

LIBER AMICORUM
PER
PASQUALE COSTANZO

ELENA BASSOLI

**LA CIBER-ETICA: LUCI E OMBRE DELLA
PREDITTIVITÀ ALGORITMIZZATA**

17 APRILE 2020



Elena Bassoli

La Ciber-Etica: luci e ombre della predittività algoritmizzata

SOMMARIO: 1. I cinque principi della Carta Etica Europea 2018/2019. – 2. L'uso di sistemi AI nei sistemi giudiziari. – 3. Panoramica delle politiche relative agli *open data* relative alle decisioni giudiziarie nei sistemi giudiziari degli Stati membri del Consiglio d'Europa. – 4. La definizione di *open data* su decisioni giudiziarie. – 5. Lo sviluppo degli *open data* in materia giudiziaria negli Stati membri del Consiglio d'Europa e le conseguenze per lo sviluppo della giurisprudenza. – 6. Protezione dei dati personali negli *open data* delle decisioni giudiziarie. – 6.1. I nomi delle parti e dei testimoni. – 6.2 I nomi dei professionisti, inclusi i giudici. – 7. Caratteristiche operative del *machine learning* applicato alle decisioni giudiziarie. – 8. Le funzionalità teoriche del software di “giustizia predittiva”. – 9. Il funzionamento pratico dell'intelligenza artificiale: macchine statistiche che costruiscono modelli basati sul passato. – 10. L'intelligenza artificiale può modellare in anticipo il ragionamento legale? – 11. L'AI può spiegare il comportamento dei giudici *a posteriori*? – 12. Come si applica l'AI nella giustizia civile, commerciale e amministrativa. – 13. Un nuovo strumento di calcolo per calcolare i parametri di risarcimento. – 14. La risoluzione delle controversie online (ODR). – 15. Le principali garanzie da riaffermare nei procedimenti civili, commerciali e amministrativi. – Il diritto di accesso a un tribunale. – 15.2. Il principio del contraddittorio. – 15.3. Il principio della parità di armi. – 15.4. Il principio di imparzialità e indipendenza del giudice. – 15.5. Il diritto di assistenza legale. – 16. Le questioni specifiche della giustizia penale: prevenzione dei reati, rischio di recidiva e valutazione del livello di pericolo. – 17. Gli strumenti utilizzati dalle autorità investigative prima del processo penale. – 18. Gli strumenti per il processo penale. – 19. Le sfide della “predittività” in materia penale. – 20. Le questioni specifiche relative alla protezione dei dati personali. – 21. Il potenziale e i limiti degli strumenti di giustizia predittiva. – 22. La necessità di un dibattito pubblico per il rispetto dei diritti fondamentali. – 23. La creazione di un quadro etico.

1. I cinque principi della Carta Etica Europea 2018/2019.

L'intelligenza artificiale applicata al diritto è stata finalmente oggetto di attenzione da parte del Consiglio Europeo per mezzo della Commissione europea per l'efficienza della giustizia (CEPEJ).

L'organismo ha adottato la “Carta Etica europea sull'uso dell'intelligenza artificiale nei sistemi giudiziari e loro ambiente” a Strasburgo nella 31a riunione plenaria del 3-4 dicembre 2018, e l'ha aggiornata il 13 settembre 2019¹.

La Carta si apre con una dichiarazione di intenti esplicitata in cinque punti fondamentali che dovranno informare tutta la disciplina a livello europeo, e che ricordano in qualche modo le tre leggi della robotica di Asimov.

- 1. Principio del rispetto dei diritti fondamentali: assicurare che il progetto e l'implementazione di strumenti e servizi di intelligenza artificiale sia compatibile con i diritti fondamentali;
- 2. Principio di non discriminazione: prevenire specificamente lo sviluppo o l'intensificazione di ogni discriminazione tra individui o gruppi di individui;
- 3. Principio di qualità e sicurezza: per quanto riguarda il trattamento di decisioni e dati giudiziari, utilizzare fonti certificate e dati immateriali con modelli concepiti in modo multidisciplinare, in un ambiente tecnologico sicuro;
- 4. Principio di trasparenza, imparzialità ed equità: utilizzare metodologie accessibili e comprensibili per l'elaborazione dei dati, autorizzando audit esterni;
- 5. Principio “sotto il controllo dell'utente”: evitare un approccio prescrittivo e garantire che gli utenti siano informati e abbiano il controllo delle loro scelte.

La Commissione, lungi da un approccio statico, prescrive che la Carta sia soggetta ad una continua revisione al fine di una sempre migliore applicazione dei principi, anche ricercando ed analizzando i

¹ <https://www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-in-judicial-systems-and-their-environment>.

motivi della loro mancata o parziale attuazione, unitamente ad un piano d'azione per introdurre le misure necessarie.

Le autorità indipendenti menzionate nella Carta potrebbero valutare periodicamente il livello di approvazione dei principi della Carta tra tutti i soggetti coinvolti e proporre miglioramenti per adattarla al cambiamento tecnologico.

1) In virtù del primo principio il trattamento di decisioni e dati giudiziari deve avere scopi chiari, nel pieno rispetto dei diritti fondamentali garantiti dalla Convenzione europea dei diritti dell'uomo (CEDU) e della Convenzione di Strasburgo sulla protezione dei dati personali n. 108/1981, come modificata il 17-18 maggio 2018 dal protocollo di modifica n. 223 firmato ad Elsinore, in Danimarca.

In ossequio a tali discipline, pertanto, quando vengono utilizzati strumenti di intelligenza artificiale per risolvere una controversia, o come strumento di aiuto nel processo decisionale giudiziario, o per fornire assistenza al pubblico, è essenziale garantire che tali strumenti non minino le garanzie del diritto a un ricorso effettivo davanti a un giudice imparziale² e del diritto a un giusto processo, in virtù del principio di uguaglianza degli strumenti per difendersi e di rispetto del principio del contraddittorio.

L'Intelligenza artificiale trova i suoi limiti naturali, poi, nel rispetto del principio di legalità e nell'indipendenza dei giudici nel loro processo decisionale.

Andranno pertanto privilegiati approcci miranti all'etica umana *by design* fin dalla loro progettazione, ponendo cura a che ogni fase dell'apprendimento di sistemi di AI sia informato da regole che impediscano violazioni dirette o indirette dei valori fondamentali dell'uomo.

2) Sulla base del secondo principio di non discriminazione, data la capacità di questi metodi di elaborazione di rivelare informazioni discriminatorie, raggruppando o classificando dati relativi a individui o a gruppi di individui, occorrerà garantire che i metodi non riproducano o aggravino tale discriminazione.

Per far sì che ciò avvenga è necessario prestare particolare attenzione sia allo sviluppo in sé dello strumento di AI, sia alle fasi in cui lo sviluppo si snoda, in particolare quando l'elaborazione si basa direttamente o indirettamente su "dati "sensibili", come presunte origini razziali o etniche, background socioeconomico, opinioni politiche, credenze religiose o filosofiche, appartenenza a sindacati, dati genetici, dati biometrici, dati relativi alla salute o alla vita sessuale o all'orientamento sessuale.

Così, quando in un sistema di AI viene identificato un processo che conduce di fatto ad una discriminazione, debbono essere adottate misure correttive per neutralizzare questi rischi unitamente alla sensibilizzazione dei soggetti coinvolti.

3) In base al terzo principio di qualità e sicurezza, i progettisti di modelli di machine learning dovrebbero essere in grado di attingere ampiamente all'esperienza degli attori coinvolti nel sistema giudiziario: da un lato giudici, pubblici ministeri, avvocati, cancellieri, ufficiali giudiziari; dall'altro ricercatori e docenti nei settori del diritto e delle scienze sociali, quindi non solo giuristi, ma anche economisti, sociologi e filosofi.

L'approccio all'implementazione di un algoritmo, tanto più se dominato da logiche di Intelligenza artificiale che ne possono modificare il comportamento tramite auto-apprendimento, dovrà partire fin

² Così, riprendendo l'art. 13 della CEDU, la Carta dei diritti fondamentali dell'UE - Articolo 47 (Diritto a un ricorso effettivo e a un giudice imparziale). Ogni persona i cui diritti e le cui libertà garantiti dal diritto dell'Unione siano stati violati ha diritto a un ricorso effettivo dinanzi a un giudice, nel rispetto delle condizioni previste nel presente articolo. Ogni persona ha diritto a che la sua causa sia esaminata equamente, pubblicamente ed entro un termine ragionevole da un giudice indipendente e imparziale, precostituito per legge. Ogni persona ha la facoltà di farsi consigliare, difendere e rappresentare. A coloro che non dispongono di mezzi sufficienti è concesso il patrocinio a spese dello Stato, qualora ciò sia necessario per assicurare un accesso effettivo alla giustizia.

dalla fase di progettazione con un team di progetto di tipo interdisciplinare servendosi di tutte le professionalità coinvolte.

Il team stesso si trova quindi ad essere il primo validatore delle regole etiche, che condivise da tutti i soggetti coinvolti, andranno a vincolare il sistema di AI anche grazie a continue e progressive analisi e processi di feedback.

4) In virtù del quarto principio di trasparenza, imparzialità ed equità, un rilevante elemento di correttezza della procedura è costituito dal fatto che le decisioni giudiziarie inserite nel software che implementa un algoritmo di *machine learning* debbano provenire da fonti certificate e non debbano essere modificate fino a quando non siano state utilizzate dal meccanismo di apprendimento.

L'intero processo deve quindi essere tracciabile, così come i modelli e gli algoritmi creati devono essere memorizzati ed eseguiti in ambienti sicuri, in modo da garantire l'integrità del sistema e la sua inalterabilità.

Le misure di garanzia devono essere applicate, a valle, all'intera catena di progettazione e gestione; a monte al processo di selezione ed organizzazione dei dati che influenzano direttamente la fase di apprendimento.

5) Infine, con riferimento al quinto principio di controllo del sistema da parte dell'utente, la Carta rileva che, come noto, uno dei maggiori ostacoli alla conoscibilità del metodo di funzionamento dell'algoritmo sia costituito dai diritti di proprietà intellettuale degli inventori³.

La Carta Etica sul punto ribadisce la necessità di operare un equilibrio tra la proprietà intellettuale di alcuni metodi di elaborazione e la necessità di trasparenza, imparzialità, equità ed integrità intellettuale, dando la priorità agli interessi della giustizia, soprattutto quando vengono utilizzati strumenti che producono conseguenza giuridiche o che incidono in modo significativo sulla vita delle persone.

Esistono diverse alternative sul punto, ma l'opzione preferibile è quella orientata alla completa trasparenza tecnica mediante conoscibilità del codice sorgente e della relativa documentazione, ma, come visto, tale possibilità si scontra con i diritti di proprietà intellettuale e dei segreti commerciali.

Il sistema potrebbe anche essere spiegato in modo chiaro e con un linguaggio familiare che descriva come vengono prodotti i risultati, comunicando, per esempio, la natura dei servizi offerti, gli strumenti che sono stati sviluppati, le prestazioni e i rischi di errore. Autorità indipendenti o esperti potrebbero essere incaricati di certificare e verificare i metodi di elaborazione o fornire consulenza preventiva.

Poiché l'autonomia dell'utente deve essere aumentata, e non limitata, dall'uso di strumenti e servizi di intelligenza artificiale i professionisti del sistema giudiziario dovrebbero, in qualsiasi momento, essere in grado di mettere in discussione le decisioni giudiziarie assunte dall'algoritmo e i dati utilizzati per produrle senza esserne necessariamente vincolati alla luce delle caratteristiche specifiche della fattispecie concreta in decisione.

È fondamentale, poi, che in virtù del principio di trasparenza l'utente sia informato in un linguaggio chiaro e comprensibile in merito alla vincolatività o meno delle soluzioni offerte dagli strumenti di intelligenza artificiale, conoscendo le opzioni disponibili, e mantenendo fermo il diritto a ricorrere ad una consulenza legale e al potere di accedere comunque a un tribunale imparziale.

³ Cfr. p. 38 del Consiglio d'Europa Studio MSI-NET su "*Algorithms and Human Rights*": "La disclosure di interi algoritmi o del codice software sottostante è una soluzione improbabile in questo contesto, in quanto le aziende private considerano il loro algoritmo come un software proprietario protetto. Tuttavia, potrebbe esserci la possibilità di richiedere quei sottoinsiemi di informazioni chiave sugli algoritmi forniti al pubblico, ad esempio quali variabili sono in uso, quali obiettivi hanno, come vengono ottimizzati per gli algoritmi, quali sono i dati di trial, i valori medi e le deviazioni standard, i risultati prodotti o la quantità e il tipo di dati elaborati dall'algoritmo". O anche i suggerimenti che appaiono a pagina 117 della relazione "AI for humanity" redatta da CÉDRIC VILLANI, membro dell'Assemblea nazionale francese: "I revisori possono essere soddisfatti semplicemente dalla verifica dell'equità di un programma testando ad esempio la reazione del sistema immettendo dati di input falsi o creando una grande quantità di profili utente secondo precise linee guida."

L'interessato deve anche essere chiaramente informato di qualsiasi precedente elaborazione di un caso mediante intelligenza artificiale prima o durante un processo giudiziario, con il contestuale diritto di opporsi, in modo che il suo caso possa essere ascoltato direttamente davanti a un tribunale ai sensi dell'articolo 6 della CEDU.

Anche il GDPR è intervenuto sul punto, e di esso occorrerà tenere conto, posto che all'art. 22, par. 1 recita: "L'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona".

2. *L'uso di sistemi AI nei sistemi giudiziari*

L'ondata di trasformazione digitale nelle nostre società comporta ancora una certa disparità di approcci sui sistemi giudiziari degli Stati membri. Molti paesi europei sembrano aver già sviluppato un livello estremamente avanzato all'utilizzo di applicazioni pratiche (in termini sia di tecnologia che di supporto legale), mentre per altri questo è ancora solo un problema emergente e l'attenzione è rivolta esclusivamente ad una efficace gestione degli strumenti IT.

Tra le tecnologie al lavoro in questa grande trasformazione digitale, l'intelligenza artificiale (AI) sembra essere la più spettacolare e la più sorprendente.

Negli Stati Uniti, "avvocati robot" sono già al lavoro e sembrano conversare in linguaggio naturale con gli umani; start-up *legal-tech* offrono nuovi prodotti ad avvocati, servizi legali e assicuratori, permettendo un accesso approfondito ad informazioni giudiziarie e giurisprudenza che mirano a prevedere le decisioni dei giudici.

Appare tuttavia di prima evidenza che questi strumenti privati non siano propriamente definibili come sistemi di giustizia predittiva.

