



UNIVERSITÀ DEGLI STUDI DI MILANO

**DIPARTIMENTO DI DIRITTO PUBBLICO
ITALIANO E SOVRANAZIONALE**

Corso di Dottorato in Diritto pubblico, internazionale ed europeo
XXXVII Ciclo

TESI DI DOTTORATO

**LA GIUSTIZIA ALGORITMICA ALLA PROVA DELLE
GARANZIE COSTITUZIONALI**

Tesi presentata da:

Dott. Andrea Valsecchi

Matricola n. R13277

Tutor: Prof.ssa Barbara Randazzo

Coordinatrice: Prof.ssa Francesca Biondi

Anno Accademico 2023/2024

INDICE

CAPITOLO INTRODUTTIVO-METODOLOGICO

LA GIUSTIZIA ALGORITMICA: IL PERIMETRO DELL'INDAGINE E LA PROSPETTIVA COSTITUZIONALISTICA

1. La novità delle questioni e l'intreccio di "linguaggi" diversi	11
2. Algoritmi e intelligenza artificiale	14
2.1. Il problema definitorio dell'intelligenza artificiale	16
2.2. Cenni sul funzionamento degli algoritmi	21
2.3. Definire la giustizia algoritmica	24
3. La prospettiva costituzionalistica della giustizia algoritmica	28
3.1. La prospettiva multilivello	29
3.2. L'influenza dell'algoritmo sulla decisione del giudice	30
3.3. Algoritmi e giusto processo	33
4. L'approccio metodologico	36
4.1. Le difficoltà di ridurre ad unum le esperienze di giustizia algoritmica	36
4.2. Il metodo induttivo e i suoi limiti	38
4.3. Il metodo comparativo	40
4.3.1. (segue) Comparazione con le esperienze europee	40
4.3.2. (segue) L'esperienza di giustizia algoritmica statunitense come benchmark dell'indagine	41
5. Obiettivi e limiti della ricerca	45
5.1. Verso il giusto processo algoritmico	45
5.2. Pluralità di applicazioni, evoluzione tecnologica e quesiti aperti	48
6. Struttura del lavoro	50

PARTE I

**LA GIUSTIZIA ALGORITMICA TRA STATI UNITI ED
EUROPA: STATO DELL'ARTE, ESPERIENZE
PRATICHE E PROSPETTIVE FUTURE**

Premessa: i differenti sviluppi della giustizia algoritmica negli Stati Uniti e in Europa.....53

CAPITOLO I

**LA GIUSTIZIA ALGORITMICA NEGLI STATI UNITI:
L'IMPIEGO DI STRUMENTI PREDITTIVI NEL
PROCEDIMENTO PENALE**

1. La giustizia algoritmica nel processo penale statunitense.....	56
2. Il dibattito filosofico di fondo sull'impiego dei <i>tools</i>	59
3. L'evoluzione storica della valutazione del rischio di recidiva.....	61
4. Il funzionamento dei <i>risk assessment tools</i>	62
5. L'impiego dei tools nel contesto giudiziario statunitense: in particolare il <i>sentencing</i>	67
6. L'acquisizione da parte del giudice del risultato algoritmico	73

CAPITOLO II

**LA GIUSTIZIA ALGORITMICA NEGLI ORDINAMENTI
EUROPEI: IL PARADIGMA DELLA GIUSTIZIA
'PREDITTIVA'**

1. Giustizia algoritmica e giustizia predittiva	75
1.1. Il fondamento teorico della giustizia predittiva in Italia e in Europa.....	78
1.2. Il funzionamento dei moderni sistemi di giustizia predittiva.....	79
2. Studi e sperimentazioni di giustizia predittiva in Europa	83
2.1. Il machine learning e il natural language processing per prevedere le decisioni della Corte europea dei diritti dell'uomo	83
2.2. Il caso francese e gli open data	86

2.3. Esperimenti di giustizia predittiva in alcune giurisdizioni italiane.....	91
--	----

CAPITOLO III

LE PROSPETTIVE FUTURIBILI DELLA GIUSTIZIA ALGORITMICA

1. L'avvento dell'IA generativa nel mercato dell'intelligenza artificiale.....	99
2. L'impiego di strumenti di <i>Generative AI</i> da parte dei giudici	101
3. Brevi riflessioni conclusive: esperienze diverse, diversi interrogativi	104

PARTE II

REGOLARE LA GIUSTIZIA ALGORITMICA: I DIVERSI APPROCCI DI STATI UNITI ED EUROPA

Nota introduttiva: perché regolare l'intelligenza artificiale... ..	108
... e la giustizia algoritmica	112

CAPITOLO IV

LA REGOLAZIONE DELL'IA E DELLA GIUSTIZIA ALGORITMICA NEGLI STATI UNITI

1. La regolazione frammentata dell'intelligenza artificiale negli Stati Uniti..	114
2. La regolazione federale: gli approcci strategici della Casa Bianca.....	117
2.1. La self-regulation come paradigma regolatorio	117
2.2. Verso la co-regolamentazione	122
3. La regolazione statale dei <i>risk assessment tools</i> (rinvio).....	131
4. Quale giustizia algoritmica nel contesto regolatorio statunitense?	131

CAPITOLO V

LA REGOLAZIONE DELL'IA E DELLA GIUSTIZIA ALGORITMICA IN EUROPA

1. La regolazione della giustizia algoritmica tra Unione europea e Consiglio d'Europa.....	135
2. Il percorso di regolazione <i>soft</i> dell'Unione europea sino all'approvazione del regolamento sull'intelligenza artificiale	136
2.1. I primi approcci strategici delle Commissione europea	136
2.2. Un primo strumenti di regolazione <i>soft</i> : gli “Orientamenti etici per un'IA affidabile”	138
2.3. La crescente necessità di un quadro normativo per l'IA: il Libro Bianco sull'intelligenza artificiale	145
3. L' <i>AI Act</i> : la regolazione <i>hard</i> (con elementi <i>soft</i>) dell'intelligenza artificiale	148
3.1. La RoadMap: dalla Proposta della Commissione all'approvazione del Regolamento sull'intelligenza artificiale	148
3.2. Profili generali del Regolamento	150
4. La giustizia algoritmica nel prisma dell' <i>AI Act</i>	160
4.1. La giustizia algoritmica nella Proposta della Commissione e il parere del Comitato economico sociale europeo	161
4.2. L'idea di giustizia algoritmica nelle previsioni del Regolamento	163
4.3. Gli strumenti di previsione del rischio di recidiva all'interno dell' <i>AI Act</i>	167
4.4. L'implementazione dell' <i>AI Act</i> nel disegno di legge di delega italiano sull'intelligenza artificiale	170
5. La recente strategia della Corte di Giustizia Europa sulla giustizia algoritmica	172
6. La giustizia algoritmica nei lavori del Consiglio d'Europa	177
6.1. Il trattato sull'intelligenza artificiale e le specifiche previsioni sulla giustizia algoritmica.....	178
6.2. I lavori della CEPEJ e i principi della Carta Etica sull'impiego dell'IA nei sistemi giudiziari (e il suo seguito)	185
7. Breve nota conclusiva: l'azione regolatoria europea come sentiero sicuro della giustizia algoritmica.....	189

PARTE III

GIUSTIZIA ALGORITMICA E GIUSTO PROCESSO

Nota introduttiva: la giustizia algoritmica nel prisma dell' <i>AI Act</i>	193
--	-----

CAPITOLO VI

I NODI PROBLEMATICI TEORICI E PRATICI DEGLI ALGORITMI IMPIEGATI NELLA GIUSTIZIA

Premessa.....	196
1. Algoritmi e dati	198
1.1. La disponibilità e la selezione dei dati	198
2. L'opacità algoritmica	204
2.1. L'opacità intrinseca dei sistemi di IA	205
2.2. Il ruolo dei privati e il trade secret	207
3. L'autorità dell'algoritmo nei processi decisionali e l' <i>automation bias</i>	210
Breve conclusione: l'intreccio dei nodi.....	213

CAPITOLO VII

IL *DUE PROCESS OF LAW* STATUNITENSE E IL *FAIR TRIAL* NELLA PROSPETTIVA EUROPEA MULTILIVELLO

Premessa.....	215
1. Cenni generali sul <i>due process of law</i> nella Costituzione degli Stati Uniti	218
1.1. Il dibattito giurisprudenziale tra abuse of discretion e violazione del due process of law	218
1.2. L'abuse of discretion standard	220
1.3. Le principali garanzie del due process of law	221
2. Il <i>fair trial</i> nella prospettiva multilivello: le disposizioni rilevanti.....	222

CAPITOLO VIII

L'INCIDENZA DELLA GIUSTIZIA ALGORITMICA SULLA DECISIONE GIUDIZIARIA E LE RICADUTE SULLE GARANZIE DEL GIUSTO PROCESSO

1. Premessa: l'impatto degli algoritmi sulla discrezionalità come prodromo alle violazioni del giusto processo	226
--	-----

SEZIONE I. LA GIURISPRUDENZA AMERICANA SULL'IMPIEGO DEI *RISK ASSESSMENT TOOLS* NEL PROCEDIMENTO PENALE

2. L'insegnamento giurisprudenza statunitense: l'impatto dei <i>risk assessment tools</i> sulla discrezionalità giudiziaria.....	229
2.1. I primi giudizi sull'impiego dei tools algoritmici nel sentencing: la deresponsabilizzazione del giudice.....	231
2.2. (segue) I primi argomenti a favore dell'impiego dei tools nel sentencing	234
3. (Segue) L'insegnamento della giurisprudenza statunitense: i <i>risk assessment tools</i> alla prova del <i>due process of law</i>	239
3.1. (segue) Il caso Loomis	239
3.2. (segue) La giurisprudenza successiva a Loomis	244

SEZIONE II. I DIVERSI LIVELLI DI INFLUENZA DELL'ALGORITMO SULLA DECISIONE DEL GIUDICE

4. L'autorità algoritmica nel procedimento decisionale del giudice	250
4.1. Autorità algoritmica nel giudizio penale nordamericano: la pretesa non decisività dell'algoritmo come garanzia di una sentenza individualizzata	251
4.2. L'influenza dell'autorità algoritmica nelle altre ipotesi di giustizia algoritmica	256
5. Opacità dell'algoritmo ed esternalizzazione della funzione giudicante.....	261

SEZIONE III. LE RICADUTE SULLE GARANZIE DEL GIUSTO PROCESSO

6. La giustizia algoritmica alla prova dell'indipendenza e dell'imparzialità del giudice.....	264
7. La <i>fairness</i> del processo giurisdizionale	271
7.1. La giustizia algoritmica alla prova del contraddittorio e della parità delle armi	272
7.2. La giustizia algoritmica alla prova della motivazione della decisione	284

8. Nota conclusiva: la giustizia algoritmica alla prova del diritto di difesa e a una tutela giurisdizionale effettiva	289
--	-----

PARTE IV

CONCLUSIONI

Nota introduttiva: sciogliere i nodi	295
--	-----

CAPITOLO IX

IL GIUSTO PROCESSO ALGORITMICO E IL RINNOVATO RUOLO DEL GIUDICE

1. Algoritmi e giusto processo: un intreccio possibile	297
2. Controllo del giudiziario sull'algoritmo e sui dati	298
3. La trasparenza algoritmica come tassello irrinunciabile per il rispetto del giusto processo.....	302
3.1. Obblighi di trasparenza per un alghoritmik due process negli Stati Uniti	303
3.2. Obblighi di trasparenza per un giusto processo algoritmico in Europa e in Italia	312
4. Alfabetizzazione e autonomia del giudice	321
 Un tentativo (parziale ma ambizioso) di conclusione: la giustizia algoritmica come opportunità	 329
 <i>Bibliografia</i>	 336

