations Special Rapporteur on Contemporary Forms of Racism, Racial Discrimination, Xenophobia, and Related Intolerance Regarding Racial Disparities in the United States Criminal Justice System, The Sentencing Project, 19 Aprile 2018.

¹⁸ I dati più recenti del BJS si riferiscono al 2016.



condanna sulla fedina penale. Sebbene i cittadini afroamericani e di origine ispanica costituiscano solo il 29% del totale della popolazione statunitense, compongono più della metà della popolazione carceraria. Questi dati si traducono anche in isolamento e difficoltà di partecipazione nella società dopo aver scontato la pena poiché per molti impieghi è richiesta la fedina penale pulita, così come per avere un alloggio o ricevere gli ammortizzatori sociali, e nel sistema penitenziario degli USA non è previsto un trattamento individualizzato volto alla rieducazione e al reinserimento dei detenuti come avviene nel nostro Paese. Le disuguaglianze riportate da The Sentencing Project si manifestano non solo in tribunale, ma anche da parte delle forze di polizia e ancor prima da politiche che penalizzano le classi sociali più indigenti e quindi, di conseguenza, spesso riguardano le minoranze etniche. Ad esempio, tra le possibili soluzioni, il report propone l'interruzione della War on Drugs o guerra al narcotraffico che si concentra sui cittadini afro e ispanici facendo sì che il 56% dei condannati per reati legati agli stupefacenti siano proprio di queste origini. Inoltre, il report accenna anche a strumenti di risk-assessment precisi ed efficaci.

La sfida riguardo la qualità dei sistemi di IA dipende strettamente dai dati che si utilizzano per il loro allenamento, i quali però come abbiamo appena visto sono fortemente sbilanciati e discriminatori. Sistemi allenati con dati contenenti bias e disparità tra le etnie produrranno risultati pregiudizievole, che amplificheranno tali problematiche. Per quanto riguarda gli strumenti di predictive policing, l'identificazione di pattern e hot spot dove intensificare i controlli a causa delle alte probabilità che vengano commessi crimini legati alle droghe o rapine portano ad una maggiore presenza delle forze di polizia nei quartieri più periferici dove solitamente vivono cittadini di origine afro o ispanica, portando dunque a molti più arresti rispetto ai bianchi. Il pattern è bene ricordare che sono delle correlazioni che l'algoritmo identifica all'interno dei training data e non delle relazioni di causalità, pertanto quelli che un sistema di IA applica dipendono interamente dai dati utilizzati per il suo allenamento.

Inoltre, anche i sistemi di riconoscimento facciale hanno prodotto risultati discriminatori per genere ed etnia, producendo errori di identificazione dovuti alla sottorappresentazione di donne e minoranze etniche tra i training data¹⁹.

Per quanto riguarda gli algoritmi di risk-assessment, sono molti gli studi che riportano quanto i dati siano sbilanciati e non equi tra gli individui di diverse etnie. Negli USA vengono utilizzati sistemi di IA per le previsioni di probabilità di recidiva, per decidere in merito al rilascio su cauzione e per stabilirne l'importo, fino al giudizio di innocenza o colpevolezza. Come già illustrato in precedenza,

¹⁹ <https://www.jhunewsletter.com/article/2020/10/facial-recognition-technology-isnt-ready-for-police-use>