Per ora nella maggior parte degli Stati europei i giudici non sembrano molto attratti dall'uso quotidiano di software predittivo e la causa è da ricercarsi nella risposta tiepida a livello governativo; mentre nel settore privato la ricerca prosegue a favore però di altri operatori del diritto, un po' come negli USA, vale a dire, avvocati, consulenti, commercialisti, notai e assicurazioni, che mediante tali strumenti tentano di ridurre l'incertezza giuridica e l'imprevedibilità della decisione giudiziarie. Tuttavia, i decisori pubblici stanno iniziando ad essere sempre più sollecitati da un settore privato che desidera vedere questi strumenti - talvolta in versioni "beta" - integrati in politiche pubbliche, anche per il ritorno economico che ciò comporterebbe.

Il CEPEJ si propone di fornire ai decisori pubblici e ai professionisti della giustizia le chiavi di lettura per una migliore comprensione del fenomeno della "giustizia predittiva".

Il primo compito è chiarire la natura di questi sistemi di elaborazione dei dati di giurisprudenza di massa, e le loro limitazioni tecniche e teoriche.

Tra l'altro è interessante notare come questi aspetti non siano stati spesso menzionati nel dibattito in ambito giudiziario, mentre sono molto conosciuti e discussi da matematici, statistici e informatici.

Così da un lato si trovano i sostenitori degli strumenti di AI che evidenziano gli aspetti positivi in termini di trasparenza, prevedibilità e standardizzazione della giurisprudenza, mentre coloro che si attestano su posizioni critiche puntano alle limitazioni intrinseche dei sistemi di AI che possono persino trascendere l'atto di giudicare ed influire invece su elementi funzionali essenziali dello Stato di diritto e del sistema giudiziario, primo fra tutti il principio di legalità.

Ciò può avvenire quando i sistemi siano strutturati in maniera tale da non avere una mera funzione consultiva, ma prescrittiva, creando, di fatto una normazione occulta che potrebbe integrare, o addirittura sostituire, il *dictum* normativo, intaccando il principio di interpretazione della legge da parte del giudice, potenzialmente conducendo, a lungo termine, verso una standardizzazione delle decisioni giudiziarie basate non più sul ragionamento caso per caso da parte dei tribunali, ma su un puro calcolo statistico collegato al compenso medio precedentemente concesso da altri tribunali.

È inoltre necessario considerare se queste soluzioni siano compatibili con i diritti individuali sanciti dalla CEDU, come il diritto a un processo equo (in particolare il diritto a giudice naturale prestabilito per legge, il diritto a un tribunale indipendente e imparziale e l'uguaglianza delle armi nei procedimenti giudiziari e il diritto al rispetto della vita privata e familiare.

Tenendo conto di questi problemi, il documento evidenzia il grande potenziale dell'AI per aiutare i professionisti legali nel loro lavoro. Non v'è dubbio che alcune applicazioni di AI che sono ancora in fase di sviluppo o testing, come quelli progettati per migliorare la ricerca legale, potrebbero essere molto utili nel rendere più veloce ed efficiente l'elaborazione del carico di lavoro giudiziario.

Nell'ambito dell'Unione europea che, come visto, ha un approccio alla tematica ancora larvale sono stati individuati a livello giudiziario attualmente i seguenti strumenti:

- ▶ Motori di ricerca giurisprudenziali avanzati;
- ▶ Risoluzione delle controversie online (ODR- *Online Dispute Resolution*);
- ▶ Assistenza nella redazione di atti;
- ▶ Analisi (predittiva, tabelle);
- ▶ Classificazione dei contratti in base a criteri e rilevazione di clausole contrattuali divergenti o incompatibili;
- ▶ “Chatbot” per informare le parti in causa o supportarle nei loro procedimenti legali.

La Lettonia ha dichiarato che sta esplorando le possibilità del machine learning applicato alla migliore allocazione di risorse umane e finanziarie nell'amministrazione della giustizia.

Di seguito un elenco non esaustivo di servizi legali che fanno uso di sistemi di intelligenza artificiale:

Software	State	Type
Doctrine.fr	France	Search engine
Prédicte	France	Analysis (except criminal cases)
Case Law Analytics	France	Analysis (except criminal cases)
JurisData Analytics (LexisNexis)	France	Search engine, Analysis (except criminal cases)
Luminance	United Kingdom	Analysis
Watson/Ross (IBM)	USA	Analysis
HART	United Kingdom	Analysis (criminal, risk of reoffending)
Lex Machina (LexisNexis)	USA	Analysis

3. *Panoramica delle politiche relative agli open data relative alle decisioni giudiziarie nei sistemi giudiziari degli Stati membri del Consiglio d'Europa*

La disponibilità dei dati è una condizione essenziale per lo sviluppo dell'intelligenza artificiale, che consente di eseguire determinati compiti precedentemente svolti dagli esseri umani in modo non automatizzato. Più dati disponibili l'AI ha a disposizione, più è in grado di affinare i modelli e migliorare la loro capacità predittiva.

D'altra parte, l'approccio di open data alle decisioni giudiziarie costituisce un prerequisito per il lavoro delle società *legal-tech* per i motori di ricerca e per la “giustizia predittiva”.

Il trattamento di questi dati solleva una serie di problemi, come i cambiamenti nella formazione della giurisprudenza e la protezione dei dati personali (compresi i nomi dei professionisti).

Si dice che i dati raccolti dal computer siano il “petrolio” del 21° secolo così come il loro uso e il riferimento incrociato tra i dati stanno producendo una ricchezza completamente nuova. Nonostante

alcuni stakeholder e autori contestino questo argomento, i successi globali dell'industria digitale negli ultimi decenni hanno confermato l'enorme potenziale di crescita di questo campo di attività.

La quantificazione delle attività umane, ora su scala globale, potrebbe non mancare di toccare i dati prodotti dal settore pubblico ed è questo che ha indotto verso un'apertura dei dati pubblici, basati su imperativi molto più vecchi, che sono i principi fondanti dei nostri stati costituzionali.

Il grande cambiamento negli ultimi anni è stato causato dall'emergere di dati pubblici scaricabili (open data), in particolare nel contesto della "Partnership for Open Government" (OGP)⁴.

Lo scopo di questa apertura è migliorare la trasparenza delle attività pubbliche, incoraggiare i cittadini nello sviluppo e nella valutazione di politiche pubbliche e garantire l'integrità del servizio pubblico e di coloro che per eseguirlo elaborano notevoli quantità di informazioni, organizzate in database (big data).

4. La definizione di open data su decisioni giudiziarie

Anzitutto la Carta mira a ridefinire e disambiguare lo stesso concetto di "open data", introducendo la distinzione tra l'accesso alle informazioni e l'accesso ai dati (più precisamente, accesso alle informazioni sotto forma di database).

Una certa quantità di informazioni pubbliche, che richiede ampia pubblicità, è già diffusa utilizzando la tecnologia dell'informazione.

In Francia, il sito governativo Légifrance.fr è la principale fonte online di informazioni pubbliche certificate, che comprende non solo testi legislativi e regolamentari, ma anche giurisprudenza e informazioni sui concorsi pubblici. Questa informazione unitaria, sebbene disponibile su Internet, differisce completamente dall'accesso diretto a dati organizzati e inclusi in un database che può essere scaricato ed elaborato da un computer.

Pertanto, gli open data implicano solo la diffusione di dati "grezzi" contenuti all'interno di banche dati informatiche strutturate.

Questi dati, aggregati in tutto o in parte con altre fonti strutturate, costituiscono ciò che viene definito *big data*.

Il Comitato consultivo della Convenzione 108 del Consiglio d'Europa definisce i big data come "La crescente capacità tecnologica di raccogliere ed elaborare rapidamente un grande volume di vari dati ed estrarne di nuovi al fine di condurre alla conoscenza predittiva. In termini di protezione dei dati, i problemi principali non riguardano solo il volume, la velocità e la varietà dei dati elaborati, ma anche l'analisi dei dati utilizzando software per estrarre conoscenze nuove e predittive per il processo decisionale riguardanti individui o gruppi di individui. A tali fini, la definizione di Big Data comprende quindi sia i Big Data che i Big Data Analytics"⁵.

5. Lo sviluppo degli open data in materia giudiziaria negli Stati membri del Consiglio d'Europa e le conseguenze per lo sviluppo della giurisprudenza

Il ciclo di valutazione CEPEJ 2016-2018 si è focalizzato per la prima volta sulla questione delle decisioni giudiziarie fornite in open data, cui venga applicata l'intelligenza artificiale valutando anzitutto la questione dell'anonimizzazione o della pseudonimizzazione dei dati nell'ambito del quadro europeo per la protezione dei dati previsto dal Regolamento generale sulla protezione dei dati (GDPR, Regolamento UE 2016/679) e dalla Convenzione n. 108 del Consiglio d'Europa.

⁴ L'OGP è un'organizzazione non governativa che riunisce quasi 70 Stati membri (inclusi molti degli Stati membri del Consiglio d'Europa) con rappresentanti della società civile e Big digitali.

⁵ T-PD(2017)1, *Guidelines on the protection of individuals with regard to the processing of personal data in a world of big data*.

Di tutti gli Stati e gli osservatori intervistati, nel 2016 solo 5 hanno dichiarato di non aver attuato una politica in materia di open data per le decisioni giudiziarie. Peraltro, questo tasso di risposta dovrebbe essere meglio ponderato visto che alcuni Stati hanno confuso il fatto di rendere accessibili al pubblico le decisioni giudiziarie, con il diverso concetto di open data (Armenia, Belgio, Bosnia ed Erzegovina, Lussemburgo, Federazione Russa, Spagna, Turchia).

Ciò rivela, da un lato, il desiderio di trasparenza da parte delle istituzioni giudiziarie europee e dall'altro, un desiderio da parte di molti paesi di rendere pubbliche le decisioni giudiziarie al fine di sottoporle in seguito ad un trattamento per mezzo di strumenti di Intelligenza Artificiale.

Per quanto riguarda la protezione dei dati personali, 23 paesi hanno dichiarato che stanno pseudonimizzando i dati riguardanti alcuni tipi di cause, come quelle in materia di famiglia o stato delle persone, cancellando i riferimenti che rendono identificabili le parti o i testimoni (nomi, indirizzi, numeri di telefono, numeri di identità, numeri di conto bancario, codici fiscali, stato sanitario, ecc.). Il lavoro di oscuramento dei dati identificativi in questi casi in alcuni paesi è affidato direttamente ai giudici (ad esempio Israele, Repubblica di Moldavia) o ai cancellieri (ad es. Bosnia Erzegovina, Spagna).

Nonostante ciò esiste una reale difficoltà nella misurazione dell'impatto degli open data sull'efficienza e la qualità della giustizia, visto che, come detto più sopra, il riutilizzo di questi dati sembra essere affidato al settore privato che si rivolge a professionisti privati.

Così, se da un lato appaiono numerosi i vantaggi dati dall'ampia diffusione della produzione giurisprudenziale, primo fra tutti una maggiore consapevolezza dell'attività giudiziaria e delle tendenze interpretative, la presumibile maggiore qualità di un sistema giudiziario che sa di essere osservato e la creazione di una base di riferimento fattuale omogenea, dall'altro questo desiderio teorico di "regolare" la giurisprudenza attraverso la leva digitale solleva una serie di domande generali che sono rilevanti per tutti i paesi che tendono ad un simile approccio.

Anzitutto, la Corte europea dei diritti dell'uomo sottolinea chiaramente la necessità di bilanciare la certezza del diritto, il che rende le decisioni più prevedibili, contro la vitalità dell'interpretazione giudiziaria e poi vi sono una serie di osservazioni, non secondarie, anche di natura tecnica e sostanziale:

1. La raccolta di tutte le decisioni giudiziarie pubblicabili attualmente non è ben coordinata tra tutti i gradi di giudizio: alcune delle applicazioni informatiche utilizzate nei tribunali europei non sono state progettate per questo, in particolare per quanto riguarda le decisioni di primo grado, e alcuni paesi devono ancora stabilire nuovi processi per la raccolta delle decisioni, se vogliono rendere la base di dati esaustiva.

2. Nonostante alcuni sviluppi promettenti, non è stato ancora ideato un meccanismo efficace di post-identificazione che possa prevenire qualsiasi rischio di identificazione o nuova identificazione a tutela dei dati personali.

3. L'implementazione di sistemi che raccolgono le decisioni giurisprudenziali ai fini di una maggiore conoscibilità e prevedibilità comporta una profonda trasformazione della logica stessa della produzione giurisprudenziale al fine di rispondere all'interrogativo su quale sia il valore giuridico dello "standard" derivante dal numero di decisioni già prese su una singola fattispecie. Vale a dire se tale standard si affianchi al *corpus* normativo e se quindi si possa ragionare di una nuova fonte para-normativa.

4. Una questione di non poco momento riguarda il dubbio se i giudici, che nelle proprie decisioni si discostino dallo standard, debbano motivare o meno tale discrasia.

Questa domanda non è secondaria ed essa non potrà ricevere una risposta unitaria valida per tutti i sistemi giudiziari europei. Nella maggior parte degli stati europei, quelli di *civil law*, i giudici devono porre alla base della propria decisione la legge. In tali contesti, se venisse richiesto ai giudici di motivare le proprie decisioni, fornendo tutte le ragioni per le quali hanno deviato dalla presunta tendenza della giurisprudenza maggioritaria, ciò equivarrebbe ad uno svuotamento sostanziale della funzione giudiziaria.

A questo punto i giudici non sarebbero più sottoposti soltanto alla legge, nell'assumere le decisioni, poiché ad essa si affiancherebbe o sostituirebbe lo standard della giurisprudenza maggioritaria, derivante da statistiche compilate da uno strumento digitale (che potrebbe anche essere distorto o sviluppato da un operatore privato senza alcun controllo esterno).

Il quesito va quindi ben oltre un aspetto nazionale specifico, relazionandosi in modo più ampio sul ruolo degli strumenti di intelligenza artificiale nei processi decisionali dei giudici.

5. Inoltre, parrebbe questione non del tutto peregrina quella legata al rischio dello snaturamento della funzione nomofilattica delle Corti superiori.

Se infatti i giudici di grado inferiore dovessero tenere in alta considerazione lo standard maggioritario frutto di algoritmizzazione vi sarebbe il rischio concreto di una eccessiva uniformità di decisioni non più ordinate secondo la gerarchia dei tribunali da cui emanano.

Così, se lo standard divenisse un criterio non meramente programmatico, ma dotato di una certa forza precettiva, nell'assunzione delle proprie decisioni i giudici si troverebbero a dover giustificare non più lo scostamento dalla giurisprudenza delle Corti superiori, ma dallo standard.

6. Infine, esiste la possibilità che le decisioni giudiziarie vengano annullate secondo una logica riproduttiva che non tiene conto del mutare del contesto. Infatti, mentre di norma le decisioni giudiziarie vengono assunte sia sulla base dell'evoluzione del quadro normativo (nazionale, europeo o internazionale), sia sulla base della giurisprudenza disponibile come riferimento (per esempio, da tribunali supremi e tribunali europei) sia del contesto socioeconomico, lo standard rischierebbe di indurre all'eccessiva standardizzazione non ponderata delle decisioni giudiziarie.