CAPITOLO INTRODUTTIVO-METODOLOGICO

LA GIUSTIZIA ALGORITMICA: IL PERIMETRO DELL'INDAGINE E LA PROSPETTIVA COSTITUZIONALISTICA

SOMMARIO: 1. La novità delle questioni e l'intreccio di “linguaggi” diversi - 2. Algoritmi e intelligenza artificiale - 2.1. Il problema definitorio dell'intelligenza artificiale - 2.2. Cenni sul funzionamento degli algoritmi - 2.3. Definire la giustizia algoritmica - 3. La prospettiva costituzionalistica della giustizia algoritmica - 3.1. La prospettiva multilivello - 3.2. L'influenza dell'algoritmo sulla decisione del giudice - 3.3. Algoritmi e giusto processo - 4. L'approccio metodologico - 4.1. Le difficoltà di ridurre ad unum le esperienze di giustizia algoritmica - 4.2. Il metodo induttivo e i suoi limiti - 4.3. Il metodo comparativo - 4.3.1. (segue) Comparazione con le esperienze europee - 4.3.2. (segue) L'esperienza di giustizia algoritmica statunitense come *benchmark* dell'indagine - 5. Obiettivi e limiti della ricerca - 5.1. Verso il giusto processo algoritmico - 5.2. Pluralità di applicazioni, evoluzione tecnologica e quesiti aperti - 6. Struttura del lavoro

1. La novità delle questioni e l'intreccio di “linguaggi” diversi

Vista la nota difficoltà di delimitare il campo nel contesto ormai “onnivoro” dell'intersezione tra tecnologia e diritto, ai fini di chiarezza e organicità del lavoro, riteniamo necessario dichiarare sin da subito sia di cosa si occuperà questo lavoro e, soprattutto, di cosa *non* si occuperà, con la speranza di sgombrare il campo da possibili fraintendimenti e dubbi.

Il capitolo introduttivo è dunque finalizzato a fornire una minima cornice sulla prospettiva che si intende adottare, sul perimetro dell'indagine e sulla metodologia adottata.

Ed è vero che sono cautele necessarie per qualunque lavoro che intende affrontare scientificamente una problematica giuridica, ma è ancora più urgente delimitare il campo quando si affrontano tematiche – come quelle oggetto del presente lavoro – che affondano le proprie radici non solo nel campo del diritto, ma in altri settori del sapere quali l'ingegneria, l'informatica, l'etica.

Non solo: il tema dell'intelligenza artificiale nel campo della giustizia è relativamente nuovo¹, perché “nuove” solo le potenzialità e i rischi che di recente l'intelligenza artificiale ha mostrato. Essendo un ambito di ricerca ancora tutto da esplorare, e ancora povero di solidi dati di realtà su cui ancorare la riflessione giuridica, esso ha dato spazio – in tutto il mondo come anche in Italia – a ragionamenti di ampio respiro e onnicomprensivi che, sebbene tutt'altro che inutili e privi di interesse, hanno tuttavia scontato una poca chiarezza metodologica proprio per la difficoltà di comprendere i confini attuali dei possibili utilizzi di intelligenza artificiale e algoritmi nel processo giurisdizionale.

La difficoltà di una ricerca giuridica in questo campo si lega altresì alla miriade di “linguaggi” diversi che si intrecciano e rendono complicata la comunicabilità tra diritto e tecnologia.

Ciò avviene su livelli distinti.

Il linguaggio algoritmico è un linguaggio essenzialmente matematico, fatto di codici e formule che la macchina assimila per svolgere i compiti assegnati²: un primo problema che si pone nell'intersezione con il diritto è quindi il fatto che il programmatore si trova innanzi al compito complesso di tradurre il linguaggio giuridico – come noto tecnico e complesso, ed espresso in linguaggio naturale – nel

¹ L'applicazione della tecnologia alle decisioni giudiziarie è tutt'altro che una novità, ma le evoluzioni recenti della tecnologia dell'IA e il fenomeno dei *Big Data* hanno totalmente cambiato la prospettiva, sì da consentire oggi di trattare il tema della giustizia algoritmica come un fenomeno ancora giovane. Per i primi studi sull'uso della tecnologia per prevedere le decisioni dei giudici si cfr. D.J. ELARDO, *A Case for Computers in Law Practice* in *Cleveland State Law Rev.*, 17(3), 1968, pp. 552-566; R.C. LAWLOR, *What Computers Can Do: Analysis and Prediction of Judicial Decisions*, in *Am Bar Assoc J*, 49 (4), 1963, pp. 337-344. Quanto agli studi pionieristici rispetto all'applicazione di tecniche computazionali al ragionamento giuridico si cfr. B.G. BUCHANAN - T.E. HEADRICK THOMAS, *Some Speculation About Artificial Intelligence and Legal Reasoning*, in *Harvard Law Review*, 23, 1970, 40 ss.; L. THORNE MC CARTY, *Reflection on 'Taxman': An Experiment in Artificial Intelligence and Legal Reasoning*, in *Harvard Law Review*, 90, 5, 1977, 837-893. Sulle difficoltà di razionalizzare la discrezionalità tramite IA si cfr. uno scritto precursore di M. TARUFFO, *Judicial Decisions and Artificial Intelligence*, in *Artificial Intelligence and Law*, 1998, pp. 311-324. Persino la possibilità di una sostituzione del giudice con le macchine si radica in ricerche lontane: si cfr. A. D'AMATO, *Can/Should Computers Replace Judges?*, in *Georgia Law Review*, 11, 1977, pp. 1277-1301.

² Problema tipico dell'informatica giuridica: si cfr. G. SARTOR, *Linguaggi (e sistemi) informatici e linguaggio giuridico*, in A. ARTOSI - G. BONGIOVANNI - S. VIDA (a cura di), *Problemi della produzione e dell'attuazione normativa*, III, *Analisi del linguaggio giuridico, legistica e legimatica*, Bologna, 2001.

linguaggio informatico di programmazione, fatto di regole sintattiche e schemi propri³.

Sempre legato al tema della progettazione, è fondamentale costruire sistemi artificiali in grado di *comprendere* il linguaggio (nel nostro caso giuridico), problema che si era posto già agli albori dell'intelligenza artificiale, quando i primi studiosi dichiaravano che avrebbero tentato di fare in modo che le macchine potessero *usare* il linguaggio⁴. Come è stato scritto: «*[p]olicy is often distorted when programmers translate it into code. Although all translations shade meaning, the translation of policy from human language into code is more likely to result in a significant alteration of meaning than would the translation of policy from English into another human language. This is, in part, because the artificial languages intelligible to computers have a more limited vocabulary than human languages. Computer languages may be unable to capture the nuances of a particular policy*»⁵.

Il problema del linguaggio algoritmico si manifesta altresì nella comprensione della tecnologia: in un'era in cui le quest'ultima non è più solo un mezzo per eseguire decisioni ma è essa stessa strumento decisionale⁶, occorre dotarsi di strumenti da una parte per poter *spiegare* la decisione dell'algoritmo, espressa in un linguaggio diverso da quello giuridico, dall'altra per *comprenderla*. Nel campo della giustizia algoritmica tale problematica si riflette, come vedremo, sia sui diritti di chi è soggetto alle decisioni algoritmiche (le parti del processo), sia di chi si serve di strumenti automatizzati per assumere le decisioni (il giudice). In una parola, le differenze di linguaggio imporranno di ragionare sull'*alfabetizzazione* digitale, sia da parte dei giudici, sia in generale della popolazione (si v. *infra*, capitolo IX § 4).

Un ultimo livello di linguaggio impinge direttamente nel ruolo del legislatore: la materia dell'intelligenza artificiale è intrisa di definizioni diverse, ognuna che

³ Sul tema si cfr. G. CAVAGGION - M. OROFINO, *Lingua e costituzione: l'irrompere dei linguaggi algoritmici*, in *Rivista AIC*, n. 4/2023, pp. 155-184.

⁴ J. MCCARTHY - M. MINSKY - N. ROCHESTER - C. SHANNON, *A proposal for the Dartmouth summer research project on artificial intelligence*, 1955, che riprenderemo nel successivo paragrafo.

⁵ D.K. CITRON, *Technological Due Process*, in *85 Wash. U.L. Rev.*, 2008, p. 1261.

⁶ A. SIMONCINI, *Il linguaggio dell'Intelligenza Artificiale e la tutela costituzionale dei diritti*, in *Rivista AIC*, 2023, II, 1-39.

guarda il fenomeno da una differente prospettiva. Ma il ruolo paradigmatico del legislatore è quello formulare leggi generali e astratte, sicché esso ha l'ingrato compito di impiegare delle definizioni che meglio si attagliano ai fenomeni sociali in atto. Tale compito è ancor più complicato di fronte a un fenomeno in continua e rapida evoluzione, che rischia di rendere obsolete, in tempi ristretti, regole e prassi. Proprio in riferimento al nuovo Regolamento sull'intelligenza artificiale (di cui si tratterà diffusamente) siffatto rischio di rapida obsolescenza è stato ampiamente segnalato prima della sua approvazione, sicché il legislatore europeo, come vedremo, ha introdotto meccanismi di facile (si auspica) adattamento alla realtà che si evolve.

Siffatte problematiche, qui rapidamente segnalate, verranno man mano affrontate nel corso del lavoro e successivamente calate nella realtà della giustizia algoritmica, la cui specificità rispetto al tema dell'intelligenza artificiale imporrà attente riflessioni autonome.

Il presente capitolo è quindi finalizzato a guidare il lettore nella comprensione del tema qui trattato, con la speranza che sgombrare il campo da possibili equivoci definitori o contenutistici possa essere d'aiuto per alimentare una discussione proficua su problematiche che inevitabilmente saranno, e sono già ora, all'ordine del giorno per ogni giurista.

2. Algoritmi e intelligenza artificiale

Ci sono delle ragioni per cui abbiamo scelto di titolare questo lavoro “giustizia algoritmica alla prova delle garanzie costituzionali”, senza impiegare il termine “intelligenza artificiale”.

Tale scelta mira – anche furbescamente – a superare un problema definitorio che caratterizza tutti coloro i quali intendono studiare tematiche di contatto con la tecnologia dell'IA, data la mancanza di una definizione condivisa della stessa. Ma la decisione è stata assunta anche per evitare di perdersi in discussioni prive di

sostanza sulla distinzione tra algoritmi “tradizionali”⁷ e algoritmi più “evoluti” che impiegano tecniche sofisticate di *machine learning*⁸ e presentano quindi, come vedremo, delle frizioni con le garanzie costituzionali diverse e autonome.

Tale autonomia non verrà certamente ignorata, ma la ricerca, come si vedrà, ha evidenziato che le problematiche giuridiche afferenti alla giustizia algoritmica spesso prescindono dal fatto che un algoritmo si serva o meno di tecniche evolute di intelligenza artificiale: un esempio noto a chi si occupa del tema in Italia è rappresentato dall’algoritmo impiegato dalla pubblica amministrazione nelle procedure di selezione del corpo docenti nell’ambito della riforma della cd. Buona Scuola, che ha dato origine alla casistica giurisprudenziale più rilevante sul tema della compatibilità costituzionale dell’impiego di algoritmi, benché in un campo diverso da quello della giustizia.