gli algoritmi non sono totalmente autonomi nel giudizio ma la presenza di bias cognitivi come quello dell'automazione fanno sì che i giudici si fidino ciecamente dei risultati prodotti dall'IA e ne siano fortemente influenzati. Il modello più conosciuto, utilizzato e controverso è COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) prodotto dall'azienda Northpointe. Questo sistema è in grado di stabilire non solo il livello di rischio di un individuo su una scala da 1 a 10 ma anche i cosiddetti bisogni criminogenici²⁰. Nonostante questo strumento sia largamente adottato negli USA, sono numerosi gli studi che ne criticano l'integrità. Secondo un report di valutazione statistica della validità di COMPAS pubblicato dalla stessa Northpointe nel 2009²¹, la percentuale di precisione dei suoi risultati sarebbe al 68%. Tuttavia, questo studio non tiene in considerazione le discrepanze nelle valutazioni tra individui di diverse etnie. Di questo invece si è occupato un report di ProPublica del 2016 dal titolo Machine Bias²², all'interno del quale sono riportati dati ed esempi delle diseguaglianze su base etnica degli assessment elaborati da COMPAS. In primo luogo, ProPublica rileva una precisione generale dell'algoritmo pari al 61% (più bassa di quella riportata nello studio di Northpointe). L'algoritmo si è dimostrato inaccurato nella previsione del rischio di commissione di reati in futuro sia per i neri che per i bianchi, con percentuali simili ma su risultati diametralmente opposti: tra gli individui valutati con rischio alto, il 23% dei bianchi non ha mai più commesso un crimine contro il 44,9% dei neri; al contrario, tra le valutazioni a rischio basso, solo il 28% dei neri ha commesso altri reati contro il 47% dei bianchi.

Inoltre, per i neri la probabilità di essere considerati ad alto rischio anche senza precedenti è maggiore del 77% rispetto ai bianchi. Le critiche avanzate verso l'azienda produttrice di COMPAS riguardavano il funzionamento dell'algoritmo e i parametri utilizzati per le valutazioni poiché, nonostante non venissero presi in considerazione l'origine etnica e il colore della pelle, alcune variabili proxy portavano a risultati sfavorevoli per gli afroamericani. La risposta di Brennan, uno dei fondatori dell'azienda Northpointe, fu che escludendo fattori come povertà, disoccupazione e marginalizzazione sociale i risultati non sarebbero stati sufficientemente precisi. Tuttavia, questi ulteriori indicatori sono senza alcun dubbio delle variabili proxy che senza essere esplicitamente legate all'etnia ne sono comunque correlate, portando a valutazioni non equilibrate tra individui bianchi e neri. Interessanti sono gli esempi riportati nello studio di ProPublica in merito: Brisha

²⁰ "I bisogni criminogenici, invece, sono fattori di rischio psicologici dinamici collegati direttamente al comportamento antisociale e alla sua reiterazione. Riguardano, dunque, aspetti dell'individuo e del contesto in cui vive che, se modificati, potrebbero facilitare cambiamenti nella carriera criminale."
<http://www.psychodesk.it/risk-assessment-la-valutazione-del-rischio-le-sue-applicazioni-nel-sistema-penale-italiano/>

²¹ <http://www.northpointeinc.com/files/publications/Criminal-Justice-Behavior-COMPAS.pdf>

²² <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>



Borden e Vernon Prater sono stati accusati entrambi di furto di beni per un valore di circa \$80 e sebbene Borden avesse solo piccoli precedenti, venne assegnato il grado 8 sulla scala di rischio di commissione di altri reati, mentre il livello di rischio di Prater fu stabilito a 3 nonostante tra i suoi precedenti ci fossero due rapine a mano armata. Borden, afroamericana valutata ad alto rischio, non commise più reati; Prater, bianco a basso rischio, commise successivamente un furto di grande valore. Tuttavia, sebbene ci siano molte associazioni impegnate contro la discriminazione nella giustizia e negli algoritmi, nella maggior parte degli stati USA si fa utilizzo di sistemi di IA dalle attività delle forze dell'ordine fino alle decisioni prese dai giudici.

Conclusioni

Sin dagli albori della tecnologia e del settore digitale, la fantasia dell'uomo si è immaginata gli scenari più apocalittici in cui le macchine prendono sempre più il controllo fino ad avere il sopravvento sulla specie umana. Sebbene non siamo ancora arrivati a situazioni di tale autonomia, le macchine oggi hanno le capacità per svolgere molte azioni prima riservate agli uomini, dalla produzione industriale alla chirurgia. Un ramo della tecnologia che solleva diverse questioni etiche e legali è quello dell'intelligenza artificiale. L'IA si basa sugli algoritmi, ossia delle sequenze di calcoli che hanno come scopo quello di risolvere un problema o raggiungere un obiettivo, e trova applicazioni in moltissimi ambiti sia della nostra vita digitale che di quella reale. Questo paper nello specifico ha voluto concentrarsi sull'utilizzo di sistemi di intelligenza artificiale da parte di forze dell'ordine e organi di giustizia. Sebbene tali modelli abbiano ormai raggiunto un certo grado di autonomia nelle decisioni, in questo ambito il loro compito è relegato all'ausilio e all'affrancamento del processo di decision making svolto dall'uomo. Non per questo però la loro influenza è minore poiché entrano in gioco i cosiddetti bias cognitivi nella psiche umana, come il bias di automazione che spinge l'uomo a credere che l'analisi di un algoritmo sia sempre più precisa e corretta di quella svolta da noi umani e a fidarsi quasi ciecamente dell'IA.