Potrebbero anche sorgere dubbi sulle conseguenze per gli utenti, che dovrebbero beneficiare direttamente della trasparenza delle attività decisionali. Vi è il concreto dubbio che il *quisque de populo*, anziché trarre beneficio dalla pubblicazione di tutte le decisioni giudiziarie su Internet, sia sopraffatto dall'enorme quantità di informazioni senza necessariamente trarne benefici qualitativi per mancanza di conoscenza legale e di apparato critico per analizzarli.

Paradossalmente, gli open data sono in grado di fornire informazioni significative proprio perché rendono possibile una considerevole mole di dati non elaborati con un semplice clic su un collegamento, ma il significato del file probabilmente rimane totalmente oscuro per la maggioranza dei cittadini.

Il primo requisito affinché i dati aperti siano trasparenti e informativi è che terze parti possano usarli a fini di analisi e sviluppo, ma un modello economico in cui i dati della giurisprudenza pubblica, vengano elaborati dalle stesse autorità giudiziarie al fine di renderli tecnicamente "leggibili" dai sistemi di intelligenza artificiale e conformi ai requisiti legali per la protezione dei dati personali, per essere poi rivenduti dal settore privato a tribunali, professionisti e cittadini, non può andare esente da qualche perplessità.

6. Protezione dei dati personali negli open data delle decisioni giudiziarie

6.1. I nomi delle parti e dei testimoni

Al fine di trovare un giusto equilibrio tra la necessità di rendere pubbliche le decisioni giudiziarie e il rispetto dei diritti fondamentali delle parti o dei testimoni, i loro nomi e indirizzi non dovrebbero essere pubblicati nelle decisioni, onde evitare il rischio di furti di dati o di riutilizzo illecito di tali informazioni personali e sensibili.

I processi automatizzati possono essere utilizzati per nascondere sistematicamente tali informazioni, così come numeri di telefono, indirizzi e-mail, date di nascita, nomi dei bambini, luoghi, ecc.

In termini di principi di protezione dei dati personali, questo occultamento equivale a una semplice pseudonimizzazione dei dati, non alla completa anonimizzazione, poiché la tipologia di informazioni contenute nelle decisioni giudiziarie, combinate con la facilità crescente di riferimenti incrociati con

altri database, in pratica rende impossibile garantire che la persona interessata non possa essere nuovamente identificata.

Nell'assenza di tale garanzia, non potendo questi dati essere qualificati come anonimi, bensì come pseudonimizzati, potendo astrattamente sempre risalire alla versione completa della sentenza contenente i dati nella loro interezza, essi sono soggetti alle norme sulla protezione dei dati personali. Alcuni dati personali, poi, sono particolarmente sensibili e necessitano di particolare attenzione, come previsto dall'articolo 6 della Convenzione 108, poiché la loro diffusione presenterebbe gravi rischi di discriminazione, profilazione e violazione della dignità umana.

Ciò vale per i dati che rivelano l'origine razziale o etnica, le opinioni politiche, l'appartenenza a sindacati, le credenze religiose o di altro tipo, la salute fisica o mentale, la vita sessuale o gli orientamenti sessuali, i dati genetici e biometrici, i dati giudiziari penali e che spesso confluiscono all'interno del testo di decisioni giurisprudenziali.

6.2 I nomi dei professionisti, inclusi i giudici

Ovviamente, per un avvocato riuscire a prevedere l'esito del giudizio è un elemento essenziale ed alcuni ritengono che conoscere il giudice sia tanto importante quanto conoscere la legge.

Empiricamente, grazie all'esperienza, molti avvocati hanno avuto modo di confrontare commissioni e corti così da fornire una migliore consulenza ai clienti; se però questo metodo poteva ritenersi valido in relazione a un numero limitato di tribunali locali, ora, con il progressivo allentamento delle restrizioni in molti paesi e la libertà di muoversi e lavorare all'interno dell'intera UE tale approccio non appare più adeguato dovendo l'avvocato conoscere la giurisprudenza di ogni giurisdizione nazionale ed europea.

È facile prevedere che, in futuro, vi sarà una crescita incontrollata dell'utilità - e quindi della richiesta e dei relativi costi - delle applicazioni di *machine learning*, molto più efficaci della singola esperienza tradizionale del professionista.

L'uso di tali applicazioni però potrebbe ulteriormente accentuare la distorsione della concorrenza e della disuguaglianza tra studi legali che hanno, o non hanno, disponibilità all'utilizzo di tali software di analisi "predittiva".

Esiste un rischio reale che, a causa di tale vantaggio competitivo, il diritto ad un equo processo stabilito dalla legge venga compromesso.

La possibilità di attuare la profilazione mediante riferimenti incrociati di dati pubblici e privati potrebbe consentire alle società private e ai loro avvocati di impegnarsi ancora di più per pratiche ai limiti della correttezza.

Questa tattica, peraltro, è già stata attuata negli Stati Uniti per i reati di stampa, compresa la diffamazione a la violazione della privacy a mezzo stampa, ove i querelanti sono già informati su quale sia il tribunale migliore che concede maggiori risarcimenti.

Sul punto occorre non confondere gli *open data* con la necessità di pubblicare un certo tipo di informazioni che sono, comunque per loro natura, pubbliche.

La fornitura di dati informatici giurisprudenziali è infatti questione diversa dai criteri legali di pubblicazione di copie di decisioni originali o di certificati. L'obiettivo degli *open data* è infatti consentire l'elaborazione automatizzata di giurisprudenza a basso costo e in questo progetto non rientra la pubblicizzazione dei nomi dei professionisti, mentre, è un obbligo comune previsto dagli ordinamenti degli stati europei inserire nelle decisioni il nome dei giudici, in ossequio al principio del processo pubblico di cui all'articolo 6, paragrafo 1, della CEDU, al fine di garantire l'imparzialità dei giudici (che devono essere identificabili) e il rispetto delle norme procedurali sulla pubblicità e collegialità.

La sfida più che altro, risiede nel conciliare esigenze spesso contrastanti, come rendere trasparenti le attività pubbliche consentendo ai cittadini di conoscere e valutare i propri giudici, da un lato, proteggendo al contempo la loro privacy dall'altro.

Esistono sfide rigorose per garantire l'imparzialità dei giudici e persino degli organismi giudiziari nel loro insieme, e in questo senso gli *open data* possono aiutare; ma ci si può d'altro canto interrogare su quali possano essere le misure pratiche per proteggerli da potenziali tentativi di attacco esterno che potrebbero essere indotti dall'incrociare i dati personali dei giudici di riferimento inclusi nelle banche dati con altre fonti, come social network, o risultati di motori di ricerca generalisti, pur salvando l'individuazione di ipotetici pregiudizi politici, religiosi e di altro genere.

Queste domande non si presentano nella stessa forma ovunque in Europa e sono strettamente dipendenti dalle caratteristiche specifiche del sistema giudiziario interessato (e dalla natura dell'organo di gestione della carriera giudiziaria), dall'emissione collegiale o meno della sentenza e dal livello del tribunale interessato.

In Svizzera, per esempio, in cui i giudici sono eletti, la pubblicazione è una garanzia di trasparenza e responsabilità sociale dei giudici nei confronti di cittadini e gruppi politici, e queste informazioni sono già disponibili nei motori di ricerca online (che tecnicamente non sono *open data* in senso stretto).

Né queste domande sorgono nella stessa forma a seconda del livello di giurisdizione. La valutazione delle caratteristiche della giurisprudenza dei giudici delle corti inferiori potrebbe non essere la stessa per i giudici delle corti supreme o dei tribunali internazionali.

Ad esempio, la Corte europea dei diritti dell'uomo autorizza le ricerche delle sentenze con i nomi dei giudici membri del comitato decisionale, ma non consente il calcolo di statistiche relative a un determinato giudice. D'altra parte, ai paesi in cui gli organi giudiziari non hanno familiarità con la pratica delle *dissenting opinions*, può sembrare ingiusto assegnare a un giudice la responsabilità personale per una decisione contro cui ha votato durante la deliberazione in un tribunale collegiale.

Questi dibattiti sono stati ben definiti da una missione di studio condotta nel 2006 in Francia del professor Loïc Cadiet⁶. Le conclusioni della missione rimangono limitate poiché non raccomandano di vietare la pubblicazione ma di riservarla per alcuni tipi di contenzioso, escludendola per altri, ad esempio in speciali materie penali.

Allo stato attuale, potrebbe essere applicato un semplice principio di precauzione in ordine alla valutazione dell'interesse, del tipo di controversia e del grado di giurisdizione, nel pubblicare i nomi dei giudici in un database scaricabile. Allo stesso modo, non si può escludere la possibilità che le stesse istituzioni giudiziarie o terzi autorizzati possano sfruttare queste informazioni al di fuori del contesto degli *open data* per conoscere la giurisprudenza.

In ogni caso la pubblicazione di tali dati dovrebbe essere ancora esaminata alla luce delle norme europee internazionali applicabili in materia di protezione dei dati personali, come quelle della Convenzione n. 108, del GDPR e della direttiva 2016/680, in ambito penale.

7. Caratteristiche operative del machine learning applicato alle decisioni giudiziarie

Sono due le tecniche centrali dell'intelligenza artificiale applicata alle decisioni giudiziarie: l'elaborazione del linguaggio naturale e l'apprendimento automatico.

Nella maggior parte dei casi, l'obiettivo di questi sistemi non è riprodurre il ragionamento legale ma identificare le correlazioni tra i diversi parametri di una decisione (ad esempio, in caso di divorzio rivendicare la durata del matrimonio, il reddito dei coniugi, l'esistenza di adulterio, l'entità del beneficio richiesto, ecc.) e, attraverso l'uso dell'apprendimento automatico, inferire uno o più modelli per "prevedere" una futura decisione giudiziaria.

Lo scopo di questo capitolo è chiarire la natura intrinseca del software descritto come "predittivo" - a volte in versione "beta", - sia in termini di potenziale che di limiti spiegando in maniera semplice

⁶ L.C. CADIET, *Introduction to French Civil Justice System and Civil Procedural Law* <http://www.ritsumei.ac.jp/acd/cg/law/lex/rlr28/CADIET3.pdf>

gli algoritmi di apprendimento automatico, che sono il cuore dell'analisi automatizzata della giurisprudenza.

8. *Le funzionalità teoriche del software di “giustizia predittiva”*

I software “predittivi” si propongono di individuare le probabilità di successo o di fallimento di un caso davanti a un tribunale mediante la modellizzazione statistica di decisioni precedenti, attingendo a due ampi settori informatici: l'elaborazione del linguaggio naturale e l'apprendimento automatico.

Questi approcci di modellizzazione sono spesso indicati come AI, ma in realtà, si tratta di AI “deboli”⁷. Il termine AI è oggetto di dibattito in quanto conduce a molte ambiguità poiché, se è pur vero che esso è entrato nel nostro linguaggio quotidiano per descrivere una vasta gamma di scienze e tecnologie che consentono ai computer di battere i migliori campioni di scacchi, guidare un'auto, conversare con l'uomo, ecc. è altrettanto vero che occorre distinguere le diverse applicazioni mediante le esatte tecnologie sottostanti, incluso l'apprendimento automatico.

Quando ci si riferisce a tutte queste risorse altamente specializzate esse andrebbero correttamente inquadrare in forme di AI “deboli”, ben lontane dal sogno di strumenti di reale autoapprendimento autonomo.

In relazione specificamente alla giustizia, sono progettati sistemi di giustizia predittiva sia ad uso degli uffici legali, delle assicurazioni e degli avvocati affinché anticipino l'esito di un contenzioso, sia per aiutare i giudici nel loro processo decisionale. Tali strumenti forniscono una rappresentazione grafica della probabilità di successo per ogni risultato di una controversia basato su criteri inseriti dall'utente (specifico per ogni tipo di controversia) e sono in grado, sempre su base statistica, di calcolare il probabile importo del risarcimento dei danni deciso dal tribunale.

9. *Il funzionamento pratico dell'intelligenza artificiale: macchine statistiche che costruiscono modelli basati sul passato*

Deve essere fatta una distinzione dall'inizio tra ciò che è una “predizione” e ciò che è una “previsione”. La previsione è l'atto di annunciare ciò che accadrà (*prae*, prima - *dictare*, dire) prima che gli eventi si verifichino (per ispirazione soprannaturale, per chiaroveggenza o premonizione).

Previsione è invece il risultato dell'osservazione di un insieme di dati al fine di prevedere una situazione futura. Questo fraintendimento è probabilmente dovuto al fatto che il termine sia stato mutuato da varietà di tecniche di *data science* derivate da matematica, statistica e teoria dei giochi, che analizzano fatti presenti e passati per fare ipotesi su eventi futuri.

Va inoltre notato come la logica del funzionamento dei software di giustizia predittiva si basi essenzialmente sia su classificazioni generative⁸ (comunemente indicate come bayesiane) sia su classificazioni discriminative che alla fine cercano di stimare l'intervallo di valori attuale o futuro di una variabile (ad es. il risultato di una sperimentazione) dall'analisi di esempi passati.

⁷ L'intelligenza artificiale è definita nel glossario finale della Carta Etica come “*A set of scientific methods, theories and techniques whose aim is to reproduce, by a machine, the cognitive abilities of human beings. Current developments seek to have machines perform complex tasks previously carried out by humans. However, the term artificial intelligence is criticised by experts who distinguish between “strong” AIs (yet able to contextualise specialised and varied problems in a completely autonomous manner) and “weak” or “moderate” AIs (high performance in their field of training). Some experts argue that “strong” AIs would require significant advances in basic research, and not just simple improvements in the performance of existing systems, to be able to model the world as a whole*”. Quindi quando si parla, come in questo contesto, di “Machine learning”, si intende riferirsi ad una intelligenza artificiale di tipo debole.

⁸ Molto semplicisticamente, senza alcuna pretesa di esaustività, con l'approccio generativo si stimano le probabilità a priori e le probabilità condizionali, quindi si combinano a formare le probabilità a posteriori e si classificano con la regola di Bayes.

Al proposito è anche importante sfatare il mito dell'idea antropomorfa che le macchine informatiche siano intelligenti e che i loro progettisti siano riusciti ad implementare una sorta di mente umana nei loro meccanismi. Sfortunatamente, questa idea permea ancora molte analisi della giustizia predittiva e induce a ritenere erroneamente che queste funzionalità immediate o future, possano replicare in qualche modo l'intelligenza umana.

Questo atteggiamento, alimentato ogni giorno dal continuo progresso e immissione sul mercato di nuovi dispositivi tecnologici sempre più performanti induce ad avvicinarsi a queste previsioni con una certa dose, più o meno consapevole, di misticismo, a volte affermando che ciò che non è del tutto possibile oggi, lo sarà inevitabilmente domani.