In tale contesto l’algoritmo impiegato era molto semplice e lontano da quello che oggi identifichiamo con intelligenza artificiale, e ciò nonostante la sua utilizzazione, come vedremo, ha destato non poche perplessità sia nella giurisprudenza, che ha posto una serie di limiti all’impiego di algoritmi da parte della Pubblica Amministrazione, sia nel mondo accademico, che non ha mancato di evidenziare le storture di un siffatto impiego (di questo ci si occuperà *infra*, cap. IX § 3.2).

Consapevoli di tali differenze, in questo lavoro signaleremo di volta in volta i casi in cui gli algoritmi più evoluti presentano concretamente delle caratteristiche tali per cui diventa imprescindibile affrontarle sotto una diversa lente giuridica. In

⁷ Nella letteratura giuridica vengono definiti “condizionali” o “deterministici”. Sono algoritmi che utilizzano la cd. *Symbolic AI* (o *Human readable AI*), vale a dire algoritmi che si basano su regole stabilite dal programmatore e assumono decisioni sulla base di modelli predefiniti. A una serie di condizioni l’algoritmo produce il proprio output secondo una logica *if-then*. Il procedimento logico-inferenziale di tali algoritmi è dunque facilmente verificabile, essendo sempre possibile risalire alle condizioni (*id est* i dati) sui cui l’algoritmo ha prodotto la propria determinazione. Per un approfondimento si cfr. E. PERES, *Che cosa sono gli algoritmi*, Salani, Firenze, 2020. Si cfr. anche G. CARULLO, *Decisione amministrativa e intelligenza artificiale*, in *Diritto dell’informazione e dell’informatica*, n. 3, 2021, p. 434.

⁸ Il *machine learning* è definibile come la capacità dei sistemi di IA di «*sophisticatedly perform the task without any intervention of human beings on the basis of learning and constantly increasing experience to understand the problem complexity and need for adaptability*», si v. J. ALZUBI - A. NAYYAR - A. KUMAR, *Machine Learning from Theory to Algorithms: An Overview*, in *J. Phys.: Conf. Ser.*, 2018, pp. 1142 ss.

definitiva, non si farà una netta distinzione tra algoritmi e intelligenza artificiale, se non quando essa sarà necessaria per portare il ragionamento su un piano diverso.

2.1. Il problema definitorio dell'intelligenza artificiale

Come detto, la prima esigenza è quella di smarcarci da un ingombrante problema definitorio: mentre infatti l'algoritmo è un concetto – matematicamente, ma anche nel senso comune – piuttosto definito⁹, l'intelligenza artificiale (d'ora in avanti anche "IA") non ha ancora trovato una definizione internazionalmente condivisa, e addirittura per alcuni l'ambiguità e la divergenza delle definizioni sono funzionali a un uso ideologico della tecnologia, permettendone una reinterpretazione sulla base degli interessi di volta in volta in gioco¹⁰.

Senza qui dover passare in rassegna le miriadi di classificazioni che sono state elaborate rispetto alla suddetta tecnologia, intendiamo prendere in considerazione solo due di esse che contribuiscono a comprendere tanto l'evoluzione che l'intelligenza artificiale ha avuto dall'origine ai giorni nostri quanto la complessità nel racchiudere in una sola definizione un concetto così sfaccettato¹¹.

Nel documento del 1955 che per primo diede il via alla ricerca sull'intelligenza artificiale¹² compare una definizione molto concisa, chiaramente ispirata alle

⁹ Secondo la definizione fornita dall'enciclopedia Treccani, un algoritmo è, in generale, «un procedimento di calcolo esplicito e descrivibile con un numero finito di regole che conduce al risultato dopo un numero finito di operazioni, cioè di applicazioni delle regole». In informatica un «insieme di istruzioni che deve essere applicato per eseguire un'elaborazione o risolvere un problema». Si cfr. <https://www.treccani.it/vocabolario/algoritmo/>.

¹⁰ Y. KATZ, *Artificial Whiteness: Politics and Ideology in Artificial Intelligence*, in *Columbia University Press*, New York, 2020.

¹¹ Una rassegna di tutte le definizioni di intelligenza artificiale raccolte dal 1955 al 2019 si può trovare in una *technical report* della Commissione europea: si cfr. S. SAMOILI - M. LÓPEZ COBO - E. GÓMEZ - G. DE PRATO - F. MARTÍNEZ-PLUMED - B. DELIPETREV, *AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence*, Publications Office of the European Union, Lussemburgo, 2020.

¹² J. MCCARTHY - M. MINSKY - N. ROCHESTER - C. SHANNON, *A proposal for the Dartmouth summer research project on artificial intelligence*, 1955. La nascita dell'intelligenza artificiale come disciplina autonoma viene comunemente collocata all'indomani della pubblicazione di tale scritto, quando un gruppo di scienziati si riunì al Dartmouth College di Hanover nel New Hampshire con lo scopo di introdurre alla comunità scientifica lo studio di tale materia.

ricerche di Alan Turing¹³, il quale si chiedeva in uno scritto «*Can machines think?*»¹⁴.

Sulla scia di tale domanda, nel documento si definiva intelligenza artificiale quel problema «*to be making a machine behave in ways that would be called intelligent if a human were so behaving*». Nella spiegazione della definizione gli Autori spiegano che «*every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves*»¹⁵.

La seconda definizione è invece assai più dettagliata e compare nel nuovo Regolamento europeo sull'intelligenza artificiale (d'ora in avanti anche il "Regolamento" o l'"AI Act"), di recente approvato e pubblicato nella Gazzetta Ufficiale dell'Unione Europea il 12 luglio 2024¹⁶, il quale definisce un sistema di IA «un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali

¹³ Turing enuncia le caratteristiche del suo test attraverso un gioco, quello che egli chiama il *gioco dell'imitazione*. Questo si compone di tre partecipanti, un uomo (A), una donna (B) e una terza persona (C) che Turing chiama *interrogator*. Quest'ultimo è situato in una stanza separata dagli altri due e il suo scopo è quello di comprendere chi dei due partecipanti sia un uomo e chi una donna. Per fare ciò C pone delle domande ad A e B e le loro risposte, per evitare che la voce possa aiutare l'interrogante, devono essere battute a macchina o comunque trasmesse tramite scrittura.

Non è tutto: lo scopo di A nel gioco è quello di ingannare l'interrogante per indurlo a fornire un'identificazione errata mentre lo scopo di B è quello di aiutarlo a dare la soluzione corretta.

Il test si fonda a questo punto sulla sostituzione di A con una macchina. Quello che si chiede Turing è che cosa accadrà se una macchina prenderà il posto di A nel gioco; l'interrogante darà una risposta errata altrettanto spesso di quando il gioco viene giocato tra un uomo e una donna? Ciò che sosteneva Turing è che entro circa 50 anni sarà possibile programmare calcolatori per fare giocare loro il gioco dell'imitazione così bene che un esaminatore medio non avrà più del 70 per cento di probabilità di compiere identificazione esatta dopo 5 minuti di interrogazione. Si v. A.M. TURING, *Computing machinery and intelligence*, in *Mind*, 59, 1950, pp. 433-460.

¹⁴ Si v. A.M. TURING, *Computing machinery and intelligence*, cit.

¹⁵ J. MCCARTHY - M. MINSKY - N. ROCHESTER - C. SHANNON, *A proposal for the Dartmouth summer research project on artificial intelligence*, cit.

¹⁶ Regulation (Eu) 2024/1689 Of The European Parliament And Of The Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (*Artificial Intelligence Act*). Il regolamento sarà oggetto di analisi infra nel cap. V.

previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali» (Art. 3, § 1, n. 1).

Come si intuisce, sono definizioni che si collocano a due estremi opposti: la prima che vede l'intelligenza artificiale come una replica dell'intelligenza umana; la seconda che rinuncia perfino a definirla, assolvendo il compito – più limitato e semplice – di chiarire che cosa si intenda per “sistema” di intelligenza artificiale.

Tale distinzione esiste ancora oggi ed è quella che intende scindere i concetti di cd. *strong AI* (anche *artificial general intelligence - AGI*) e *weak AI* (o *narrow AI*): mentre la prima intende replicare l'intelligenza umana e addirittura, in futuro, superarla, la seconda assolve il più limitato obiettivo di svolgere compiti specifici e circoscritti, molto spesso in modo più rapido e più efficiente dell'essere umano¹⁷. Mentre si scrive un'intelligenza artificiale generale non esiste e quindi non sembra utile trattare la tematica della giustizia algoritmica seguendo tale prospettiva: tuttavia, di recente sono stati presentati al mercato modelli di IA “con finalità generali”¹⁸ (l'esempio più celebre è ChatGPT) i quali, sebbene non siano definibili *intelligenti*, non sono progettati per svolgere un singolo compito in luogo dell'essere umano, ma possono venire impiegati per una serie di compiti differenti (tra cui, per quanto qui interessa, la ricerca giurisprudenziale e la scrittura delle sentenze).

Rispetto alle definizioni che abbiamo riportato, ovviamente, nel mezzo esiste una vasta gamma di ulteriori classificazioni che privilegiano l'una o l'altra prospettiva, anche a seconda della natura delle diverse applicazioni di IA ad oggi in uso e alla materia a cui si riferiscono, che sono in continua (e rapidissima) evoluzione¹⁹: abbiamo scelto le due indicate sia per mostrare il percorso che ha avuto l'elaborazione del concetto dalla sua “invenzione” ad oggi, sia per dimostrare

¹⁷ G. SARTOR, *Intelligenza artificiale e diritto. Un'introduzione*, Giuffrè, Milano, 1996, p.12. Le intelligenze artificiali “forti” descritte dalla fantascienza non esistono. Questo tipo di intelligenza artificiale, dotato non soltanto di intelligenza ma anche di coscienza, è puramente immaginario.

¹⁸ Secondo la definizione fornita dal nuovo Regolamento sull'IA. Su tali modelli ci si soffermerà *infra* (cap. III).

¹⁹ Si cfr. i dati riportati dall'ultimo report dell'Università di Stanford, *Artificial Intelligence Index Report*, 2024, visionabile al seguente indirizzo: https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_2024_AI-Index-Report.pdf.

che quando si parla di intelligenza artificiale si parla di miriadi di cose diverse, e per svolgere una riflessione giuridica sulle conseguenze dell'ingresso dell'IA nella giustizia abbiamo bisogno di una base certa per evitare di perdersi in elaborazioni prive di sostegni solidi.

Ciò si collega anche alla seconda esigenza segnalata: a causa dell'idea che il termine "intelligenza artificiale" ha generato nel senso comune – come conseguenza sia della parola *intelligenza*, che più di tutte dovrebbe definire l'essere umano, sia del seguito letterario e cinematografico che essa ha avuto – il dibattito scientifico ma anche l'opinione pubblica si sono concentrati maggiormente su scenari apocalittici di tensioni tra macchine superintelligenti ed esseri umani²⁰.

Come è stato scritto da alcuni studiosi, «*focusing on the possibility of a rogue superintelligence killing off the human species may itself be harmful. It could distract regulators, the public and other AI researchers from work that mitigates more pressing risks, such as mass surveillance, disinformation and manipulation, military misuse of AI and the inadequacy of our current economic paradigm in a world where AI plays an increasingly prominent role. Refocusing on these present concerns can align the goals of multiple stakeholders and serve to contravene longer-term existential risks*»²¹. E ciò è ancora più problematico se consideriamo che i più potenti sistemi di intelligenza artificiale non possiedono nemmeno le caratteristiche più elementari delle menti umane; non condividono con gli esseri umani ciò che chiamiamo coscienza, la capacità di provare emozioni come l'empatia, il dolore, la gioia e nemmeno possiedono il minimo senso del loro ruolo in questo mondo, tanto meno la capacità di farne esperienza²².