Come è stato largamente illustrato, ci sono diverse tipologie di sistemi di IA in base allo scopo, dai modelli di predictive policing a quelli di risk assessment. Ciò che accomuna tutti gli algoritmi e l'intelligenza artificiale sono i big data, ossia le grandi quantità di dati che costituiscono il "cibo" dell'algoritmo utilizzati per il suo allenamento. È dunque evidente che l'integrità, la qualità e la precisione dei risultati prodotti da un sistema di IA dipendano strettamente dai training data. L'allenamento di un algoritmo si basa sull'immissione di banche dati, all'interno delle quali l'IA identifica dei pattern e delle correlazioni che applicherà durante il suo utilizzo. Per evitare che un sistema produca risultati basandosi su elementi discriminatori come l'etnia, il genere o la religione,



si escludono tali parametri. Tuttavia, è stato dimostrato che le cosiddette variabili proxy come povertà, disoccupazione e marginalizzazione sociale, anche senza essere apparentemente legate all'etnia, producono risultati sbilanciati tra individui di diverse origini.

Per quanto riguarda l'Unione Europea, l'impiego dell'intelligenza artificiale è abbastanza diffuso come strumento di prevenzione del crimine (predictive policing) da parte delle forze dell'ordine attraverso l'individuazione di "zone calde" dove intensificare i controlli o con l'elaborazione di heat list di individui socialmente pericolosi. L'utilizzo di algoritmi nella giustizia europea è, invece, ancora limitato date le preoccupazioni circa l'etica e l'equità delle valutazioni. La Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi del 2018 e il report Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law del 2020 sono passi importanti che l'UE ha fatto verso lo sviluppo di un'IA etica e non discriminatoria.

Nel panorama della giustizia negli USA, dall'altra parte, sono molto frequenti non solo gli algoritmi a supporto delle forze dell'ordine ma anche nel giudizio. In quasi tutti gli stati statunitensi sono stati adottati dei sistemi di IA che affiancano i giudici nel processo decisionale. Numerosi sono tuttavia gli studi che criticano i risk assessment tools come COMPAS, un algoritmo che calcola il grado di pericolosità di un individuo per valutare quale tipo di pena assegnargli e la sua durata. Le critiche sollevate si riferiscono al rilevamento di importanti disparità di valutazione tra individui bianchi e afroamericani, i quali hanno più probabilità di essere arrestati, condannati e di scontare pene più lunghe, nonché di ricevere una valutazione di high risk.

In conclusione, per avere in un futuro non troppo lontano dei modelli di IA sempre più efficaci, etici e non discriminatori è necessario, che l'innovazione nel campo dell'intelligenza artificiale sia regolamentata verso uno sviluppo etico ed antropocentrico, che riesca a far sì che le macchine siano al servizio dell'uomo per aiutarlo a svolgere dei compiti in modo più veloce e preciso, senza però prenderne il posto. Inoltre, come è stato largamente trattato in questo paper, non bisogna sottovalutare l'importanza dei dati poiché è dalla loro integrità e qualità che dipendono i risultati dell'elaborazione algoritmica. I concetti chiave per scongiurare uno scenario apocalittico, in cui le macchine sostituiranno la specie umana, sono un framework legale solido e propenso all'etica, una maggiore attenzione alla qualità dei training data e un'importante capacità di calcolo adeguata all'elaborazione di grandi quantità di dati in modo sempre più rapido e preciso.