In realtà ancora non esistono le AI “forti” della letteratura di fantascienza, dotate non solo di intelligenza ma anche di coscienza, infatti, gli attuali sistemi di *machine learning* al momento in fase di sviluppo sono descritti come AI “deboli” e sono in grado di estrarre schemi complessi e di apprendere da grandi volumi di dati in modo efficiente e spesso con alti livelli di precisione predittiva, ma basandosi sempre e comunque sulla statistica fornita dalla mole di dati a cui attingono.

Per superare qualsiasi considerazione istintiva o fantasiosa, occorre comprendere tre concetti fondamentali di queste tecnologie di elaborazione e analisi delle informazioni:

1) L'intelligenza artificiale non è un singolo oggetto omogeneo: è in realtà un assemblaggio di scienze e tecniche (matematica, statistica e informatica) in grado di elaborare dati per progettare sistemi informatici molto complessi.

2) I motori di intelligenza artificiale non producono intelligenza di per sé ma procedono utilizzando un approccio induttivo: associano cioè in modo quasi automatizzato una serie di possibili risultati (output) utilizzando un insieme di osservazioni (input) preconfigurate. In particolare, per la giustizia predittiva, il motore crea collegamenti tra i diversi gruppi lessicali che compongono decisioni giudiziarie. Questi gruppi sono correlati tra quelli identificati nella fase di input (fatti e ragionamento) e quelli nella fase di output (il dispositivo della decisione).

3) L'affidabilità del modello (o della funzione) costruito dipende fortemente sia dalla qualità dei dati utilizzati sia dalla scelta della tecnica di apprendimento automatico.

Per meglio comprendere può valere la pena ripercorrere la storia dell'AI tornando brevemente agli antenati di questi sistemi, vale a dire ai sistemi esperti che si basavano su regole di elaborazione scritte da scienziati informatici. I sistemi esperti (ES) si sono sviluppati rapidamente alla fine degli anni '80 e Anni '90, soprattutto in medicina e finanza. Questi sistemi erano in grado di rispondere a domande specializzate e a ragionare utilizzando fatti noti, eseguendo le regole predefinite in un motore. Nonostante il successo degli ES, come Deep Blu contro Garry Kasparov in una serie di partite di scacchi nel 1997, questi sistemi fallirono, in particolare perché non erano in grado di interpretare “l'infinita varietà di situazioni e contesti”⁹ e diventavano inefficaci superando le 200 o 300 regole di codifica, sia in termini di prestazioni di esecuzione, sia di manutenzione ed erano programmati con regole che riproducevano la logica del ragionamento giuridico, un progetto quindi più ambizioso dell'attuale.

Oggi, l'idea non è più quella di scrivere regole di ragionamento che rispecchino il ragionamento dell'uomo, come per i sistemi esperti più vecchi, ma di lasciare che i sistemi stessi di *machine learning* identifichino modelli statistici esistenti nei dati e li correlino a risultati specifici.

Ad esempio, i traduttori online di oggi non compiono ragionamenti astratti, ma inferiscono una probabile stima della migliore corrispondenza tra gruppi di strutture lessicali e traduzioni già fatte, anche basandosi sulle azioni degli utenti che contribuiscono così al miglioramento della ricerca delle corrispondenze, ma l'algoritmo di machine learning in realtà non esegue una traduzione delle frasi elaborate comprendendone il significato¹⁰.

⁹ H. DREYFUS, *What Computers Still Can't Do. A Critique of Artificial Reason*, Cambridge, The MIT Press, 1992 citato in D. CARDON in *A quoi rêvent les algorithmes, nos vies à l'heure des Big Data*, 59.

¹⁰ L. GONG, *La traduction automatique statistique, comment ça marche?*, Interstices.info, published on 29 October 2013, <https://interstices.info/la-traduction-automatique-statistique-comment-ca-marche/>.

Questo esempio mostra come sia diverso l'approccio di avvocati e ricercatori: un avvocato cercherà di comprendere l'efficacia delle regole preesistenti, che il ricercatore può solo prevedere attraverso il linguaggio della sua scienza, il che talvolta significa limitarsi ad interpretare migliaia di parametri da un sistema di *machine learning*.

L'apprendimento automatico è un campo dell'informatica in cui i programmi dei computer apprendono dall'esperienza.

Gli algoritmi fanno sì che una macchina esegua un processo di addestramento, come un bambino che impara nel suo ambiente. In sintesi, queste tecniche di apprendimento possono o meno essere supervisionate da un essere umano.

La categoria più importante dell'apprendimento automatico è attualmente quella dell'apprendimento per rinforzo: la sola macchina raccoglie "ricompense" virtuali se l'azione produce il risultato atteso in un determinato ambiente.

I metodi di *machine learning* includono le reti neurali (o la loro versione più complessa nota come *deep learning*), che sono state ampiamente pubblicizzate in vista della loro autonomia e delle loro applicazioni piuttosto sorprendenti, come quella che riesce a ottenere punteggi elevati sui vecchi videogiochi Atari 2600¹¹ (basati esclusivamente sul posizionamento). Altre applicazioni pratiche per queste tecnologie stanno già influenzando la nostra vita quotidiana e stanno iniziando ad apparire nel mondo professionale della giustizia¹².

Tuttavia, questi algoritmi rimangono altamente specializzati in un compito specifico e presentano problemi di discernimento di fronte a situazioni caotiche o con dati insufficienti per consentire la previsione (come la comprensione effettiva del linguaggio naturale).

Nelle scienze sociali, a cui appartengono la legge e la giustizia, il fallimento sembrerebbe persino inevitabile in assenza di un modello cognitivo convincente. Per Pierre Lévy, l'intelligenza artificiale è in realtà una cassetta degli attrezzi eterogenea (regole logiche, sintassi formali, metodi statistici, simulazioni neurali o socio-biologiche, ecc.) che non offre una soluzione generale al problema della modellistica matematica della cognizione umana¹³.

Pertanto, il vero apprendimento predittivo dovrebbe in realtà basarsi su una buona rappresentazione sistemica del mondo, che il ricercatore di intelligenza artificiale Yann LeCun ritiene sia una questione scientifica e matematica fondamentale, non una questione tecnologica¹⁴.

Il problema è che gli attuali sistemi di elaborazione dei *big data* non tentano di riprodurre il nostro modello di cognizione ma producono statistiche di contesto su una dimensione di dati senza precedenti, senza alcuna reale garanzia di escludere false correlazioni¹⁵.

Mentre è comunemente accettato che i modelli statistici e probabilistici siano migliorabili aumentando i dati che li alimentano, alcuni matematici avvertono del rischio che aumentando i dati vi sia un correlativo aumento di false correlazioni (cioè di collegamenti tra fattori che non hanno assolutamente alcun nesso causale).

I matematici Cristian Sorin Calude e Giuseppe Longo indicano il rischio di un diluvio di false correlazioni nei *big data*: maggiore è un database utilizzato per le correlazioni, maggiori sono le possibilità di trovare schemi ricorrenti e maggiori sono le possibilità di errori¹⁶.

¹¹ <https://www.nature.com/articles/nature14236>.

¹² L'Intelligenza Artificiale Watson di IBM fornisce un servizio in campo medico con uno strumento di ricerca chiamato "Ross", presentato come un avvocato virtuale, "*Do more than humanly possible*" <https://www.obtanywhere.com/blog/2017/03/01/worlds-first-ai-lawyers-legal-apocalypse>.

¹³ P. LEVY, *Intelligence artificielle et sciences humaines*, Pierre Levy's blog, 8 October 2014. [Online], <https://pierrelevyblog.com/2014/10/08/intelligence-artificielle-et-sciences-humaines/>.

¹⁴ Y. LECUN, *Qu'est-ce que l'intelligence artificielle*, Collège de France website, https://www.college-de-france.fr/media/yann-lecun/UPL4485925235409209505_Intelligence_Artificielle_Y._LeCun.pdf.

¹⁵ D. CARDON, *A quoi rêvent les algorithmes, nos vies à l'heure des big data*, La République des idées, Editions du Seuil, 2015, 60.

¹⁶ C. SORIN CALUDE, G. LONGO, *Le déluge des corrélations fallacieuses dans le big data*, in Bernard Stiegler (dir.) (2017), *La toile que nous voulons*, FYP editions, 156.

Ciò che può apparire come regolarità per un'intelligenza artificiale (ricorrenti collegamenti tra dati, concetti, contesti o gruppi lessicali diversi) potrebbe effettivamente essere frutto di casualità, così come è stato notato in alcune vaste serie di numeri, punti o oggetti, possono apparire schemi casuali regolari impossibili da distinguere da schemi algoritmici non casuali¹⁷.

Infine, gli algoritmi, indipendentemente dal loro attuale livello di sofisticazione, si riducono automaticamente alle interazioni stabilite nella fase di apprendimento e quindi al loro passato. Il contenuto e la qualità dei flussi di dati utilizzati nella composizione dei calcoli sono quindi fondamentali per comprendere i risultati ottenuti e per identificare possibili distorsioni analitiche. Ci sono molte sfide anche in questo caso poiché, in un approccio esaustivo, l'analisi della maggior quantità possibile di dati relativi ad un'attività produrrà risultati il cui significato deve essere chiarito rispetto a tutti i fattori che hanno avuto un'influenza.

In un approccio più restrittivo in cui vengono campionati i dati in entrata, sorgeranno anche rischi legati al compromesso necessario per selezionare un dato piuttosto che un altro.

10. *L'intelligenza artificiale può modellare in anticipo il ragionamento legale?*

L'intelligenza artificiale supera le difficoltà incontrate con i sistemi esperti più vecchi poiché non tenta di riprodurre il ragionamento legale di un essere umano esperto nel campo, ma conduce a categorizzazioni tra i diversi parametri identificati dai progettisti o tra quelli scoperti dalla macchina.

I sistemi giuridici continentali sono molto lontani dall'ideale di razionalità che informava, ad esempio, le *code civil* del 1804 in Francia. Esistono una moltitudine di fonti che non si inquadrano perfettamente e che si riferiscono ad un insieme di regole il cui significato rimane indeterminato, che il teorico giuridico Herbert LA Hart ha definito la "struttura aperta della legge"¹⁸.

Sistemi di common law, sebbene considerati più efficienti, poiché si adattano dinamicamente alle nuove esigenze legali, stanno anche loro evolvendo e non offrono più un elevato grado di certezza giuridica.

Quindi si può affermare che le regole giuridiche non si evolvono in modo lineare, a differenza delle leggi empiriche, quelle delle "scienze esatte", in cui ogni nuova regola integra generalmente le precedenti e non invalida un insieme completo di ragionamento.

Ciò ha determinato una generale complessità dei sistemi legali, che i vecchi sistemi esperti non hanno saputo affrontare, anzitutto perché, i vecchi ES non reggevano più di 200/300 regole logiche, poi perché la divisione della legge in regole di produzione non era sufficiente a fornire una valida rappresentazione del *corpus* di conoscenze e metodi che guidano un avvocato.

Di conseguenza, risultava altrettanto impossibile modellare il ragionamento dei giudici su un computer avente come modello un sistema giuridico positivo.

Come ha evidenziato la teoria del diritto, il ragionamento giudiziario è soprattutto una questione di valutazione e interpretazione dei fatti comprovati e pertinenti di un caso, delle norme giuridiche applicabili (testuali o precedenti) - il cui significato rimane indeterminato -,¹⁹ e dell'interpretazione soggettiva dei giudici del concetto di equità.

Xavier Linant de Bellefonds ha sottolineato che la complessità della legge risiede nella sua natura teleologica e controversa: due argomentazioni coerenti possono condurre a giudizi diversi secondo due priorità differenti²⁰.

Questo perché il famoso sillogismo giuridico è più un modo di presentare ragionamenti legali che la sua traduzione formale. Non riflette l'intero ragionamento del giudice, che in realtà è costituito da

¹⁷ R. L. GRAHAM, J. H. SPENCER – *Ramsey Theory*, Scientific American, vol.263, No.1, July 1990, 112-117 citato da C. SORIN CALUDE, G. LONGO, *Le déluge des corrélations fallacieuses dans le big data*, *ibid*.

¹⁸ H. L. A. HART (1976), *Le concept de droit*, Saint-Louis university departments, Brussels.

¹⁹ Sul punto cfr., M. TROPER (2001), *La théorie du droit, le droit, l'Etat*, PUF, Paris, spec. 69-84.

²⁰ X. LINANT DE BELLEFONDS (1994), *L'utilisation des systèmes experts en droit comparé*, *Revue internationale de droit comparé*, Vol. 46, No. 2, 703-718.

una moltitudine di fattori decisionali, che non possono essere formalizzati *a priori*, e talvolta si basano sulla discrezionalità con la quale il giudice stesso risponde ai quesiti su quali siano i fatti rilevanti, quali i fatti provati, quale disposizione applicare, quale sia il significato frutto di interpretazione della disposizione alla fattispecie concreta, quale norma debba prevalere tra una gamma di norme in conflitto.

La coerenza complessiva delle decisioni giudiziarie non viene mai raggiunta ed è più una questione di un resoconto *a posteriori* che i giudici usano nel loro ragionamento, avendo più una funzione di convincersi della validità di una soluzione specifica che di descrivere rigorosamente e obiettivamente tutte le fasi che hanno portato alla decisione presa.

Tuttavia, questo lavoro di interpretazione è esattamente ciò che le tecniche di apprendimento automatico non fanno - e non tentano di fare - oggi, poiché svolgono, come abbiamo visto, elaborazioni automatizzate in cui si suppone che la correlazione di grandi quantità di informazioni sostituisca la comprensione delle reali cause di una decisione.

Lungi dal tentare di formalizzare il ragionamento giuridico, fanno solo sperare che i modelli catturati possano anticipare le probabili decisioni di un giudice in situazioni simili.

I risultati raggiunti dall'AI non sono in realtà correlati alla questione della conformità giuridica di una particolare soluzione e non possono discriminare tra argomenti legali e illegali.

Una revisione del lavoro svolto dall'University College di Londra (UCL) sulla giurisprudenza della Corte europea dei diritti dell'uomo conferma questo assunto.

Lo studio dell'UCL ha ipotizzato che un semplice modello di *machine learning* potesse prevedere l'esito di un caso con un'accuratezza del 79%. Il modello di *machine learning* si è rivelato più accurato rispetto alla parte descrittiva dei fatti delle decisioni studiate rispetto al ragionamento relativo all'applicazione della Convenzione al caso in questione²¹.

L'esame dei termini correlati all'accertamento di una violazione (con una ponderazione positiva) e quelli correlati a una non violazione della Convenzione europea dei diritti dell'uomo (con una ponderazione negativa) fa luce inequivocabile sulla meccanica sul lavoro, che non è in alcun modo paragonabile al ragionamento legale.

La scelta dei fatti rilevanti e la loro interpretazione costituiscono uno degli elementi della decisione del giudice. In altre parole, lo studio UCL ha evidenziato che il sistema è stato solo in grado di indicare una probabilità sulla base delle sole occorrenze di materiale lessicale ampiamente derivato dal ragionamento e dalla motivazione del giudice e non da quello assemblato dal richiedente.