Anche l'elaborazione giuridica, quanto meno agli inizi, è inciampata nel tentativo di mettere in guardia rispetto a scenari che ad oggi non appaiono ancora nemmeno ipotizzabili, come la scomparsa definitiva del giudice a favore di *robot*

²⁰ Su tali temi si cfr. gli scenari suggestivi e spaventosi descritti in N. BOSTROM, *Superintelligence: path, dangers, strategies*, Oxford University Press, New York, 2014.

²¹ B. RICHARDS - B. AGÜERA - Y. ARCAS - G. LAJOIE - D. SRIDHAR, *The illusion of AI's existential risk*, in *Noema Magazine*, 18 luglio 2023.

²² S. VALLOR, *The danger of Superhuman AI is not what you think*, in *Noema Magazine*, 23 maggio 2024.

intelligenti, o l'imminente comparsa di un'AGI in grado di rompere del tutto gli schemi a cui siamo abituati, anche in sede processuale.

Non solo l'accademia, ma anche il legislatore europeo e la giurisprudenza italiana non sono rimasti immuni da tali abbagli. Solo per fare degli esempi reali, in una risoluzione del Parlamento europeo del 16 febbraio 2017 "*Norme di diritto civile sulla robotica*"²³, nel primo *recital* si scrive: «considerando che, dal mostro di Frankenstein ideato da Mary Shelley al mito classico di Pigmalione, passando per la storia del Golem di Praga e il robot di Karel Čapek, che ha coniato la parola, gli esseri umani hanno fantasticato sulla possibilità di costruire macchine intelligenti, spesso androidi con caratteristiche umane».

Come è evidente, sono considerazioni che non hanno nulla di giuridico e, sebbene scritte ormai sette anni fa, ancora non hanno nessun contatto con la realtà dell'IA²⁴, tant'è che l'idea al fondo della risoluzione – dotare i sistemi di IA di autonoma personalità giuridica – non è più sul tavolo della discussione.

Anche il giudice amministrativo italiano sembra essere stato influenzato dalla letteratura fantascientifica quando in una sentenza scriveva che la devoluzione della decisione amministrative a sistemi automatizzati avrebbe realizzato la «deleteria prospettiva orwelliana di dismissione delle redini della funzione istruttoria e di abdicazione a quella provvedimentale»²⁵.

Siffatte erronee considerazioni hanno spostato l'attenzione su problematiche irrealistiche e hanno dato largo spazio a contributi dottrinali non di rado proiettati in un futuro immaginifico. Essi, peraltro, non sono inutili o insignificanti, tutt'altro: hanno il grande merito di animare una discussione sui possibili scenari futuri e sugli strumenti che il diritto possiede (o che dovrà implementare) per guidare tale

²³ PARLAMENTO EUROPEO, *Risoluzione del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*.

²⁴ La risoluzione è stata prontamente criticata da un gruppo di esperti attraverso una lettera aperta indirizzata alla Commissione europea, si cfr. il seguente link: <https://www.politico.eu/wp-content/uploads/2018/04/RoboticsOpenLetter.pdf>.

²⁵ T.A.R. Lazio, Sez. III-bis, sentt. nn. 9924-9230 del 2018. Tali sentenze si inseriscono nel panorama giurisprudenziale già menzionato sull'impiego di algoritmi da parte della PA, e verrà considerato *infra*, cap. IX, § 3.2.

cambiamento (*recte* rivoluzione)²⁶. Allo stesso tempo, tuttavia, si rischia di volgere lo sguardo troppo *al di là della siepe* e di non accorgersi che già oggi l'ingresso (ancora timido) dell'intelligenza artificiale nel campo della giustizia pone al giurista delle sfide complicate, che meritano di essere adeguatamente indagate su un piano di realtà per scongiurare un avvento improvviso di nuove forme di giustizia non accompagnate da adeguate guarentigie giuridiche.

Tutto ciò spiega quindi la scelta di non utilizzare il termine "intelligenza artificiale". Ciò non significa, come detto, che il termine non comparirà nel corso dell'elaborato (anche perché la produzione legislativa degli ultimi anni parla apertamente di IA), ma che non si distinguerà nettamente tra algoritmi e intelligenza artificiale, sia perché la distinzione anche nell'elaborazione scientifica piano piano sta sfumando, sia perché le conseguenze giuridiche, come detto, sono spesso le medesime.

Questo, peraltro, è un approccio impiegato nei recenti studi di diritto costituzionale in cui le problematiche derivanti dall'uso pervasivo della tecnologia sono istruite partendo dai concetti di società algoritmica (*algorithmic society*) invece che di società dell'intelligenza artificiale²⁷.

2.2. Cenni sul funzionamento degli algoritmi

Avendo escluso – quantomeno nel titolo – la parola "intelligenza artificiale", ci rimane di capire cosa sia, in effetti, l'algoritmo.

Per comprendere le problematiche giuridiche che gli algoritmi pongono sul tavolo è necessario possedere una – anche minima – conoscenza di come funziona un algoritmo. Sembra un compito sconcertante per il giurista, non abituato a doversi

²⁶ In questi termini si v. E. LONGO, *La "giustizia digitale" tra nuove forme di efficienza e garanzia dei diritti costituzionali*, in *La Nuova Giuridica - Florence Law Review*, 2023, 2, p. 207. In generale sulla rivoluzione dell'IA si cfr. L. FLORIDI, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Raffaello Cortina Editore, Milano, 2017.

²⁷ Sul punto, si rinvia, *ex multis*, a M. SCHUILENBURG - R. PEETERS (a cura di), *The Algorithmic Society. Technology, Power, and Knowledge*, Abingdon-New York, 2021; AA.VV., *Constitutional Challenges in the Algorithmic Society*, Cambridge, 2021; G. DE GREGORIO, *Digital Constitutionalism in Europe. Reframing Rights and Powers in the Algorithmic Society*, Cambridge, 2022.

interfacciare con un elemento esotico appartenente al campo della progettazione informatica e della logica matematica. Fortunatamente, come è stato peraltro già mostrato dalla riflessione costituzionalista, non è necessario essere ingegneri informatici per ragionare sulle opportunità e sui rischi che l'algoritmo pone nel campo della giustizia, come d'altronde non serve essere ingegneri informatici per comprendere come in generale gli algoritmi stanno plasmando la nostra quotidianità, dal modo in cui prendiamo le nostre decisioni, al modo in cui otteniamo un credito da una banca fino al modo in cui veniamo assunti da un'azienda. Insomma, ci sembra che l'impiego diffuso degli algoritmi nella vita di tutti i giorni ci permetta di considerarli come qualcosa che ormai fa parte del senso comune: si può dire che forse non sappiamo esattamente come funziona un algoritmo, ma se ne vediamo uno sappiamo riconoscerlo.

Ciò detto, un cenno a cosa sia un algoritmo e come esso funziona non può che essere d'aiuto per il lettore. L'algoritmo è un insieme di regole e parametri che un progettista crea attraverso un linguaggio informatico (il "codice") e che un *software* utilizzerà per svolgere i propri compiti, produrre i propri *output*, prendere decisioni.

Gli algoritmi, come noto, sono impiegati ovunque: nei social network che profilano i nostri gusti e ci mostrano contenuti sulla base del nostro traffico su internet; nelle piattaforme di *streaming* che ci consigliano cosa guardare o cosa vedere; in imprese assicurative e banche per determinare i rischi di riscossione dei crediti; nelle amministrazioni pubbliche per le procedure di assunzione; nelle strutture sanitarie per le diagnosi mediche; nel settore finanziario per effettuare scambi sul mercato azionario; nelle indagini e nelle investigazioni per la prevenzione dei reati.

Precisando che l'elenco è solo esemplificativo, gli algoritmi stanno piano piano venendo impiegati anche nel campo giurisdizionale, in diverse forme che vedremo nel corso nel lavoro.

Il grande successo che gli algoritmi stanno riscuotendo nei campi più svariati della società lo si deve al noto fenomeno dei *Big Data*, l'aumento incessante e senza

precedenti della circolazione di dati nel mondo *online*, dovuto a Internet e in generale alla digitalizzazione di quasi tutte le attività che normalmente svolgiamo²⁸.

Non dobbiamo sorprenderci dunque che il campo della giustizia non sia oggi il terreno più fertile su cui sperimentare applicazioni algoritmiche: la situazione in Italia sulla digitalizzazione della giustizia, nonostante il periodo pandemico abbia dato una grossa spinta in tal senso, si trova ancora in una fase iniziale in alcuni settori – si pensi alla digitalizzazione del processo penale – in altri in una fase di aggiustamento e modernizzazione. In una situazione di tal guisa è complicato integrare applicazioni algoritmiche che si nutrono di dati digitali e *software* informatici.

In ogni caso, questa enorme quantità di dati²⁹, unita alla progettazione di *software* molto sofisticati e con potenza di calcolo sempre più elevata permette di affinare le procedure automatizzate e garantire maggiore efficienza tramite l'impiego di algoritmi sempre più accurati, soprattutto quelli evoluti di intelligenza artificiale che tramite procedure di *machine learning* autoapprendono e migliorano senza necessità di un intervento umano, essendo tuttavia più propensi a commettere errori³⁰.

Infatti, mentre gli algoritmi “tradizionali”, alla base dei cd. *sistemi esperti*, vengono progettati indicando nel codice una serie di regole e informazioni che l'algoritmo deve seguire per raggiungere l'obiettivo richiesto, gli algoritmi di *machine learning* sono in grado autonomamente, se addestrati con ampi *dataset*, tramite procedure inferenziali di ricavare direttamente dai dati regole e modelli che poi utilizza per decisioni³¹.

²⁸ Si v. *ex multis*, J.M. BALKIN, *The three laws of robotics in the age of big data*, in, *Ohio State Law Journal*, 78, 2017 e J. VAN DIJCK, *Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology*, in *Surveillance & society*, XII, 2, 2014, pp.197-208.

²⁹ La quantità di dati generata annualmente è cresciuta di anno in anno dal 2010. Si stima che il 90% dei dati mondiali sia stato generato solo negli ultimi due anni (!), si cfr. i dati presenti al seguente indirizzo: <https://explodingtopics.com/blog/data-generated-per-day>.

³⁰ T. HSU - S.A. THOMPSON, *Disinformation Researchers Raise Alarms About A.I. Chatbots*, in *New York Times*, 20 giugno 2023, si cfr. <https://www.nytimes.com/2023/02/08/technology/ai-chatbots-disinformation.html>.

³¹ Ne sono un esempio i modelli di IA generativa come Chat GPT, di cui si parlerà *infra*.

Questi cenni al funzionamento degli algoritmi consentono comunque di capire già quali potrebbero essere i problemi di decisioni pubbliche, come decisioni di natura giurisdizionale, assunte tramite algoritmi: che garanzie dovrà fornire il progettista di tali algoritmi affinché possano venire utilizzati in un settore così delicato come quello della giustizia? Se il *machine learning*, come noto, presenta problemi di trasparenza in quanto non è facile ricavare il procedimento che da alcuni *input* genera il risultato algoritmico, come questo influisce sulle garanzie di pubblicità della giustizia e di conoscenza delle prove? Che competenze dovranno maturare gli operatori per comprendere il funzionamento dei sistemi algoritmici?

Sono solo alcune delle domande che il presente lavoro ha l'ambizione di prendere in considerazione: alcune di esse, come si vedrà, hanno trovato una prima risposta normativa da parte del legislatore europeo, in alcuni casi soddisfacente, in altri ancora non sufficiente; altre sono invece ancora sul tavolo della riflessione dottrinale e meritano di essere considerate dalla prospettiva dello studioso del diritto costituzionale.