Questo sistema di intelligenza artificiale stabiliva quindi un'alta probabilità di corrispondenza tra gruppi di parole e una decisione che era già stata formalizzata e poteva solo dare origine a un numero limitato di possibili esiti.

In nessun caso si può parlare di riproduzione del ragionamento dei giudici europei né, soprattutto, prevedere un risultato, sulla base, ad esempio, del dato grezzo di un futuro richiedente dinanzi al tribunale di Strasburgo, la cui domanda sarà soggetta a un esame di ammissibilità molto rigoroso (quasi 70.356 domande sono state dichiarate inammissibili o cancellate dall'elenco nel 2017) in gran parte basate sull'applicazione di standard di valutazione (importanza e gravità della denuncia, ecc.) lasciando un notevole margine di manovra all'interno del processo decisionale.

11. L'AI può spiegare il comportamento dei giudici a posteriori?

Una spiegazione *a posteriori* del comportamento di un giudice, in particolare la rivelazione di pregiudizi, richiederebbe che tutti i fattori potenzialmente causativi siano identificati attraverso un quadro interpretativo e un'analisi contestualizzata. Il fatto che, statisticamente, l'assistenza

²¹ Work on a sample of 584 decisions of the European Court of Human Rights: N. ALETRAS, D. TSARAPATSANIS, D. PREOȚIUC-PIETRO, V. LAMPOS, *Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective*, 24.10.2016, <https://peerj.com/articles/cs-93/>.

all'infanzia sia più spesso affidata alle madri che ai padri non dimostra una propensione da parte dei giudici, ma rivela la necessità di mobilitare discipline diverse dalle scienze sociali per far luce su questo fenomeno.

Da un punto di vista scientifico, spiegare un fenomeno o, per quanto ci riguarda, un pezzo di comportamento umano, equivale a determinare i meccanismi causali che hanno portato a questo comportamento usando una certa quantità di dati contestuali.

Ciò richiede, in modo molto schematico, la costituzione preliminare di un quadro interpretativo, esso stesso derivato dalla ripetuta osservazione di questo tipo di evento o comportamento in presenza di determinati fattori o elementi.

Il quadro interpretativo è costituito da ipotesi o da punti di vista adottati dalle diverse discipline delle scienze sociali. Questo è un ulteriore passo analitico che può essere inserito negli algoritmi, ma che essi non possono eseguire da soli.

Alcune società di *legal-tech* sono andate oltre e hanno pensato di poter identificare possibili pregiudizi personali dei giudici e alimentare sospetti di parzialità. I dati in chiaro dei nomi di alcuni presidenti di tribunali e corti superiori amministrativi in Francia hanno permesso di sviluppare un indicatore del tasso di rigetto dei ricorsi contro l'obbligo di lasciare il territorio francese assunto dalle autorità amministrative.

Alcuni commentatori hanno sostenuto con fervore che la presunta imparzialità dei giudici è stata quindi messa in dubbio dall'intelligenza artificiale, ma perché si possa fondatamente affermare l'esistenza di un pregiudizio personale nei processi decisionali dei giudici (diversi dalle loro dichiarazioni personali e pubbliche nel caso in questione) il loro comportamento, o in questo caso la loro decisione, devono essere determinati da tratti di personalità, opinioni o religione. Tuttavia, come è stato detto, una simile spiegazione causale non può essere semplicemente dedotta dal risultato probabilistico fornito dagli algoritmi; al contrario, richiede un ulteriore lavoro analitico al fine di isolare, tra i molti fattori correlati (compresa l'identità dei membri del gruppo di giudici), quelli che sono realmente causativi.

Ad esempio, il fatto che un tribunale di famiglia decida statisticamente più spesso che i bambini dovrebbero vivere con la madre non riflette necessariamente la propensione del giudice a favore delle donne, ma piuttosto l'esistenza di fattori psicosociali, economici e persino culturali specifici della giurisdizione, come l'orario di lavoro di ciascuno dei genitori, il loro reddito, la disponibilità locale di assistenza all'infanzia collettiva, indipendentemente dal fatto che il bambino sia in età scolare, che uno dei genitori abbia o meno una nuova relazione o anche semplicemente che vi sia la mancanza di interesse da parte di genitore a prendersi cura di un bambino piccolo.

Allo stesso modo, le decisioni di espulsione da un paese emesse da un tribunale amministrativo situato vicino a un grande centro di detenzione non possono essere paragonate equamente a quelle di un tribunale che si occupa di tali controversie solo occasionalmente.

Inoltre, indipendentemente dall'ubicazione del tribunale, la questione della giurisprudenza di un giudice unico che si occupa solo occasionalmente di un certo tipo di controversia, ma che utilizza (o ignora) la giurisprudenza dei suoi colleghi, è particolarmente interessante e può legittimamente sollevare la questione dell'uguaglianza dei cittadini nei procedimenti giudiziari. Tuttavia, l'attenzione dovrà essere puntata sulla revisione o sulla conservazione della natura collegiale del sistema giudiziario, piuttosto che sulla classificazione o stigmatizzazione tramite strumenti di *machine learning*.

Giunti a tal punto occorre chiedersi cosa si possa dedurre dalla personalità del presidente di un tribunale collegiale, quando il suo nome sia l'unica informazione nominativa visibile nelle decisioni giudiziarie amministrative pubblicate, o ancora come si possano spiegare due distinti approcci filosofici e culturali delle decisioni giudiziarie, per cui, in alcuni paesi europei esiste una cultura del precedente e una conoscenza dettagliata da parte dei giudici dei dati fattuali di tutto il primo e il secondo grado, mentre altri paesi o sistemi favoriscono l'indipendenza intellettuale di ciascun tribunale, insieme al desiderio di affrontare ogni situazione caso per caso.

Fornire una spiegazione accurata di una decisione del tribunale richiede quindi un'analisi molto più dettagliata dei dati contingenti di ciascun caso e delle norme di legge applicabili, piuttosto che sostenere la vana speranza che la massa di collegamenti abbia senso.

12. Come si applica l'AI nella giustizia civile, commerciale e amministrativa

Lo stato di sviluppo delle tecniche di *machine learning* non consente oggi di raggiungere risultati affidabili per quanto riguarda la “previsione” delle decisioni giudiziarie. D'altra parte, la loro applicazione nella giustizia civile, commerciale e amministrativa deve essere presa in considerazione per la creazione di bilanciamenti o per la risoluzione precontenziosa delle controversie online, quando resta possibile un ricorso successivo al giudice.

Come possiamo vedere, la prima domanda sollevata da tale uso dell'intelligenza artificiale non è tanto se sia benefica o dannosa, desiderabile o meno, ma se gli algoritmi proposti possano raggiungere il tipo di risultato cercato. Le conclusioni degli esperimenti condotti nelle corti di appello di Douai e Rennes in Francia dimostrano chiaramente che a fronte di un discorso stimolante che promuove un prodotto di AI, è possibile nascondere difetti di progettazione inaccettabili e risultati di analisi totalmente errati.

Indipendentemente dalla qualità del software testato, anticipare le decisioni dei giudici in materia civile, commerciale e amministrativa sembrerebbe essere un beneficio potenzialmente desiderabile, sia per i responsabili delle politiche pubbliche giudiziarie sia per professionisti privati.

Qualunque sia la tradizione giuridica del paese, l'incertezza giuridica, vale a dire il rischio di far convalidare o respingere il proprio reclamo legale, spinge il desiderio di poter quantificare questi fattori utilizzando queste nuove applicazioni tecnologiche.

Gli avvocati (o l'ufficio legale di una società) vedono la possibilità di utilizzare questa tecnologia per fornire ai propri clienti una consulenza meglio informata valutando empiricamente e sistematicamente le possibilità di successo di una procedura, e incoraggiando la conclusione di transazioni che, se necessario, consentono di evitare una prova lunga e costosa. Alcuni assicuratori offrono già ai loro clienti l'uso di sistemi predittivi per valutare i meriti della loro attività²².

Allo stesso tempo, i decisori pubblici vedono questo approccio come un'opportunità per regolare meglio il flusso di nuovi procedimenti attraverso i tribunali e fornire loro una leva per ridurre i costi operativi giudiziari anche perché si ritiene che ciò incoraggi le parti in causa ad utilizzare metodi alternativi di risoluzione delle controversie (conciliazione, mediazione o arbitrato).

L'approccio che già esiste in molti sistemi giudiziari finalizzato all'armonizzazione di decisioni in molte materie (divorzio, licenziamento, risarcimento per lesioni personali) potrebbe essere migliorato attraverso un approccio probabilistico o attuariale.

Sono stati persino creati servizi alternativi di risoluzione delle controversie online per aiutare a valutare l'importo del risarcimento per le controversie di piccole dimensioni.

Tuttavia, questi approcci interessanti non sono imparziali e non devono privare i cittadini dell'accesso a un giudice o mettere in discussione il principio del contraddittorio.

In Francia, su iniziativa del Ministero della Giustizia, nella primavera del 2017 le due corti di appello di Rennes e Douai hanno concordato di testare un software di giustizia predittiva su vari ricorsi in contenzioso, che in realtà era un'analisi delle decisioni civili, sociali e commerciali di tutti le corti d'appello francesi.

Sebbene questi dati interni ed esaustivi sulla giurisprudenza fossero già disponibili gratuitamente per molti anni (banca dati JURICA), il Ministero li ha resi disponibili ad una società *legal-tech* quando questa si è offerta di valutare il valore di un'analisi quantificata delle somme risarcite dai due tribunali, oltre a una classificazione geografica delle discrepanze rilevate per domande e processi simili.

²² Il software predittivo viene offerto agli assicurati Allianz, ad esempio.

L'obiettivo dichiarato del software era quindi quello di creare uno strumento decisionale al fine di ridurre, se necessario, l'eccessiva variabilità delle decisioni giudiziarie, in nome del principio di uguaglianza dei cittadini di fronte alla legge. Il risultato dell'esperimento, dibattuto contraddittoriamente tra le due corti d'appello, il Ministero della Giustizia e la società *legal-tech* che ha progettato il prodotto purtroppo ha evidenziato l'assenza di valore aggiunto della versione testata del software per il lavoro di riflessione, decisione e creazione dei magistrati.

Ancora più significativo, sono stati rivelati errori di ragionamento del software che hanno portato a risultati aberranti o inappropriati a causa della confusione tra semplici elementi lessicali tipici del ragionamento giudiziario e le causalità che erano state invece decisive nel ragionamento dei giudici.

13. Un nuovo strumento di calcolo per calcolare i parametri di risarcimento

Le procedure per il calcolo delle tabelle di risarcimento in varie materie civili (ad esempio, compensazione per lesioni personali, indennità compensativa e trattamento di fine rapporto) sembrano migliorare notevolmente se combinate con altre tecniche di elaborazione, al fine di evitare il cd. "effetto performativo" o di autorealizzazione ove il rischio è che un sistema produca lo stesso output progressivamente influenzando gli elementi di input; questo effetto è spesso menzionato per quanto riguarda le tabelle giudiziarie che, quando informate da decisioni basate su queste tabelle, tendono ad essere rappresentative solo di sé stesse.

È importante evidenziare ciò che Jean-Paul Jean, presidente del gruppo di lavoro di valutazione CEPEJ, ha descritto come la sfida qualitativa durante una conferenza sugli open data tenutasi nel 2016 in Francia: la procedura condotta sul *machine learning* dovrebbe utilizzare certificati originali, la cui integrità sia stata verificata²³.

L'altro rischio già affrontato dai progettisti di questi strumenti è quello di "*snooping* dei dati", vale a dire la selezione di soli dati a monte che sono significativi per le griglie di analisi predeterminate, ad esempio escludendo dai campioni decisioni che si prestano male a correlazioni delle sequenze linguistiche attraverso il *machine learning* o qualsiasi altro metodo (ad esempio, decisioni senza presentazione del contenzioso o con un ragionamento limitato).

Ma se calcoliamo una tabella, stiamo già facendo una piccola previsione. Il limite può sembrare relativamente impreciso se non si distingue chiaramente lo scopo del processo: l'obiettivo qui non è quello di fornire informazioni prescrittive ma di fornire informazioni su uno stato di cose.

Fatte salve queste precauzioni metodologiche e operative, sono disponibili tabelle per la gestione di determinate controversie che costituiscono un potente strumento per armonizzare la giurisprudenza. Le tabelle erano precedentemente calcolate su campioni di decisioni più o meno limitati, ma gli strumenti di intelligenza artificiale consentono di rilevare più decisioni e sono suscettibili di produrre tabelle più accurate, applicando uno standard e quindi dando più peso ai risultati.

14. La risoluzione delle controversie online (ODR)

Tutti i tribunali europei affrontano in gradi diversi contenziosi civili ripetitivi di basso valore; l'idea quindi di automatizzare la procedura attraverso la tecnologia dell'informazione o esternalizzarli è ampiamente condivisa²⁴.

²³ J-P. JEAN, *Penser les finalités de la nécessaire ouverture des bases de données de jurisprudence*, Conferenza del 14 ottobre 2016 tenuta alla Corte di cassazione francese, <https://www.courdecassation.fr/IMG///Open%20data,%20par%20Jean-Paul%20Jean.pdf>.

²⁴ Per i crediti transfrontalieri, l'Unione europea ha istituito, con il regolamento n. 524/2013, un quadro comune disponibile su Internet (piccoli crediti europei).

Gran Bretagna, Paesi Bassi e Lettonia sono esempi di paesi che hanno già implementato o stanno per implementare questo tipo di soluzioni più o meno automatizzate²⁵.

Tuttavia, la portata di questi servizi di risoluzione delle controversie online (ODR-*Online Dispute Resolution*) sembra essersi gradualmente estesa. Si è infatti passati da servizi online limitati a servizi ODR prima del contenzioso giudiziale, a sistemi che operano sempre più nel processo giudiziario stesso, al punto da offrire servizi giudiziari elettronici²⁶.

Tali sistemi ormai non riguardano solo controversie sul valore del risarcimento, ma anche controversie fiscali o relative ai servizi di sicurezza sociale o procedimenti di divorzio.

Per coloro che sostengono tali soluzioni, che interessano una serie di professioni legali e il settore privato, l'accesso alla giustizia potrebbe essere notevolmente migliorato da un'ampia soluzione che combini ODR e AI. L'idea è quella di fornire agli istanti una diagnosi automatica della controversia mediante survey, le cui risposte vengono quindi elaborate dalla macchina che emette proposte per una soluzione. Il lavoro del laboratorio *Cyberjustice de Montréal*, che riunisce le varie fasi precontenziose e contenziose in un processo basato su computer per controversie a bassa intensità (ad esempio, tribunali per le controversie di modesta entità in Quebec), è un buon esempio di ibridazione²⁷.