2.3. Definire la giustizia algoritmica

Spiegate le ragioni per le quali nel presente lavoro si parlerà di giustizia algoritmica, i problemi teorici e definatori non terminano qui.

L'espressione "giustizia algoritmica" non è infatti priva di sfumature in quanto, soprattutto nella visione che dà la letteratura straniera del termine inglese *algorithmic justice*³², il concetto è ancora poco definito³³. È quindi necessario spiegare come esso sarà inteso in questo lavoro.

Talvolta tale termine viene utilizzato per denunciare le conseguenze negative dell'uso degli algoritmi sul rispetto dei diritti fondamentali, adottando quindi un concetto di giustizia di impostazione valoriale e indicando piuttosto quando l'uso

³² Si veda, per una teorizzazione del concetto, O. MARJANOVIC - D. CECEZ-KECMANOVIC - R. VIDGEN, *Theorizing algorithmic justice*, in *European Journal of Information Systems*, 31(3), 269-287.

³³ T. SOURDIN, *Judges, Technology and Artificial Intelligence. The Artificial Judge*, Cheltenham (UK), 2021, p. 64.

degli algoritmi viene visto come ingiusto³⁴. Da questa prospettiva si vuole gettare luce essenzialmente sulle potenzialità discriminatorie degli algoritmi impiegati nel settore della giustizia, derivate dalla esacerbazione delle iniquità già presenti nei dati reali che vengono impiegati per addestrare l'algoritmo, che non fa altro che riprodurre tale iniquità. In questo senso la giustizia algoritmica sembra sovrapporsi agli studi sui *bias* algoritmici, e non è l'intenzione del presente lavoro.

Altre volte si adotta un concetto di giustizia algoritmica che ricomprende qualunque attività pubblica in cui viene in gioco l'impiego di strumenti automatizzati e di intelligenza artificiale, dall'impiego nelle cd. attività di *polizia predittiva*³⁵ al procedimento amministrativo svolto con l'ausilio di strumenti algoritmici³⁶.

Anche in tale caso, come già indicato, l'impiego di algoritmi in tali settori verrà toccato solo lateralmente, quando le problematiche che esso presenta sono a nostro avviso trapiantabili nel campo giudiziario.

Ancora, si parla di giustizia algoritmica per quelle applicazioni di IA *al di fuori del processo*³⁷, vale a dire le esperienze già esistenti di giustizia automatizzata come alternativa al processo, che nel linguaggio corrente vengono chiamate ODR (*Online Dispute Resolution*), l'evoluzione delle tradizionali ADR (*Alternative Dispute Resolution*)³⁸.

³⁴ Si v., per esempio, L. TOOHEY - M. MOORE - K. DART - D.J. TOOHEY, *Meeting the access to civil justice challenges: digital inclusion, algorithmic justice and Human-Centred Design*, in *19 Macquarie Law Journal*, 133, 2019; K. FREEMAN, *Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis*, in *North Carolina Journal of Law & Technology*, 18, 2016.

³⁵ Si cfr. *ex multis*, C. CASTETS-RENARD, *Human Rights and Algorithmic Impact Assessment for Predictive Policing*, in H.-W. MICKLITZ *et al.* (a cura di), *Constitutional Challenges in the Algorithmic Society*, Cambridge, 2022, 106 ss.

³⁶ Per esempio in C. COGLIANESE - D. LEHR, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, in *University of Pennsylvania Carey Law School, All Faculty Scholarship*, vol. 105, pp. 1147 ss.

³⁷ Utilizza questa classificazione Erik Longo in E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, Milano, 2023, in cui l'A. identifica tre dimensioni di impiego della tecnologia nel processo: il processo dematerializzato, il processo fuori dal processo, il futuro nel processo. Secondo tale classificazione il presente si colloca certamente nell'ultima dimensione indicata da Longo.

³⁸ Si v. principalmente R.E. SUSSKIND, *Online Courts and the Future of Justice*, Oxford, 2019; In Italia si cfr. L. MINGARDO, *Online Dispute Resolution. Involuzioni ed evoluzioni di telematica giuridica*, in P. MORO - C. SARRA (a cura di), *Tecnodiritto: temi e problemi di informatica e robotica giuridica*, Milano 2017, pp. 121 ss.

Meno frequentemente – anche se non mancano impostazioni di tal guisa³⁹ – il termine viene impiegato nel senso più circoscritto dell'utilizzo di algoritmi e dell'intelligenza artificiale all'interno dei sistemi giudiziari e di come tali nuove tecnologie da un lato impattato sulle garanzie costituzionali legate alla giurisdizione e dall'altro contribuiscono a ridefinire l'amministrazione della giustizia e il ruolo del giudice. Ciò in quanto l'esperienza dell'uso dell'IA nel settore della giustizia, come già indicato, è più giovane rispetto all'impiego di algoritmi in altri contesti⁴⁰ e la riflessione giuridica ha per lungo tempo preferito, come già spiegato, parlare di scenari apocalittici legati al giudice-robot e alle sue conseguenze sul piano giuridico.

Il nucleo centrale dell'indagine che verrà qui presentata è invece proprio l'impiego degli algoritmi nel settore della giustizia intesa come funzione giurisdizionale, secondo una concezione del termine giustizia che è fatto proprio anche dalla Costituzione: solo all'art. 11 si fa riferimento alla giustizia come un *valore*⁴¹, mentre nelle disposizioni costituzionali nel Titolo IV in cui compare il termine⁴², essa viene intesa come sinonimo di *giurisdizione*⁴³.

A tal riguardo, sovente la letteratura che si occupa di giustizia algoritmica secondo la prospettiva qui indicata estende la riflessione anche alla *digitalizzazione* della giustizia, includendo nell'indagine il processo telematico, il *case management*⁴⁴, la dematerializzazione dei documenti processuali, la “despazializzazione”⁴⁵ del processo attraverso le tecnologie dell'informazione e

³⁹ A. ZAVRŠNIK, *Algorithmic justice: Algorithms and big data in criminal justice settings*, in *European Journal of Criminology*, 2021, Vol. 18 (5), pp. 623–642.

⁴⁰ T. SOURDIN, *Judges, Technology and Artificial Intelligence. The Artificial Judge*, cit., p. 65

⁴¹ Art. 11 Cost.: «L'Italia ripudia la guerra come strumento di offesa alla libertà degli altri popoli e come mezzo di risoluzione delle controversie internazionali; consente, in condizioni di parità con gli altri Stati, alle limitazioni di sovranità necessarie ad un ordinamento che assicuri la pace e la giustizia fra le Nazioni; promuove e favorisce le organizzazioni internazionali rivolte a tale scopo»

⁴² In particolare gli artt. 101, 102, 103, 108 e 110 Cost.

⁴³ M. LUCIANI, *La decisione giudiziaria robotica*, in *Rivista AIC*, n. 3/2018, p. 877.

⁴⁴ T. SOURDIN, *Judges, Technology and Artificial Intelligence. The Artificial Judge*, cit., pp. 116 ss.

⁴⁵ A. GARAPON, *La despazializzazione della giustizia*, Milano-Udine, 2021.

della comunicazione (cd. *ICT*), considerati come antecedenti e prodromici all'ingresso dell'IA nel processo⁴⁶.

Ora, pur condividendo l'idea per cui la giustizia algoritmica non può e non potrà prescindere da una sempre più diffusa giustizia digitale, riteniamo che per i risultati importanti già ottenuti dalla dottrina su queste tematiche – soprattutto a seguito dell'accelerazione verso la digitalizzazione della giustizia avvenuta durante il periodo pandemico in Italia e nel mondo – non sia qui necessario muovere l'indagine da questi fenomeni già profondamente studiati da altri Autori⁴⁷ e ormai generalizzati in tutto il mondo, con diversi livelli di efficienza⁴⁸. Inoltre, come si vedrà nel corso della trattazione, un lavoro monografico sulla giustizia algoritmica per come intesa qui possiede i sostegni per reggersi autonomamente e benché, come anticipato, ci troviamo ancora agli inizi di questa rivoluzione, la normativa, la giurisprudenza e la prassi – basi ancora oggi imprescindibili per il giurista – appaiono già da sole idonee per sviluppare un discorso giuridico approfondito e sistematico.

Per tali ragioni le applicazioni algoritmiche impiegate al di fuori del processo giurisdizionale – tra cui le applicazioni di IA per la risoluzione alternativa delle controversie, gli algoritmi impiegati nei procedimenti amministrativi, o l'IA impiegata da imprese o studi legali per migliorare l'efficienza del lavoro⁴⁹ – non verranno *specificamente* analizzate nel corso dell'elaborato. Piuttosto, la loro considerazione in alcuni capitoli sarà finalizzata a dotarci di uno sguardo prognostico verso le applicazioni di IA nel campo giudiziario, come detto ancora non pervaso da strumenti automatizzati.

⁴⁶ Sul tema della digitalizzazione della giustizia si v. E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit.; A. GARAPON - J. LASSÈGUE, *Justice digitale*, PUF, Paris, 2018.

⁴⁷ Vedi quelli citati *supra* alle note 37 e 38.

⁴⁸ M. LIBERTINI - M. ROSARIA MAUGERI - E. VINCENTI, *Intelligenza artificiale e giurisdizione ordinaria. una ricognizione delle esperienze in corso*, in A. PAJNO - F. DONATI - A. PERRUCCI (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, volume 2, 2022, pp. 483-514.

⁴⁹ Fenomeno che oggi cade sotto il nome di *legal tech*.

3. La prospettiva costituzionalistica della giustizia algoritmica

Delimitato il campo della giustizia algoritmica nel presente lavoro, occorre chiarire in ultimo la prospettiva da cui si guarda il fenomeno e su cui si è costruita la ricerca.

Lo studio della giustizia algoritmica può assumere le più svariate forme, che ovviamente dipendono dal settore del diritto da cui si intende indagarla. Non è un caso che di intelligenza artificiale applicata alla giustizia si siano occupati non solo i costituzionalisti, ma anche studiosi di diritto penale⁵⁰, processualisti (sia penalisti⁵¹ sia civilisti⁵²) nonché, ovviamente, gli studiosi di diritto amministrativo sia sostanziale sia processuale⁵³.

⁵⁰ Tra gli altri si cfr. V. MANES, *L'oracolo algoritmico e la giustizia penale*, in *DisCrimen*, 15 maggio 2020; F. BASILE, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Dir. pen. e uomo*, 29 settembre 2019; C. BURCHARD, *L'intelligenza artificiale come fine del diritto penale? Sulla trasformazione algoritmica della società*, in *Riv. it. Dir. Proc. Pen.*, vol. 4, 2019; O. DI GIOVINE, *Il judge bot e le sequenze giuridiche in materia penale (intelligenza artificiale e stabilizzazione giurisprudenziale)*, in *Cass. pen.*, 2020, pp. 951 ss.

⁵¹ S. QUATTROCOLO, *Artificial Intelligence, Computational Modelling and Criminal Proceedings. A Framework for A European Legal Discussion*, 2020; ID, *Quesiti nuovi e soluzioni antiche? Consolidati paradigmi normativi vs rischi e paure della giustizia digitale "predittiva"*, in *Cassazione penale*, 4, 2019, 1748 ss.; ID, *Equità del processo penale e automated evidence alla luce della convenzione europea dei diritti dell'uomo*, in *Revista Ítalo-Española de Derecho Procesal*, 2019, pp. 107-123; ID, *Equo processo penale e sfide della società algoritmica*, in *BioLaw Journal – Rivista di BioDiritto*, n. 1/2019, pp. 135-144; M. GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti e Europa*, in *Dir. Pen. Cont.*, 29 maggio 2019; G. UBERTIS, *Intelligenza artificiale, giustizia penale, controllo umano significativo*, in *Sistema penale*, 11 novembre 2020; E. KOSTORIS, *Intelligenza artificiale, strumenti predittivi e processo penale*, in *Discrimen*, 5 marzo 2024.

⁵² A. CARRATTA, *Decisione robotica e valori del processo*, in *Riv. dir. proc.*, 2/2020; U. RUFFOLO, *Giustizia predittiva, decisione algoritmica e ruolo ancillare della machina come ausiliare necessario del giudice umano*, in A. PAJNO - F. DONATI - A. PERRUCCI (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, Bologna, Il Mulino, 2022, pp. 453-481; G. RESTA, *Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza*, in *Politica del diritto*, 2019, p. 214.

⁵³ Tra i tanti si cfr. D.U. GALETTA, *Algoritmi, procedimento amministrativo e garanzie: brevi riflessioni, anche alla luce degli ultimi arresti giurisprudenziali in materia*, in *Rivista Italiana di Diritto Pubblico Comunitario*, 3, 2020, pp. 501 ss.; P. OTRANTO, *Riflessioni in tema di decisione amministrativa, intelligenza artificiale e legalità*, n. 7/2021; S. C. MATTEUCCI, «Umano troppo umano». *Decisioni amministrative automatizzate e principio di legalità*, in *Riv. Dir. Pubbl.*, 1/2019; L. MUSSELLI, *La decisione amministrativa nell'età degli algoritmi: primi spunti*, in *MediaLaws-Rivista di diritto dei media*, 1/2020; F. PATRONI GRIFFI, *Decisione robotica e giudice amministrativo*, in A. CARLEO (a cura di), *Decisione robotica*, Il Mulino, Bologna, 2019.

La difficoltà, quindi, di identificare chiaramente una disciplina privilegiata o designata per studiare la giustizia algoritmica ci obbliga a chiarire quale sarà la prospettiva di questo lavoro.

L'intento del paragrafo è quindi vedere quali strumenti ha il diritto costituzionale per accompagnare il percorso della giustizia algoritmica, tenendo ben presenti le basi su cui si fonda da sempre il diritto costituzionale, come incisi nell'art. 16 della Dichiarazione dei diritti dell'uomo e del cittadino del 1789: la garanzia dei diritti e la separazione (e quindi limitazione) dei poteri⁵⁴.

3.1. La prospettiva multilivello

Occorre fin da subito chiarire che le problematiche della giustizia algoritmica verranno istruite in chiave costituzionalistica secondo la prospettiva del diritto costituzionale multilivello, oggi non più trascurabile e che plasma non solo il discorso sulla garanzia dei diritti ma anche, come sostenuto da autorevole dottrina, il rapporto tra i diversi poteri⁵⁵. Soprattutto, l'incidenza del diritto internazionale e del diritto europeo (sia in ambito eurolunitario sia nell'ambito del Consiglio d'Europa) sul diritto nazionale – in ogni suo formante, legislativo, giurisprudenziale e dottrinale – e la conseguente influenza sul diritto nazionale delle giurisprudenze delle Corti europee, in particolare la Corte di giustizia dell'Unione Europea e la Corte Europea dei diritti dell'uomo, obbliga oggi lo studioso del diritto ad adottare tale impostazione.

A maggior ragione in un fenomeno mondiale come quello dell'intelligenza artificiale un'analisi completa in chiave costituzionalistica della giustizia algoritmica non può prescindere dal contesto giuridico europeo. A questo riguardo anche la giurisprudenza europea e convenzionale – per quanto oggi non

⁵⁴ Art. 16: «*Toute société dans laquelle la garantie des droits n'est pas assurée ni la séparation des pouvoirs déterminée, n'a point de Constitution*».

⁵⁵ Si cfr. le lucide riflessioni svolte in B. RANDAZZO, *Access to Justice in a Multilevel Constitutional System, Protecting Human Rights*, Giappichelli, 2024 nonché in ID, *Imparzialità del giudice e fiducia nella magistratura nella prospettiva costituzionale europea*, in *Questione Giustizia*, 1/2/2024, pp. 231-241. V. anche P. BILANCIA, F.G. PIZZETTI, *Aspetti e problemi del costituzionalismo multilivello*, Milano, Giuffrè, 2004.

specificamente indirizzata a sindacare la legittimità degli algoritmi impiegati in sede giurisdizionale ma in generale alle problematiche concernenti gli algoritmi impiegati per assumere decisioni – sarà importante per indirizzare il percorso.

Tale approccio, inoltre, assume oggi una decisività ancora maggiore per l'ormai sviluppato interesse dei legislatori europei rispetto al tema oggetto della presente indagine. Nell'ambito dell'Unione europea, le norme sulla giustizia algoritmica previste nel Regolamento sull'IA, per come verranno interpretate dal giudice europeo, guideranno l'interpretazione della normativa nazionale e costituiranno altresì parametro interposto di legittimità costituzionale delle norme interne. Nell'ambito del Consiglio d'Europa, invece, il 17 maggio 2024 è stata approvata dal Comitato dei ministri la “*Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*”, primo trattato internazionale sull'intelligenza artificiale che da settembre 2024 è aperto alla firma degli Stati⁵⁶ e avrà senz'altro un impatto sulle giurisprudenze e sulle legislazioni nazionali ed europee oltre che sulla stessa giurisprudenza delle Corti sovranazionali.

Infine, come già accennato, l'impiego di algoritmi nei sistemi giudiziari varia ancora molto da Stato a Stato, da continente a continente. In Italia, oltre a sporadiche sperimentazioni in alcune Corti di cui si darà conto non esistono ancora vere e proprie esperienze su cui ragionare in chiave costituzionalistica. Questo ci obbliga, sempre in una visione integrata del diritto costituzionale, a indagare esperienze che provengono da altri ordinamenti nazionali: le incursioni di esperienze e giurisprudenze di altri Stati europei, oltre a un'indagine approfondita sul contesto statunitense, avranno lo scopo di alimentare una discussione quanto più possibile empirica e fondata su esperienze reali, per valutare la conformità di tali esperienze alla luce delle garanzie costituzionali.

3.2. L'influenza dell'algoritmo sulla decisione del giudice

⁵⁶ Mentre scriviamo hanno firmato il trattato Unione europea, Stati Uniti, Regno Unito, Andorra, Georgia, Islanda, Norvegia, Moldavia, San Marino e Israele.

Senza sottacere delle numerose problematiche – anche di natura costituzionale – che permeano in generale il tema della tecnologia applicata alla giustizia, riteniamo che l’approccio più coerente con la prospettiva costituzionalistica segnalata della giustizia algoritmica sia indagare l’impatto che gli algoritmi possono avere sul percorso decisionale del giudice.

Come nel tempo siamo stati spogliati della prerogativa di esclusivi decisori, proprio in virtù dello sviluppo tecnologico e del manifestarsi di complessi procedimenti algoritmici in grado di “decidere per noi” o, comunque, di incidere profondamente sul nostro processo decisionale, così anche il ruolo del giudice è destinato a scontrarsi con tale trasformazione⁵⁷.

All’interno dello spazio giudiziario, le applicazioni di IA che già sono state utilizzate o quantomeno sperimentate spaziano da strumenti che incidono direttamente sulla decisione del giudice a strumenti che riguardano più che altro l’organizzazione del “retro-bottega” dei tribunali (pensiamo alla gestione dei fascicoli, a strumenti di sintesi e traduzioni dei documenti, a sistemi automatizzati di gestione del personale).

Senza qui voler svilire l’importanza di tutto ciò che giace dietro al processo come rito formalizzato destinato a una decisione sostanziale, maggiore priorità nel corso del lavoro verrà data a quelle applicazioni che finora hanno dimostrato di poter porre sotto maggiore *stress* la tenuta delle garanzie costituzionali legate alla giurisdizione e al giusto processo, che sono quelle idonee ad incidere sulla decisione giudiziaria. E l’attenzione particolare alle garanzie processuali e la loro tenuta di fronte all’evoluzione tecnologica si giustifica per l’importanza che ha assunto oggi – a fronte della globalizzazione, della proliferazione delle fonti nazionali e sovranazionali⁵⁸, della crisi della legge⁵⁹ e il contestuale affermarsi del

⁵⁷ R. PARDOLESI - A. DAVOLA, *Algorithmic legal decision making: la fine del mondo (del diritto) o il paese delle meraviglie?*, in *Questione Giustizia*, trimestrale n. 1, 2020 pp. 104 ss.

⁵⁸ Si cfr., *ex multis*, G. CONTI, *Il sistema di tutela multilivello e l’interazione tra ordinamento interno e fonti sovranazionali*, in *Questione Giustizia*, n. 4/2016, pp. 89-114.

⁵⁹ N. IRTI, *La crisi della fattispecie*, in ID., *Un diritto incalcolabile*, Torino, Giappichelli, 2016.

cd. diritto giurisprudenziale⁶⁰ – il momento processuale⁶¹, il cui accesso è, come noto, sia un diritto fondamentale autonomo sia lo strumento per la tutela di tutti gli altri diritti fondamentali⁶².

L'importanza del processo come strumento oggi decisivo per la tutela dei diritti giustifica quindi una trattazione specifica sulle implicazioni della giustizia algoritmica come qui definita e delimitata, tenendo a mente, come è stato autorevolmente affermato, che il processo e le sue garanzie rappresentano una «cartina di tornasole della qualità della vita democratica»⁶³.

L'impostazione qui segnalata non è peraltro il frutto della deliberata volontà di rinunciare alla trattazione di alcuni argomenti che spesso vengono fatti rientrare dalla letteratura giuridica nel campo della giustizia algoritmica: al contrario, essa si fonda, oggi, sul dato legislativo. Come si dirà diffusamente (*infra*, cap. V) il Regolamento sull'intelligenza artificiale approvato di recente classifica come “ad alto rischio” i sistemi di IA impiegati nel campo della giustizia (Articolo 6 e Allegato III al Regolamento). E tuttavia, all'art. 6 § 3 del Regolamento si precisa che, in deroga alla classificazione fornita nel regolamento, non sono considerati ad alto rischio i sistemi che non presentano «un rischio significativo di danno per la salute, la sicurezza o i diritti fondamentali delle persone fisiche» e ciò «anche nel senso di non influenzare materialmente il risultato del processo decisionale».

In tutti questi casi il Regolamento presume quindi che un sistema astrattamente ad alto rischio non abbia un'influenza sul processo decisionale dell'utilizzatore e quindi non debba essere classificato come *high risk*.

Seguendo quindi l'impostazione del legislatore europeo, anche nel presente lavoro l'attenzione sarà dedicata all'impatto della giustizia algoritmica *dentro* il

⁶⁰ Sull'incidenza del diritto giurisprudenziale nel discorso sulla giustizia algoritmica si v. A. PAJNO, *Intelligenza artificiale e autonomia del giudice. il ruolo del fatto e il valore del precedente*, in A. PAJNO - F. DONATI - A. PERRUCCI (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, volume 2, 2022, pp. 389-432.

⁶¹ A. DI PORTO, *Calcolo giuridico secondo la legge nell'età della giurisdizione, il ritorno del testo normativo*, in A. CARLEO (a cura di), *Calcolabilità giuridica*, Il Mulino, Bologna, 2017.

⁶² B. RANDAZZO, *Accesso alla giustizia e diritti fondamentali nella dimensione costituzionale nazionale e sovranazionale*, in C. AMALFITANO - D.U. GALETTA - L. VIOLINI, “*Law, Justice and sustainable development*”. *L'accesso alla giustizia nel quadro del Sd goal 16*, 2022 pp. 117-138.