Secondo per i progettisti, vi sarebbero chiari vantaggi in termini di efficienza e qualità ma l'attenzione sul punto va rivolta al parametro e al metodo sul quale sarebbe calcolata una compensazione proposta da tale sistema. E ancora occorrerebbe predefinire il grado di equità dell'algoritmo, prevedere se la proposta possa essere discussa tra le parti, con l'aiuto di un terzo qualificato e se sia sempre possibile il ricorso a un processo effettivo.

Alcuni autori vedono persino l'uso diffuso di questi metodi di risoluzione delle controversie come una nuova manifestazione di "soluzionismo" digitale, vale a dire l'uso sistematico di tecnologie per cercare di risolvere problemi che non necessariamente sono alla loro portata²⁸.

Va anche notato che in Europa, recentemente è stato istituito un quadro normativo più protettivo e vincolante per gli Stati membri: l'articolo 22 del GDPR prevede esplicitamente che le persone possano rifiutare di essere oggetto di una decisione basata esclusivamente sul trattamento automatizzato, con alcune eccezioni.

L'articolo 22 del Regolamento EU/2016/679 prevede: "(Processo decisionale automatizzato relativo alle persone fisiche, compresa la profilazione) 1. L'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona". Sono previste eccezioni (come il consenso), ma "il titolare del trattamento attua misure appropriate per tutelare i diritti, le libertà e i legittimi interessi dell'interessato, almeno il diritto di ottenere l'intervento umano da parte del titolare del trattamento, di esprimere la propria opinione e di contestare la decisione".

Nello stesso senso si pone la Convenzione sulla protezione delle persone rispetto al trattamento automatizzato di dati a carattere personale firmata a Strasburgo nel 1981, n. 108 del Consiglio d'Europa, come modificata dal Protocollo adottato nel maggio 2018 ad Elsinore, che all'art. 9.1

²⁵ Un'analisi della risoluzione delle dispute on line in Gran Bretagna è rinvenibile all'url <https://www.judiciary.gov.uk/wp-content/uploads/2015/02/Online-Dispute-Resolution-Final-Web-Version1.pdf>. Cfr. anche il sistema "Pas" nei Paesi Bassi, che emette decisioni automatizzate basate su autorizzazioni precedentemente concesse e che ha dato origine a controversie a livello nazionale e dinanzi alla Corte di Giustizia europea: vi sono due casi (c-293/17 e c-294/17) portati al Consiglio di Stato dei Paesi Bassi (Agricoltori e conservazione della natura vs. Paesi Bassi) per determinare se un sistema (Programma di regolazione dell'azoto) è autorizzato o meno a decidere se gli agricoltori stanno violando la direttiva Habitat.

²⁶ D. THOMPSON, *Creating new pathways to justice using simple artificial intelligence and online dispute resolution*, Osgoode Hall Law School of York University.

²⁷ <http://www.cyberjustice.ca/projets/odr-plateforme-daide-au-reglement-en-ligne-de-litiges/>.

²⁸ E. MOROZOV, *Pour tout résoudre cliquez ici*, FYP editions, citato da D. LARROUSSERIE, "Contre le 'solutionnisme' numérique", *Le Monde*, 6 October 2014, https://www.lemonde.fr/sciences/article/2014/10/06/contre-le-solutionnisme-numerique_4501225_1650684.html.

prevede che “Ogni individuo ha diritto: (a) a non essere soggetto a una decisione che lo riguardi in modo significativo sulla base unicamente di un trattamento automatizzato di dati senza che le sue opinioni siano prese in considerazione”.

Nonostante questo principio di divieto, l’articolo 9, paragrafo 2 afferma che “il paragrafo 1 (a) non si applica se la decisione è autorizzata da una legge a cui il titolare del trattamento è soggetto e prevede anche adeguate misure di salvaguardia dei diritti, delle libertà e degli interessi legittimi dell’interessato”.

Così come all’art. 9.1 (c) è previsto che l’interessato possa ottenere, su richiesta, la logica del ragionamento alla base del trattamento dei dati quando i risultati di tale trattamento gli siano applicati.

Nell’*Explanatory Report* al Protocollo di Elsinore del 2018, si legge al paragrafo 75, in relazione all’art. 9.1.a. che “È essenziale che un individuo soggetto ad una decisione puramente automatizzata abbia il diritto di contestare tale decisione presentando, in modo significativo, il suo punto di vista e le sue argomentazioni. In particolare, l’interessato dovrebbe avere l’opportunità di provare l’eventuale inesattezza dei dati personali prima che vengano utilizzati, l’irrilevanza del profilo da applicare alla sua situazione particolare o altri fattori che avranno un impatto sul risultato della decisione automatizzata. Questo in particolare nel caso in cui gli individui siano danneggiati dall’applicazione di ragionamenti algoritmici che comportano la limitazione di un diritto o il rifiuto di un beneficio sociale o in cui vedono la loro capacità di credito valutata solo da un software. Tuttavia, un individuo non può esercitare questo diritto se la decisione automatizzata è autorizzata da una legge alla quale è soggetto il Titolare del trattamento e che prevede anche misure idonee a salvaguardare i diritti, le libertà e gli interessi legittimi dell’interessato”.

Sempre nell’*Explanatory Report* di Elsinore, in relazione all’art. 9.1.c) si rinviene il principio in base al quale “gli interessati hanno il diritto di conoscere la logica del ragionamento alla base del trattamento dei dati, comprese le conseguenze di tale ragionamento e le conclusioni che potrebbero essere state tratte da esso, soprattutto quando si utilizzano algoritmi per il processo decisionale automatizzato, in particolare per la profilazione”. Ad esempio, nel caso di un sistema di *rating* del credito, i mutuatari hanno il diritto di conoscere la logica alla base del trattamento dei loro dati che porta alla decisione di concedere o rifiutare il credito, piuttosto che essere semplicemente informati della decisione stessa. La comprensione di questi elementi contribuisce all’efficace esercizio di altre garanzie essenziali come il diritto di opposizione e il diritto al ricorso all’autorità competente”.

Peraltro, lo stesso concetto di permettere all’interessato di ottenere informazioni significative sulla logica utilizzata nel trattamento che lo riguarda, si rinviengono anche nel GDPR all’art. 13 c.2 lett. f), all’art. 14, c.2. lett. f), all’art. 15 c.1 lett h).

I potenziali benefici di un sistema ODR, il suo grado di integrazione in un processo giudiziario completo (dal precontenzioso al contenzioso effettivo) e il ruolo decisivo dell’intelligenza artificiale nell’esecuzione del processo devono quindi essere adeguatamente valutati caso per caso.

Ora, se è innegabile che l’ODR possa contribuire all’attuazione dei servizi di conciliazione e mediazione prima o durante il giudizio, sarebbe opportuno in ogni caso prevedere una supervisione dei giudici prima che essi decidano sull’esito delle controversie in base al merito, in particolare nelle cause per le quali procedimenti di conciliazione preventivi costituiscono condizioni di procedibilità del giudizio²⁹.

Dall’altro lato occorrerà considerare l’effettivo contributo che i sistemi di intelligenza artificiale possono offrire, vale a dire se lo scopo sia utilizzare il *machine learning* solo per stabilire tabelle indicative o se invece si intenda affidare ad esso la decisione, sempre nel rispetto dei principi di trasparenza, neutralità e lealtà³⁰.

²⁹ Nei Paesi Bassi, i contratti di assicurazione sanitaria privata sembrano prevedere automaticamente il ricorso a un ODR prima che venga intrapresa un’azione legale.

³⁰ C. PAVILLON, *Concerns over a digital judge*, nrc.nl, <https://www.nrc.nl/nieuws/2018/01/19/zorgen-om-populaire-digitale-rechter-a1588963>.

Il gruppo di lavoro sulla mediazione CEPEJ (CEPEJ-GT-MED), lanciato nel 2018, ha offerto le sue prime riflessioni sul contributo della tecnologia dell'informazione ai metodi alternativi di risoluzione delle controversie³¹.

15. Le principali garanzie da riaffermare nei procedimenti civili, commerciali e amministrativi

15.1. Il diritto di accesso a un tribunale

L'utilizzo di strumenti di risoluzione delle controversie online non dovrebbe pregiudicare il diritto di accesso a un tribunale ai sensi dell'articolo 6 della Convenzione europea dei diritti dell'Uomo, anche se questo diritto non è assoluto e si presta a limitazioni implicite³².

In materia civile, ad esempio, ogni parte in giudizio ha il diritto di sottoporre a un tribunale qualsiasi controversia relativa ai suoi "diritti e doveri civili"³³.

Nel 2015, l'Assemblea parlamentare del Consiglio d'Europa ha adottato una risoluzione su "Accesso alla giustizia e Internet: potenziale e sfide" in cui ha invitato a garantire che "le parti coinvolte nelle procedure ODR mantengano il diritto di accedere ad una procedura di ricorso giudiziario che soddisfi i requisiti di un processo equo ai sensi dell'articolo 6 della Convenzione"³⁴.

15.2. Il principio del contraddittorio

Sembra indispensabile fornire ai cittadini e, soprattutto, alle parti di un processo, un certo numero di informazioni quantitative (ad esempio il numero di decisioni elaborate per ottenere la tabella) e informazioni qualitative (origine delle decisioni, rappresentatività dei campioni selezionati, distribuzione delle decisioni tra criteri diversi come il contesto economico e sociale) al fine di comprendere come sono state costruite le tabelle, misurare i loro possibili limiti e poterli discutere davanti ad un giudice.

15.3. Il principio della parità di armi

L'uso di mezzi tecnologici non dovrebbe causare squilibri tra le parti, poiché l'uso di mezzi digitali potrebbe effettivamente favorire taluni operatori (istituzioni, società con mezzi, individui avvezzi all'uso delle tecnologie) e, al contrario, porre in difficoltà alcuni tipi di popolazione, meno abbienti o con meno familiarità con i computer. È importante che nessun individuo rimanga solo davanti ai propri schermi e che venga informato che può chiedere assistenza legale ed essere assistito, ove necessario.

15.4. Il principio di imparzialità e indipendenza del giudice

³¹ Il CDCJ sta attualmente svolgendo un lavoro approfondito sugli ODR per identificare il potenziale di questi strumenti, ma anche eventuali punti problematici che equivalgono a possibili violazioni degli articoli 6, 8 e 13 della Convenzione europea sui diritti umani.

³² Art.6 §1 «1. 1. Ogni persona ha diritto a che la sua causa sia esaminata equamente, pubblicamente ed entro un termine ragionevole da un tribunale indipendente e imparziale, costituito per legge, il quale sia chiamato a pronunciarsi sulle controversie sui suoi diritti e doveri di carattere civile o sulla fondatezza di ogni accusa penale formulata nei suoi confronti». Per le limitazioni v. *Deweert c. Belgique*, § 49; *Kart c. Turquie* [GC], § 67.

³³ CEDH, *Golder c. Royaume-Uni*, §§ 28-36.

³⁴ Resolution 2054 (2015) of the Parliamentary Assembly of the Council of Europe (PACE), 10 November 2015, <http://assembly.coe.int/nw/xml/XRef/Xref-XML2HTML-EN.asp?fileid=22245&lang=en>.

È stato ipotizzato che lo standard maggioritario indotto dall'utilizzo di sistemi di AI possa ripercuotersi negativamente sull'indipendenza e sull'imparzialità della magistratura, in particolare nei sistemi in cui l'indipendenza della magistratura non è pienamente raggiunta. In questi sistemi, non si può escludere il rischio che tali standard possano esercitare pressioni indirette sui giudici a conformarsi all'orientamento maggioritario, a pena di essere sottoposti a verifica da parte del potere esecutivo qualora se ne discostino.

15.5. *Il diritto di assistenza legale*

All'inizio di questo capitolo, abbiamo menzionato i vantaggi derivanti dall'applicazione di strumenti di giustizia predittiva per gli avvocati e, in particolare, la possibilità di fornire ai loro clienti una consulenza meglio informata valutando empiricamente e sistematicamente le possibilità di successo di una procedura. Tuttavia, immaginiamo un caso in cui le possibilità di successo per la parte in causa siano estremamente scarse: ciò potrebbe influire sulla decisione dell'avvocato di assistere il suo cliente, mentre occorrerebbe minimizzare il rischio che le persone che necessitano di assistenza legale possano in definitiva essere private di esso perché anche l'avvocato decide di conformarsi allo standard senza ricercare soluzioni alternative a favore del cliente.

16. *Le questioni specifiche della giustizia penale: prevenzione dei reati, rischio di recidiva e valutazione del livello di pericolo*

Anche se non specificamente progettato per essere discriminatorio, l'uso delle statistiche e dell'intelligenza artificiale nei procedimenti penali ha mostrato il rischio di indurre la rinascita di dottrine deterministiche a discapito delle dottrine dell'individualizzazione della sanzione, che sono state ampiamente acquisite dal 1945 nella maggior parte dei sistemi giudiziari europei.

L'uso della scienza e della tecnologia dell'AI in ambito penale pone questioni specifiche sulla valutazione delle caratteristiche criminali abituali o professionali, sulla tendenza a commettere un crimine, sul carattere e la personalità dell'imputato e, in generale, sulle qualità psicologiche dell'imputato, indipendentemente da cause patologiche.

In Francia, ad esempio, la dottrina della “nuova difesa sociale” sviluppata da Marc Ancel è stata la base del diritto penale: invece di un approccio puramente punitivo e deterministico, è stato introdotto un sistema di riabilitazione sociale per impedire la commissione di un reato evitando le condizioni per il crimine. Questo approccio è condiviso da una serie di strumenti di politica penale europea incentrati sugli obiettivi di rieducazione e reinserimento dei trasgressori³⁵.

Gli strumenti di giustizia penale dovrebbero quindi essere progettati in conformità con questi principi fondamentali di riabilitazione, incluso il ruolo del giudice nell'individualizzazione della pena, sulla base di elementi oggettivi di personalità (formazione, lavoro, medicina regolare e assistenza sociale) senza alcuna altra forma di analisi rispetto a quella effettuata da professionisti appositamente formati. Queste tecniche di analisi dei *big data* potrebbero essere utilizzate da questi professionisti per centralizzare e raccogliere informazioni sulla persona accusata di un reato, e tali dati potrebbero essere archiviati da varie istituzioni e agenzie, a volte per un breve lasso di tempo.

17. *Gli strumenti utilizzati dalle autorità investigative prima del processo penale*

Gli strumenti descritti come “controlli predittivi” (prima del processo giudiziario o prima di un rinvio a giudizio) stanno già crescendo rapidamente e stanno iniziando ad essere conosciuti dal grande

³⁵ V. Corte Europea Diritti dell'Uomo, Grand Chamber, *Vinter and Others vs. United Kingdom*, par. 114-118.

pubblico³⁶. In generale, un gran numero di strumenti informatici è comunemente usato per prevenire la commissione di atti criminali (identificando possibili luoghi in cui ciò potrebbe accadere o i loro autori) o perseguirli in modo più efficace³⁷.