⁶³ A. PAJNO, *Intelligenza artificiale e autonomia del giudice. il ruolo del fatto e il valore del precedente*, cit., p. 391.

processo strettamente inteso e soprattutto al suo impatto sul *decidere* giudiziario, proprio perché riteniamo, come fa il Regolamento sull'intelligenza artificiale, che il rischio per la tenuta delle garanzie costituzionali sia – per queste applicazioni – alto.

3.3. Algoritmi e giusto processo

Una volta inserita l'indagine nel quadro del costituzionalismo multilivello e indicata la chiave del lavoro (l'influenza degli algoritmi sul processo decisionale del giudice), occorre specificare come verrà affrontata da questa prospettiva la giustizia algoritmica. Quello che è emerso nella ricerca, sia attraverso lo studio dell'esperienza statunitense, sia seguendo l'impostazione seguita dalla letteratura costituzionalistica italiana ed europea, è che il punto nodale dell'impiego di algoritmi nella giurisdizione ha principalmente a che vedere con le garanzie del giusto processo.

Senza voler qui anticipare i problemi che la giustizia algoritmica ha già mostrato di poter sollevare nel procedimento giudiziario, è utile nel presente capitolo introduttivo dare conto dei principali interrogativi.

Essi si riferiscono sia alle caratteristiche degli algoritmi e, in generale, dell'intelligenza artificiale, sia alla direzione che ha intrapreso l'industria dell'IA. Quanto al primo aspetto, gli algoritmi più evoluti presentano non trascurabili problemi di trasparenza, che portano con sé conseguenze rilevanti rispetto alla comprensibilità dell'*output* e del procedimento seguito dal sistema algoritmico (si v. *infra*, cap. VI § 2). Quanto al secondo aspetto, soprattutto per i costi elevatissimi che oggi l'intelligenza artificiale richiede per la sua progettazione e funzionamento, si è passati da un forte ruolo pubblico nella progettazione dell'IA a un'industria trainata principalmente da investimenti privati (si v. *infra*, cap. VI §§ 1-2). Come si vedrà, tali caratteristiche, emerse sia nell'indagine sulla giustizia algoritmica statunitense sia nelle più limitate esperienze europee, portano con sé conseguenze ulteriori da indagare, quali la raccolta e l'elaborazione dei dati che addestrano gli

algoritmi, l'opposizione delle società dei diritti di privacy, il controllo sui progettisti degli algoritmi.

Inoltre, l'avvento di strumenti tecnologici all'interno del processo, in particolare modo di algoritmi predittivi circondati da un'aura di *esattezza* e *infallibilità*, comporta la tendenza di coloro i quali sono chiamati ad utilizzarli a fare totale affidamento sullo strumento (cd. *anchoring effect*), soprattutto se di quello strumento non conoscono a fondo le caratteristiche (si v. *infra*, cap. 6 § 3).

Strumenti con siffatte caratteristiche comportano, come si intuisce, una serie di interrogativi rispetto alla compatibilità del loro impiego all'interno del processo con le garanzie costituzionali.

Tale compatibilità è stata oggetto di un ampio, e spesso trascurato, sindacato da parte delle Corti nell'esperienza ad oggi più evoluta di giustizia algoritmica, quella statunitense. Come si vedrà, l'oggetto del sindacato dei giudici americani ha assunto due strade distinte: prima l'abuso di discrezionalità del giudice (*abuse of discretion*) nel momento in cui questi fonda la propria decisione (anche) sulle risultanze di un algoritmo predittivo e successivamente la compatibilità costituzionale di tali algoritmi con il *due process* previsto dal V e dal XIV emendamento della Costituzione americana.

Partendo quindi dall'esperienza americana sarà possibile astrarre le conclusioni raggiunte rispetto al sistema giudiziario statunitense alla compatibilità di algoritmi decisionali rispetto al giusto processo nella prospettiva costituzionale multilivello.

Il *focus* di questo lavoro sarà pertanto il giusto processo, per come ricostruito nella Costituzione e nelle Carte dei diritti (CEDU e CDFUE) e interpretato dalle rispettive Corti (centro nevralgico sarà necessariamente l'art. 6 della Convenzione, per il contributo imprescindibile che ha fornito la Corte europea dei diritti nella ricostruzione del giusto processo e delle sue garanzie).

Tale impostazione trova peraltro conforto ora nel dato normativo, a dimostrazione che anche il legislatore europeo ha considerato particolarmente problematico l'impiego dell'IA nell'amministrazione della giustizia proprio rispetto alle garanzie processuali.

Al *considerando n. 48* del nuovo Regolamento, nell'enucleazione dei diritti garantiti dalla Carta dei diritti fondamentali che secondo il legislatore europeo giustificano la classificazione del sistema di IA ad alto rischio, compaiono «il diritto a un ricorso effettivo e a un giudice imparziale, i diritti della difesa e la presunzione di innocenza».

Ancora, per quanto concerne sistemi impiegati dalle autorità di contrasto per la prevenzione dei reati⁶⁴, il Regolamento specifica che a causa del loro impiego «potrebbe inoltre essere ostacolato l'esercizio di importanti diritti procedurali fondamentali, quali il diritto a un ricorso effettivo e a un giudice imparziale, nonché i diritti della difesa e la presunzione di innocenza, in particolare nel caso in cui tali sistemi di IA non siano sufficientemente trasparenti, spiegabili e documentati» (*considerando n. 59*).

Infine, anche specificamente per gli strumenti di IA impiegati dal giudice nell'attività giurisdizionale, essi sono da classificare ad alto rischio «in considerazione del loro impatto potenzialmente significativo [...] sulle libertà individuali e sul diritto a un ricorso effettivo e a un giudice imparziale» (*Considerando n. 61*).

È lo stesso regolamento, quindi, a identificare alcune delle garanzie che ritiene particolarmente sotto scrutinio in conseguenza dell'intervento di decisioni parzialmente automatizzate: in particolare il diritto di difesa, il diritto a un ricorso effettivo e l'imparzialità del giudice.

Come emergerà nel corso del lavoro, in realtà la cospicua casistica statunitense ha messo in luce le possibili incompatibilità anche con ulteriori garanzie processuali, là analizzate sotto la lente del processo penale ma non di rado estendibili alla generalità dei processi, quali per esempio l'indipendenza del giudice e il suo obbligo di motivare i provvedimenti, il diritto alla pubblicità delle prove e all'affidabilità di esse, il diritto al contraddittorio e la parità delle armi in sede processuale.

⁶⁴ Rientrano in questo contesto i cd. *risk assessment tools*, che negli Stati Uniti vengono utilizzati dentro al processo e che saranno oggetto di indagine approfondita (si v. *infra*, cap. I).

4. L'approccio metodologico

4.1. Le difficoltà di ridurre ad unum le esperienze di giustizia algoritmica

Pur avendo perimetrato l'indagine del presente lavoro alle applicazioni algoritmiche che in qualche modo hanno un'influenza sul processo decisionale del giudice, la ricerca ha dovuto fare i conti con il fatto che la giustizia algoritmica assume forme molto diverse, ed ognuna di esse presenta delle frizioni con le garanzie costituzionali difficilmente riducibili a unità.

Inoltre, come già accennato, spesso la letteratura costituzionalistica italiana ha tentato di adottare un approccio *top-down*, ossia cercando di calare dall'alto i principi e i diritti costituzionali alla nuova possibile realtà di applicazioni automatizzate nel processo, qualunque esse siano.

Infine, sulle riflessioni – e sugli allarmi – hanno non poco influito alcuni documenti che hanno presentato la giustizia algoritmica come un fenomeno già sostanzialmente in atto nelle giurisdizioni europee, creando qualche fraintendimento negli studi specialistici. Tra essi un ruolo senz'altro decisivo lo ha assunto la *Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi*, approvata nel 2018 dalla CEPEJ nell'ambito del Consiglio d'Europa, primo documento europeo che ha accostato l'IA e il procedimento giurisdizionale. Già il titolo del documento conduce a pensare che la giustizia algoritmica sia in effetti un fenomeno già istituzionalizzato⁶⁵, quando è invece è la stessa Carta a precisare che al 2018 – e a oggi la situazione non è sostanzialmente cambiata, salvo nuove sperimentazioni di cui si darà conto *infra* (cap. II) – «l'utilizzo di algoritmi di intelligenza artificiale nei sistemi giudiziari europei rimane principalmente un'iniziativa commerciale del settore privato destinata alle società assicuratrici, agli uffici legali, agli avvocati e alle persone»⁶⁶.

⁶⁵ Nello stesso senso si v. A. SANTOSUOSSO - G. SARTOR, *La giustizia predittiva: una visione realistica*, in *Giur. It.*, 7, 2022, pp. 1759 ss.

⁶⁶ Carta Etica, p. 15.

Nonostante, dunque, si dica chiaramente che, quantomeno negli ordinamenti europei, non sia ancora una realtà diffusa la giustizia algoritmica nel senso ristretto che qui si è delineato, e cioè la presenza di applicazioni algoritmiche impiegate direttamente dal potere giudiziario in sede processuale, la Carta ha dato avvio a una serie di ragionamenti che, per quanto degni di nota, hanno spesso cercato di immaginare un futuro ancora incerto guidandolo attraverso le categorie costituzionali⁶⁷. Non solo, a segnare la mancanza di coordinate precise su cui tracciare la rotta della riflessione giuridica ha contribuito la mancanza, fino a pochissimo tempo fa, di un dato normativo vincolante che fosse *specificamente* indirizzato all'impiego dell'intelligenza artificiale nel giudiziario. Oggi, come detto, l'Unione europea si è dotata di un Regolamento sull'intelligenza artificiale che affronta anche la tematica trattata nel presente lavoro e consente quindi di rideterminare, quantomeno in parte, il perimetro del dibattito.

Quel che è certo è che sino a ora, in Europa, la giustizia algoritmica sta assumendo la forma di esperienze locali e progetti in corso non ancora del tutto implementati in sede giurisdizionale. Inoltre, i progetti di giustizia algoritmica maggiormente studiati e in fase più avanzata sono esperimenti di cd. giustizia *predittiva* (si v. *infra*, cap. II), termine impiegato talvolta come sinonimo di giustizia algoritmica ma al quale si può fornire una definizione maggiormente circoscritta. Senza anticipare il tema, che verrà ampiamente affrontato, per come la intendiamo e la intende la dottrina maggioritaria, e per come si sta muovendo la sperimentazione in Europa, la giustizia predittiva è quel settore della giustizia algoritmica che intende progettare sistemi algoritmici per *prevedere*⁶⁸ l'esito delle controversie giudiziarie e, in generale, le decisioni dei giudici. I progetti in corso di giustizia predittiva sono presenti sia nel settore commerciale e degli studi legali, sia tuttavia – e per questi sarà necessario un approfondimento – in alcune giurisdizioni, anche in Italia.

⁶⁷ Sicuramente precursori sono i contributi usciti nel numero *Questione Giustizia, Una giustizia (im)prevedibile*, 2018.

⁶⁸ Sull'ambiguità del termine si v. *infra* cap. II.

Ma le esperienze di giustizia predittiva non esauriscono il fenomeno della giustizia algoritmica e, soprattutto, non è possibile confinare l'analisi esclusivamente alle evoluzioni in atto in Europa, considerando che il fenomeno della giustizia algoritmica, come ovviamente in generale tutta la materia che giace sotto il grande ombrello dell'intelligenza artificiale, rifiuta ontologicamente l'inserimento in rigidi confini domestici, configurandosi piuttosto come fenomeno globale⁶⁹. Soprattutto oggi con la messa sul mercato di Chat GPT (e strumenti simili) nel 2022, sono sorti dei casi giudiziari meritevoli di attenzione sull'impiego di tale strumento da parte dei giudici (si v. *infra*, cap. III).