La prima categoria include strumenti di “polizia predittiva” utilizzati per prevenire determinati tipi di reato con elementi di regolarità nella loro manifestazione, come furto con scasso, violenza di strada, furto di veicoli. L'utilità di questi strumenti deriva dalla loro capacità di determinare con precisione dove e quando questi reati potrebbero essere commessi e di riprodurre queste informazioni su una mappa geografica sotto forma di *hot spot* monitorati in tempo reale dalle pattuglie della polizia.

Questo processo è chiamato “mappatura criminale predittiva” e la maggior parte del software utilizzato in questo ambito si basa su prove storiche relative all'ubicazione del crimine, come i rapporti della polizia, ma vengono anche testate nuove tecnologie ancora più potenti che combinano vari dati provenienti da fonti diverse³⁸.

Questi strumenti, che hanno tassi molto persuasivi di efficacia, si dice anche che abbiano effetti dissuasivi sulla commissione di reati nelle aree circostanti i punti caldi, tuttavia, le capacità predittive di questi strumenti, che mostrano i loro limiti in relazione a reati di natura meno regolare o che colpiscono località diverse, come il terrorismo, devono essere attentamente valutate anche perché uno dei loro punti deboli è l'effetto “circolo vizioso” e l'effetto “profezie che si autoavverano”: i quartieri considerati a rischio attirano più attenzione della polizia e la polizia rileva più criminalità, il che porta ad un'eccessiva sorveglianza della polizia delle comunità che vivono in quei luoghi.

Vi è inoltre da considerare che i dubbi su una possibile “tirannia dell'algoritmo” che potrebbe minimizzare, o addirittura sostituire progressivamente, il giudizio umano non sono del tutto assenti all'interno degli stessi servizi di polizia, anche se, al momento, la tecnologia viene presentata come al servizio degli esseri umani al fine di fornire strumenti al processo decisionale³⁹.

L'analisi dei *big data* viene sempre più applicata nel perseguimento del crimine con strumenti come Connect, utilizzato dalla polizia del Regno Unito per analizzare miliardi di dati generati in transazioni finanziarie per trovare correlazioni o schemi di operazioni, o l'*International Child Sexual Exploitation Database*, gestito dall'Interpol, che aiuta a identificare le vittime o gli autori di reati attraverso l'analisi, ad esempio, di mobili ed altri oggetti in immagini offensive o l'analisi del rumore di fondo nei video, che si sono dimostrati particolarmente efficaci nella lotta alla criminalità. Con Connect, ad esempio, le ricerche che in precedenza richiedevano mesi di indagine possono ora essere eseguite in pochi minuti, con un livello molto elevato di complessità e volume di dati.

La dottrina mette comunque in discussione la logica gestionale della risposta al crimine fornita da questi strumenti predittivi, in cui un'analisi approfondita delle ragioni del crimine diventa meno importante del fare qualcosa qui e ora. Ciò accade in un momento in cui i budget disponibili si stanno

³⁶ Si pensi alla *no fly list*, che in realtà è un'applicazione di analisi dei *big data* che raccoglie e analizza i dati su potenziali terroristi al fine di prevenire la commissione di atti criminali o algoritmi utilizzati per rilevare frodi o riciclaggio di denaro.

³⁷ D'altro canto, dovrebbe essere incoraggiato l'uso dell'AI anche a fini terapeutici e riabilitativi (ad esempio per raccogliere dati sul trattamento somministrato o sui metodi di reintegrazione).

³⁸ Ad esempio, nell'ambito del progetto “E-Security - ICT per la sicurezza urbana predittiva” (<http://www.esecurity.trento.it/>), che è stato condotto nella città di Trento tra Novembre 2012 e maggio 2015, è stato creato un database che raccoglie informazioni sui crimini denunciati alla polizia, sui risultati delle indagini condotte dal municipio sulla vittimizzazione e la sicurezza reale e percepita dai cittadini, sul disordine urbano fisico e sociale da parte della polizia, nonché altre variabili relative a “SmartCity” (ad es. informazioni relative al contesto socio-demografico, all'ambiente urbano, all'illuminazione notturna, alla presenza di telecamere di sorveglianza e ai trasporti pubblici). I responsabili del progetto hanno testimoniato l'affidabilità delle tecniche utilizzate, che consentirebbero di prevedere gli atti criminali con un tasso di successo di circa il 60-65%. I test condotti nel Regno Unito nell'ambito di un progetto pilota per prevedere possibili ubicazioni di furto con scasso, furto e rapina dimostrano che le proiezioni del software utilizzate, denominato PREDPOL, erano accurate nel 78% dei casi, rispetto al 51% utilizzando tecniche tradizionali.

³⁹ “How technology is allowing police to predict where and when crime will happen”, The Independent, 7 October 2017.

riducendo e la polizia deve fornire lo stesso livello di protezione pubblica, ma con personale, attrezzature e risorse limitati⁴⁰.

18. Gli strumenti per il processo penale

L'uso di strumenti predittivi⁴¹ da parte dei giudici nei processi penali è molto raro in Europa, ma vi è un esempio degno di menzione: si tratta di HART (*Harm Assessment Risk Tool*), sviluppato in collaborazione con l'Università di Cambridge ed attualmente in fase di test nel Regno Unito che si basa su un sistema di *machine learning* che è stato addestrato utilizzando gli archivi della polizia di Durham dal 2008 al 2012. Apprendendo dalle decisioni assunte dagli agenti di polizia durante questo periodo su soggetti sospetti, la macchina dovrebbe essere in grado di valutare il rischio di recidiva - basso, medio o alto -, basato su una trentina di fattori, alcuni dei quali però incautamente non correlati al crimine commesso (ad esempio il codice postale del luogo di residenza e il genere).

Nei test inizialmente condotti nel 2013, durante i quali è stato osservato un comportamento sospetto in un periodo di due anni dopo la commissione del crimine, le previsioni di HART sono risultate efficaci al 98% nel prevedere un rischio basso, e dell'88% un alto rischio di recidiva. In questa fase sperimentale, HART avrà un valore puramente consultivo per il giudice. Inoltre, la polizia effettuerà periodicamente audit sul funzionamento di HART e sull'affidabilità delle sue conclusioni.

Anche se è l'unico strumento predittivo finora identificato in Europa, offre l'opportunità di considerare le sfide che i decisori pubblici potrebbero affrontare nel prossimo futuro se questo tipo di applicazione venisse testato su scala più ampia, in particolare alla luce delle esperienze americane.

Negli Stati Uniti⁴², l'ONG ProPublica⁴³ ha rivelato gli effetti discriminatori dell'algoritmo utilizzato nel software COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), che mira a valutare il rischio di recidiva al momento in cui il giudice emette sentenza.

Questo algoritmo, sviluppato da una società privata ed utilizzato dai giudici in alcuni stati federali americani, comprende 137 domande a cui il sospetto deve rispondere e informazioni estratte dal casellario giudiziario. Le domande sono piuttosto varie e comprendono la presenza di un telefono a casa, la difficoltà a pagare le bollette, la storia familiare, la storia criminale dell'accusato, ecc.⁴⁴. L'algoritmo valuta la persona su una scala da 1 (basso rischio) a 10 (alto rischio). È un aiuto al processo decisionale giudiziario, le cui conclusioni sono solo una delle variabili considerate dal giudice al momento di decidere la sentenza.

La questione è che alle popolazioni afroamericane è stato assegnato un tasso di recidiva ad alto rischio doppio rispetto a quello di altre popolazioni entro due anni dalla condanna - senza che questo effetto venisse preventivato dai progettisti⁴⁵.

Al contrario, l'algoritmo ha considerato che altre popolazioni sembravano molto meno propense a reiterare un delitto⁴⁶.

⁴⁰ A. ZAVRSNIK, *Big Data, crime and social control*, 196.

⁴¹ In dottrina questi strumenti sono spesso chiamati "*algorithmic justice*" or "*automated justice*", or "*simulated justice*".

⁴² Uno studio del 2015 aveva identificato una sessantina di strumenti predittivi negli Stati Uniti.

⁴³ www.propublica.org/article/technical-response-to-northpointe.

⁴⁴ Esistono altri algoritmi che sono stati sviluppati tenendo conto delle osservazioni critiche espresse dalla dottrina e che si basano su variabili più piccole, più direttamente correlate al crimine commesso e meno legate alla razza, al genere o allo stato socioeconomico. Un esempio è lo strumento di valutazione della sicurezza pubblica utilizzato in 30 giurisdizioni americane.

⁴⁵ Questo effetto discriminatorio può essere spiegato con la "calibrazione" del modello di algoritmo, che crea molti "falsi positivi".

⁴⁶ Le popolazioni di colore sono state più frequentemente classificate come ad alto rischio, anche senza reiterazione entro due anni dalla condanna, mentre le popolazioni bianche sono state classificate più frequentemente come a basso rischio ed hanno commesso reati entro i due anni. In breve, l'algoritmo ha sovrastimato il rischio di recidiva per i neri e lo ha sottovalutato per i bianchi (i "falsi positivi" erano per lo più neri, mentre i "falsi negativi" erano principalmente bianchi). In risposta alle accuse di ProPublica, NorthPointe (ora "Equivant") ha risposto che le popolazioni bianche e nere

Va ovviamente notato che questo tipo di interpretazione fuorviante nella realtà rivela solo la fragilità sociale ed economica di alcuni gruppi di popolazioni che non sono ovviamente criminogene per natura. I ricercatori del *Dartmouth College* hanno anche dimostrato che questo tipo di algoritmo non produce valore aggiunto poiché le persone senza precedenti penali possono riprodurre esattamente la stessa valutazione semplicemente rispondendo al questionario.

Inoltre, la mancanza di trasparenza nei processi operativi dell'algoritmo progettati da società private (che rivendicano la proprietà intellettuale) desta preoccupazione. Poiché le società private *legal tech*, creatrici dell'algoritmo attingono i dati che vanno a costituire la base informativa dalle stesse autorità statali, la loro mancanza di responsabilità nei confronti dei cittadini pone un grave problema democratico, anche perché di norma il pubblico viene informato delle operazioni sui *big data* in modo accidentale, sporadico, in occasione di *data loss* o di *data breach* o di contenziosi giudiziari: un esempio di ciò è quando ProPublica ha rivelato i difetti dell'algoritmo COMPAS a seguito del rifiuto della società proprietaria di condividerlo. La ONG ha dovuto fare appello alle autorità pubbliche affinché accedessero ai dati e inviassero un proprio perito per esaminare l'algoritmo.

19. Le sfide della “predittività” in materia penale

Gli esempi sopra riportati mostrano che possono esserci enormi opportunità ma anche rischi reali nell'applicazione di nuove tecnologie che vengono utilizzate senza le necessarie precauzioni. I responsabili delle decisioni pubbliche e le parti interessate nel settore giudiziario devono essere particolarmente vigili e svolgere un ruolo attivo nello sviluppo di queste tecnologie; il monitoraggio continuo è necessario per determinare la loro reale efficacia ed efficienza ed evitare conseguenze impreviste. Ciò è ancora più importante nei procedimenti penali a causa del loro impatto diretto sulle libertà personali degli individui⁴⁷.

I sostenitori di questi strumenti sostengono spesso che essi sono neutrali e fanno affidamento su metodi concreti e oggettivi che contribuiscono a rendere la giustizia più accurata e trasparente. Un altro grande vantaggio, si sostiene, è la loro efficienza, che a volte supera le capacità umane e può essere estremamente preziosa anche in un contesto generale di riduzione dei fondi pubblici o persino di carenza di risorse.

Ma, d'altro canto, l'inclusione di variabili algoritmiche come la storia criminale e lo sfondo della famiglia significa che il comportamento di un determinato gruppo può decidere il destino di un individuo, che è, ovviamente, un essere umano unico, con un background sociale specifico, istruzione, capacità, grado di colpa e motivazioni per la commissione del reato del tutto peculiari⁴⁸.

Sostengono inoltre che le decisioni umane possono essere basate su valori e considerazioni che non alterano invece il giudizio della macchina, ma questa prerogativa della macchina può rivelarsi

erano ugualmente rappresentate quando si consideravano i “veri positivi”. La questione di come conciliare sia l'accuratezza dell'algoritmo nel rilevare la recidiva, sia la necessità di evitare effetti discriminatori nei confronti delle popolazioni nere è stata la fonte di un intenso dibattito in letteratura; v. in particolare Chouldechova A. (2016), *A fair prediction with a disparate impact: a study on bias in recidivism prediction instruments*, disponibile su <http://arxiv.org/abs/1610.07524>; e anche “Bias in criminal risks scores is mathematically inevitable, Researchers say”, disponibile su <https://www.propublica.org/article/bias-in-criminal-risk-scores-is-mathematically-inevitable-researchers-say>. Questo dibattito riflette anche la questione della legittimità che una società privata, senza alcun controllo istituzionale, possa scegliere tra due requisiti opposti: quello di difendere la società, da un lato, e quello di rispettare i diritti dei singoli, dall'altro.

⁴⁷ Un estratto della decisione della Corte suprema del Wisconsin in “Wisconsin vs. Loomis” può fornire ispirazione anche a livello europeo: “It is important to consider that tools such as COMPAS continue to change and evolve. The concerns we address today may very well be alleviated in the future. It is incumbent upon the criminal justice system to recognize that in the coming months and years, additional research data will become available. Different and better tools may be developed. As data changes, our use of evidence-based tools will have to change as well. The justice system must keep up with the research and continually assess the use of these tools”.

⁴⁸ A. ZAVRSNIK, *Big Data, crime and social control*, 196.

un'arma a doppio taglio. Ad esempio, un giudice potrebbe decidere di ordinare la cauzione di una donna pur con rischio elevato di recidiva, sulla base di una gerarchia di valori, ad esempio attribuendo maggiore importanza al suo ruolo di madre e protettrice dei suoi figli, mentre l'algoritmo potrebbe determinare il rischio di recidiva in modo più accurato, ma non sarebbe in grado di operare una tale gerarchia di priorità.

In materia penale vi sono anche potenziali rischi di discriminazione se si considera che questi strumenti, costruiti e interpretati dagli esseri umani, possono riprodurre disparità ingiustificate già esistenti nel sistema giudiziario penale interessato; così, anziché correggere esiti discriminatori, la tecnologia potrebbe finire per legittimarli.

Alla luce di quanto precede, quando gli algoritmi vengono utilizzati in un processo penale, è essenziale poter garantire pienamente il rispetto del principio della parità delle armi e della presunzione di innocenza sancito dall'articolo 6 della CEDU.

L'interessato dovrebbe avere accesso e poter contestare la validità scientifica di un algoritmo, la ponderazione effettuata sui suoi vari elementi e qualsiasi conclusione errata a cui giunge ogni volta che un giudice suggerisce che potrebbe usarlo prima di una decisione. Come visto sopra tutte le persone hanno il diritto di non essere soggette alle decisioni che le riguardano in modo significativo, basate unicamente su un trattamento automatizzato dei dati, senza che il loro punto di vista sia stato preso in considerazione in anticipo.

A questo proposito, esiste una differenza tra Europa e Stati Uniti per quanto riguarda il diritto di accesso agli algoritmi: mentre negli Stati Uniti le autorità giudiziarie sono ancora riluttanti a riconoscere pienamente il diritto di difesa in contrapposizione con gli interessi privati e, segnatamente, con i diritti di proprietà intellettuale, in Europa il quadro è più protettivo a causa del GDPR, che stabilisce un diritto all'informazione sulla logica sottostante delle decisioni assunte usando algoritmi⁴⁹.

Le considerazioni espresse in precedenza in merito agli effetti potenzialmente negativi di questi strumenti sull'imparzialità del giudice sono valide anche in materia penale: un giudice che decide contro la previsione di un algoritmo rischia di correre rischi assumendosi una maggiore responsabilità. Non sembra irrealistico immaginare che i giudici sarebbero riluttanti ad assumersi questo onere aggiuntivo, in particolare nei sistemi in cui il loro mandato non è permanente ma soggetto al voto popolare⁵⁰, o in cui è possibile che siano chiamati a rispondere della decisione a titolo personale con una responsabilità disciplinare, civile o persino penale.

20. Le questioni specifiche relative alla protezione dei dati personali

Per sfruttare appieno il potenziale degli algoritmi nel rispetto dei principi di protezione dei dati, si dovrebbe applicare il principio di precauzione e mettere in atto politiche preventive per contrastare i potenziali rischi associati all'uso dei dati elaborati da tali algoritmi e l'impatto del loro uso sugli individui e sulla società in generale.

Il principio di liceità nel trattamento dei dati personali e l'obbligo di prevenire o minimizzare l'impatto del trattamento dei dati sui diritti e sulle libertà fondamentali degli interessati dovrebbero indurre una valutazione preliminare del rischio. Ciò dovrebbe consentire di attuare le misure appropriate, in particolare durante la fase di progettazione (*privacy by design*) e sempre (*privacy by default*), al fine di mitigare i rischi identificati.

⁴⁹ Article 15, 1. (h) of EU Regulation 2016/679: "The data subject shall have the right to obtain from the controller"... "the following information:"..." the existence of automated decision-making, including profiling, as referred to in Article 22, paragraphs 1 and 4, and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject".

⁵⁰, Subjectivity, algorithms and the courtroom su https://www.researchgate.net/publication/331283837_Subjectivity_algorithms_and_the_courtroom.

Poiché i dati personali devono essere trattati per scopi determinati e legittimi, non devono essere utilizzati in modo incompatibile con tali scopi e non devono essere ulteriormente trattati in modo tale che l'interessato possa considerare inattesi, inappropriati o discutibili (principio di lealtà), il loro riutilizzo deve essere gestito con la massima cautela.

22. Il potenziale e i limiti degli strumenti di giustizia predittiva

Il termine giustizia predittiva dovrebbe essere respinto perché è ambiguo e fuorviante. Questi strumenti si basano su metodi di analisi della giurisprudenza, utilizzando metodi statistici che non riproducono in alcun modo il ragionamento giuridico ma possono tentare di descriverlo. I pregiudizi analitici, se non possono essere completamente eliminati, devono essere identificati ed il processo di progettazione e l'uso dello strumento devono essere integrati in un chiaro quadro etico.

Già sopra sono stati evidenziati l'ambiguità e l'errore insiti nel concetto di "giustizia predittiva" e come ciò possa operare un fraintendimento nella mente collettiva, portandola a credere che le macchine, prive di qualsiasi emozione, un giorno saranno migliori e in grado di rendere più affidabile l'atto di giudicare.

Oggi più che mai, le sue promesse devono essere esaminate in modo oggettivo e scientifico, basato su solide basi della ricerca fondamentale, al fine di identificare possibili limiti. A tale proposito, va notato che i rischi di interpretazioni distorte del significato delle decisioni giudiziarie sono estremamente elevati, se basati esclusivamente su modelli statistici. Questa osservazione è ulteriormente confermata dalla mancanza di una comprensione precisa dei collegamenti tra i dati e dall'evidente presenza di false correlazioni che non possono essere individuate in grandi masse di dati.

Inoltre, la neutralità degli algoritmi è un mito, poiché i loro creatori trasferiscono consapevolmente o involontariamente i loro sistemi di valori in essi.

Il filosofo Eric Sadin⁵¹, ha osservato che, dietro la loro efficienza ed apparente neutralità vi è una catena di informazioni variabile che va dalla scala di valori del progettista, alla produzione del codice macchina, all'esecuzione dell'algoritmo, al contesto in cui opera, alle attività di manutenzione. Questa osservazione è peraltro condivisa anche dal criminologo Aleš Završnik⁵², che sottolinea come le fasi di costruzione e interpretazione degli algoritmi siano eseguite dall'uomo, per l'uomo e non possano sfuggire a errori, pre-studi, valori, interessi umani e a una rappresentazione umana del mondo, comunque siano concepiti.

Nonostante questi limiti significativi non si dovrebbe trascurare il contributo di una tecnologia con un potere senza pari. I matematici CS Calude e G. Longo⁵³ sottolineano nel loro studio sui *big data* che i loro risultati discutibili, come spesso accade, non devono distruggere la scienza dei dati ma aprire la sfida verso un nuovo metodo scientifico più ampio in grado di incorporare sia nuovi strumenti algoritmici sia strumenti classici, accompagnando il processo con una rigorosa valutazione delle prove.

Come sottolineato nell'introduzione, è probabile che l'uso dell'AI offra un supporto estremamente significativo ai professionisti, compresi giudici e avvocati, ma anche al pubblico in generale, soprattutto se un giorno consentiranno di effettuare ricerche senza pari in materia legislativa, normativa, giurisprudenziale e dottrinale e creare collegamenti dinamici tra tutte queste fonti. Ma questo tipo di applicazione va oltre lo scopo di questo articolo poiché non è progettato per prevedere l'esito di una controversia ma per analizzare la giurisprudenza in un determinato tempo e spazio.

⁵¹ E. SADIN, *La Vie algorithmique. Critique de la raison numérique*, Paris, L'Échappée, coll. «Pour en finir avec», 2015, 288.

⁵² A. ZAVRSNIK, *Big Data, Crime and Social Control*, Routledge Frontiers of Criminal Justice, Taylor & Francis Ltd, 2017, 230.

⁵³ C. S. CALUDE e G. LONGO, *The Deluge of Spurious Correlations in Big Data*, 2017, <https://www.di.ens.fr/users/longo/files/BigData-Calude-LongoAug21.pdf>.

Come già espresso sopra, fatta salva la rappresentatività dei campioni selezionati ed elaborati, l'AI ha contribuito ad elaborare tabelle molto più precise delle somme medie o mediane di denaro stanziato, *mutatis mutandis*, in vari settori (sostegno finanziario, benefici compensativi, risarcimento per lesioni personali, TFR, ecc.). Queste tabelle, che si basano più su un consenso che su un'analisi media di ciò che già esiste, forniscono già un supporto significativo per il processo decisionale e l'orientamento senza voler sostituire la legge stessa.

Come accennato in precedenza, il rischio è che, in assenza di una rappresentazione statistica della realtà, i risultati del software di giustizia predittiva siano stabiliti come standard, senza alcuna convalida da parte dell'ordinamento giuridico e in conflitto con esso.

150. Infine, consideriamo l'idea di poter fare un passo indietro a piacimento dai sistemi predittivi. Invece di bloccare gli utenti in una probabilità (o in un insieme di probabilità), l'idea sarebbe quella di consentire loro di navigare attraverso le correlazioni che hanno portato il sistema a proporre la sua valutazione e di essere in grado di prendere le distanze selezionando altri concetti più pertinenti o gruppi di parole o per escludere false correlazioni.

Per utilizzare l'esempio UCL, ciò consisterebbe nel proporre una rappresentazione grafica dei diversi termini conservati dal sistema (con i rispettivi coefficienti correttori) per scoprire che c'è stata una violazione (o non violazione) e autorizzare altri percorsi da intraprendere proponendo la selezione di altri termini o gruppi lessicali.

Per quanto audace e seducente possa essere questa proposta, presuppone che gli stessi professionisti (giudici, avvocati, università) prendano in carico collettivamente il sistema al fine di verificarne la fattibilità e non consentano agli operatori privati da soli, di progettare software bloccati nel ragionamento o nel calcolo.

Le promesse ambiziose (e non mantenute) di alcune società *legal-tech* non devono oscurare l'immenso potenziale delle tecnologie e la necessità di applicazioni adattate e integrate direttamente con ambienti di ricerca scientifica e accademica, nonché con tutti i professionisti legali, come magistrati, impiegati, avvocati, notai, ufficiali giudiziari ed esperti sul campo.

In questo contesto dinamico, sembra essenziale, in primo luogo, non assumere decisioni affrettate e prendere tempo per discutere in anticipo i rischi e le applicazioni pratiche di questi strumenti per i sistemi giudiziari e testarli nella prima fase.

Un sistema giudiziario in linea con i tempi sarebbe in grado di stabilire, amministrare e garantire una vera cibernetica, sia per il settore pubblico, sia per quello privato e insistere sulla totale trasparenza e correttezza nel funzionamento degli algoritmi, che potrebbe contribuire un giorno al processo decisionale giudiziario.

23. La necessità di un dibattito pubblico per il rispetto dei diritti fondamentali

Le questioni connesse con l'attuazione di strumenti predittivi sono così numerosi e sfaccettati che richiedono un approccio equilibrato da parte dei responsabili delle decisioni pubbliche.

In primo luogo, è essenziale tenere un dibattito pubblico su queste questioni, riunendo sia i progettisti degli strumenti che i professionisti del diritto. I consigli giudiziari, le associazioni dei giudici e quelle degli avvocati possono indubbiamente contribuire a individuare le opportunità e a cercare soluzioni sugli aspetti più controversi.

Inoltre, la formazione giudiziaria e le università di giurisprudenza possono svolgere un ruolo chiave nella sensibilizzazione e divulgazione di questi temi tra i professionisti della giustizia, in modo che tutti gli operatori possano comprendere meglio e contribuire concretamente agli sviluppi.

È inoltre essenziale effettuare ricerche sulle applicazioni proposte e testarle, sia per comprenderne le potenzialità e le debolezze, sia per svilupparle ulteriormente e adattarle alle diverse esigenze.

Il diritto di esaminare le componenti e le caratteristiche degli strumenti proposti dal settore privato (o quelli sviluppati da istituti pubblici indipendenti e specializzati, una soluzione che dovrebbe essere

incoraggiata) sembra altrettanto importante affinché il servizio di giustizia possa svolgere efficacemente la sua missione.

dovrebbe essere condotta una valutazione rigorosa dei risultati dei test prima dell'adozione di tali strumenti. Sembra inoltre fortemente consigliabile valutare regolarmente l'impatto di questi strumenti sul lavoro dei professionisti della giustizia.

24. La creazione di un quadro etico

In primo luogo, la semplice adozione di un quadro legislativo o normativo per l'AI sembra vana in un contesto digitale che è intrinsecamente transnazionale nel suo campo di applicazione. D'altra parte, è probabile che l'attenzione scrupolosa sulla natura e la qualità degli *open data* riduca al minimo i rischi di inadeguate interreferenziazioni e rafforzi la pertinenza dei risultati del trattamento automatizzato. Per quanto riguarda i nomi dei professionisti, una semplice precauzione sarebbe quella di proibire la loro diffusione pubblica in database grezzi strutturati, visti i rischi di uso improprio. Non si tratterebbe di limitare l'accesso alle informazioni già trattate (ad esempio, la composizione di una giuria) ma di filtrare i dati grezzi messi liberamente a disposizione. In breve, occorre distinguere tra l'accesso alle informazioni e l'accesso alle banche dati, che possono essere manipolate a loro volta.

I ricercatori Buttarelli⁵⁴ e Marr hanno sottolineato come i *big data* debbano essere strettamente controllati e protetti. Altri ricercatori (Pasquale e Morozov) hanno sottolineato la necessità di stabilire procedure trasparenti per la distribuzione dei *big data* e più in generale per l'AI in campo giudiziario, perché le soluzioni proposte non possono mai rappresentare la vita nella sua complessità.

Risulta in questo contesto essenziale lo sviluppo di norme di "Ciberetica" per guidare l'attività degli operatori del settore e promuovere i suddetti principi di trasparenza, equità e neutralità dello strumento.

Il monitoraggio regolare da parte di esperti indipendenti dovrebbe garantire che i sistemi di intelligenza artificiale utilizzati per assistere i giudici nelle loro decisioni non siano di parte.

Non è inappropriato anticipare l'implementazione, discreta o meno, dei sistemi di riferimento a pagamento (basati sul modello di pubblicità dei motori di ricerca di Google) che consentano a determinati operatori di dare meno peso alle decisioni che sono loro sfavorevoli. Queste norme svolgeranno un ruolo chiave nell'aumentare la fiducia dei cittadini nei loro sistemi giudiziari. A questo proposito, la qualità dei migliori sistemi potrebbe essere riconosciuta con l'assegnazione di un'etichetta o di una certificazione.

In particolare, i sistemi devono garantire una totale trasparenza e una perfetta equità nel modo in cui le informazioni vengono trattate, sia per i professionisti che per i cittadini, per evitare il ripetersi di errori come l'algoritmo COMPAS negli USA.

I professionisti della giustizia devono essere strettamente coinvolti per poter valutare adeguatamente i rischi e l'impatto di tali applicazioni sui sistemi giudiziari. Al giorno d'oggi, tutti gli esperti coinvolti nello sviluppo dell'AI, compresi ricercatori, ingegneri e sviluppatori di software, hanno responsabilità eccezionali e senza precedenti.

Il loro lavoro potrebbe essere accompagnato da un rafforzamento ancora maggiore delle discipline umanistiche. L'esempio di alcune scuole innovative di sviluppatori di computer mostra che dietro la volontà di "hackerare il sistema" si nasconde, secondo alcuni osservatori, in realtà un pragmatismo privo di qualsiasi contestualizzazione della responsabilità che ora pende su tecnici con poteri quasi-demiurgici.

⁵⁴ G. BUTTARELLI, *European data protection supervisor encourages the new debate on big data*, EDPS, 2015/11 Brussels, 19 novembre su https://edps.europa.eu/sites/edp/files/edpsweb_press_releases/edps-2015-11-edps_big_data_it.pdf.

Il Giuramento di Ippocrate ha certamente i suoi limiti in medicina, ma ritualizza la responsabilità e fornisce un quadro etico. Infine, la “Ciberetica” deve essere accompagnata da una formazione su larga scala delle parti interessate, dai progettisti di algoritmi, alle aziende *legal-tech*, ai loro utenti. Nuove discipline umanistiche interdisciplinari dovrebbero essere messe a disposizione di tutti in modo che l’AI diventi un vettore di sviluppo positivo per l’umanità.