La difficoltà riscontrata di sintetizzare le diverse strade ad oggi percorse di giustizia algoritmica ha condotto a due diverse – e tuttavia connesse – impostazioni metodologiche: la prima è stata quella di adottare un metodo induttivo – o *bottom-up* – e cioè partire dalle reali esperienze di giustizia algoritmica per poi tentare di estrapolare alcuni caratteri generali per sostenere il ragionamento giuridico; la seconda è stata quella di esplorare un ordinamento straniero, quello statunitense, in cui vi è una lunga esperienza di giustizia algoritmica nel settore penale, soprattutto nella forma di strumenti in grado di predire (*recte* prevedere) la probabilità di recidiva degli imputati.

4.2. Il metodo induttivo e i suoi limiti

Per le ragioni spiegate nel paragrafo precedente, un approccio induttivo ci appare l'unico in grado di adeguatamente selezionare le esperienze di giustizia algoritmica più rilevanti e tentare di cucirci su un ragionamento giuridico solido⁷⁰. Approccio induttivo non significa soltanto partire dalle esperienze in gioco oggi e costruirci sopra il contorno costituzionale, ma ci impegna a capire per ogni fenomeno di

⁶⁹ C. CASONATO, *L'intelligenza artificiale e il diritto pubblico comparato ed europeo*, in *DPCE Online*, 1/2022. Si v. anche A. D'ALOIA, *Ripensare al diritto al tempo dell'intelligenza artificiale*, in G. CERRINA FERONI - C. FONTANA - E. C. RAFFIOTTA, *AI Anthology. Profili giuridici, economici e sociali dell'intelligenza artificiale*, Il Mulino, 2022, p. 110.

⁷⁰ Condivide tale approccio S. PENASA, *Intelligenza artificiale e giustizia: il delicato equilibrio tra affidabilità tecnologica e sostenibilità costituzionale in prospettiva comparata*, in *DPCE Online*, 1/2022, p. 299.

giustizia algoritmica quali siano opportunità e rischi per il presente e il futuro della giustizia.

Spesso infatti, come già accennato, si sottolineano vantaggi e svantaggi di un'idea astratta di giustizia algoritmica, sintetizzata con la formula "l'intelligenza artificiale nel processo", o concetti simili. Se tuttavia tale formula non viene riempita di contenuto, essa rimane un involucro vuoto riempito di materiale gassoso e dispersivo e non, come dovrebbe, di solidi argomenti.

Ora, vale la pena qui di mettere comunque le mani avanti. L'approccio impiegato nel lavoro sconta un'eliminabile criticità, che va di pari passo con la costante evoluzione tecnologica. Proprio per la ragione che non si vuole qui prospettare un futuro ancora inesistente, già mentre si leggerà questa introduzione e poi – si auspica – il resto del lavoro, è possibile che già alcune delle riflessioni svolte risultino obsolete, proprio perché la realtà della giustizia algoritmica, come in generale quella dell'intelligenza artificiale, sembra avanzare troppo velocemente per cristallizzarla in uno scritto.

Questo porta a due diverse considerazioni: la prima è la volontà e la speranza che le riflessioni costituzionalistiche improntate sullo studio delle esperienze attuali di giustizia algoritmica possano essere sufficientemente solide da reggere nel tempo, anche quando il processo prenderà una forma ancora diversa da quella che sta assumendo oggi. In tal senso, anche quando il sistema processuale sarà ancora di più contaminato da dinamiche automatizzate, l'auspicio è che il quadro costituzionale disegnato nel presente lavoro possa in ogni caso servire come base per eventuali ragionamenti futuri, plasmati sulle nuove esperienze di giustizia algoritmica.

L'altra considerazione, legata a quanto già detto, ribalta ma solo in parte la prospettiva fino a qui indicata: la ricerca svolta ci ha costretto, per il fascino che la caratterizza e per le novità che presenta, a immaginarci anche i possibili futuri sviluppi della giustizia algoritmica, e sarebbe insensato non presentarli nel lavoro. E non sono scenari immaginifici ma previsioni fondate su quello che oggi sta accadendo, sia sul piano dell'intelligenza artificiale sia sull'evoluzione del diritto e della società. Con ciò si intende dunque avvisare che sebbene non saranno parte

centrale della ricerca, tali previsioni di tanto in tanto verranno esplicitate senza la pretesa, ma piuttosto con il sentore, che esse verranno effettivamente implementate nei panorami giudiziari.

4.3. Il metodo comparativo

Come già segnalato, l'indagine muoverà da una prospettiva di diritto costituzionale italiano, nell'accezione integrata e multilivello che si è anticipata *supra*. Coerentemente con tale prospettiva, per le caratteristiche del tema trattato – che come visto rifiuta la dimensione domestica – appare non solo utile, ma financo necessario adottare un metodo comparativo⁷¹.

In un fenomeno globale come quello della giustizia algoritmica appare ancora più importante indagare le soluzioni – normative, giurisprudenziali e accademiche – adottate in altri stati che condividono oggi tradizioni culturali e giuridiche simili.

4.3.1. (segue) Comparazione con le esperienze europee

La prospettiva multilivello ci obbliga già a condividere un approccio comparatistico, all'interno degli ordinamenti europei, sia nell'ambito dell'Unione sia nell'ambito del sistema del Consiglio d'Europa⁷².

L'indagine sulla giustizia algoritmica non può quindi sottrarsi *in primis* a guardare le soluzioni adottate dalle stesse istituzioni europee e dalle giurisprudenze delle Corti, in particolare la Corte europea dei diritti e la Corte di giustizia dell'Unione Europea. Soprattutto per un'indagine sulle problematiche afferenti al giusto processo, deve essere prioritario impiegare lo sguardo della Corte europea dei diritti sul rispetto dell'art. 6 CEDU, come meglio si chiarirà.

⁷¹ Approccio condiviso dalla scuola trentina, Si v. C. CASONATO, *L'intelligenza artificiale e il diritto pubblico comparato ed europeo*, in *DPCE Online*, 1/2022.

⁷² Del resto, la ricostruzione e la definizione delle garanzie costituzionali, nella prospettiva multilivello, passano anche attraverso la comparazione degli ordinamenti nazionali, si v. ancora B. RANDAZZO, *Imparzialità del giudice e fiducia nella magistratura nella prospettiva costituzionale europea*, cit.

Oltre alla dimensione europea, sarà necessario altresì vedere in alcuni ordinamenti europei lo stato della giustizia algoritmica in tutti i suoi formanti, gli esperimenti in corso, le pronunce dei giudici e le analisi dottrinali. Ciò risulta fondamentale considerata la stretta interrelazione che esiste tra gli Stati europei, anche per ciò che è e che sarà la giustizia algoritmica: gli Stati membri dell'Unione europea dovranno tutti rispettare le regole comuni di recente inserite nel nuovo Regolamento sull'intelligenza artificiale e saranno soggetti alla interpretazione che del Regolamento verrà data dalla Corte di Giustizia dell'Unione europea; gli Stati membri del Consiglio d'Europa sono soggetti alla giurisdizione della Corte europea dei diritti e – per chi sceglierà di ratificarlo – saranno vincolati al nuovo trattato sull'intelligenza artificiale di recente approvato.

Per tali ragioni, indagare le soluzioni adottate da alcuni Stati europei ci consentirà di trapiantare, con cautela, tali soluzioni anche nel nostro ordinamento, e di riflettere sulle opportunità e i rischi che oggi offre la giustizia algoritmica.

Ciò detto, come già specificato la giustizia algoritmica in Europa non presenta ancora esperienze davvero significative, costringendo l'indagine ad affacciarsi oltre i confini continentali e adottare il metodo comparativo orizzontale con un ordinamento che oggi presenta indubbiamente il punto più avanzato di giustizia algoritmica.

4.3.2. (segue) L'esperienza di giustizia algoritmica statunitense come benchmark dell'indagine

Non è un mistero che l'Europa non sia ad oggi l'attrice protagonista dell'industria dell'intelligenza artificiale: gli investimenti più importanti su tale tecnologia sono concentrati principalmente negli Stati Uniti e in Cina⁷³, con impianti culturali ovviamente diversi.

L'Europa ha tuttavia già qualcosa in più che è, come visto, la prima normativa organica sull'intelligenza artificiale. L'intervento legislativo, sebbene visto da

⁷³ Si v. ancora il *Report* del 2024 dell'Università di Stanford, *supra* nota 19.

alcuni come inadeguato al contesto ancora precario dell'industria dell'IA in Europa (si v. *infra* cap. V), ha comunque di recente spostato l'attenzione internazionale verso il nostro continente, interrogandosi sull'opportunità della regolazione dell'IA e sulla possibilità che si realizzi il cd. *Bruxelles effect*,⁷⁴ ossia l'affermazione fuori dai confini europei della *leadership* normativa dell'Unione, già avvenuta con l'approvazione del regolamento sulla protezione dei dati personali (GDPR)⁷⁵.

Ciò detto, il contesto europeo, sebbene stia inseguendo più che guidando l'evoluzione tecnologica e l'intelligenza artificiale, rimane il campo in cui si deve giocare la partita della giustizia algoritmica, per guidare con gli strumenti del diritto costituzionale quello che sarà il processo giurisdizionale integrato da algoritmi.

Se questa è la premessa, il lavoro di ricerca ha dovuto fare i conti con un limite insuperabile, probabilmente già intuito nelle pagine di questo capitolo introduttivo. Come visto, infatti, non essendo le tecnologie dell'intelligenza artificiale al centro dello sviluppo industriale dell'Europa, anche la giustizia algoritmica non ha avuto ancora delle manifestazioni strutturali e organiche nel panorama giudiziario europeo.

Ciò non ha impedito alla letteratura scientifica, comunque, di interrogarsi sui possibili sviluppi di tale materia e delle connessioni con le garanzie costituzionali nazionali ed europee, pur in assenza di una giurisprudenza su cui riflettere e una normativa che, fino a poco tempo fa, era disorganica e difficilmente adattabile al nuovo contesto.

Per tali ragioni nell'ambito del presente lavoro si è avvertita l'esigenza di indagare compiutamente un altro ordinamento, in cui lo sviluppo dell'intelligenza artificiale è al centro degli investimenti privati e pubblici e dove già oggi si assiste a fenomeni di giustizia algoritmica particolarmente rilevanti per il costituzionalista. Si è scelto dunque di volgere uno sguardo – anche attraverso la presenza *in loco* per

⁷⁴ A. BRADFORD, *Brussels Effect. How the European Union rules the world*, Oxford University Press, 2020.

⁷⁵ Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati).

alcuni mesi – alla giustizia algoritmica negli Stati Uniti e alle sue conseguenze sulle garanzie processuali di tale ordinamento.

Sebbene, come detto, anche la Cina presenta esperienze di giustizia algoritmica particolarmente interessanti, si è scelto l'ordinamento statunitense per diverse ragioni.

In primo luogo, la vicinanza giuridica dell'ordinamento europeo con gli Stati Uniti: benché, come noto, a differenza della maggior parte dei paesi europei, gli Stati Uniti siano un paese di *common law*, l'impianto costituzionale e le garanzie del *due process of law*, come si vedrà, sono spesso assimilabili a quelle del nostro ordinamento.

In secondo luogo l'impiego dell'intelligenza artificiale nel processo parrebbe più simile a come riteniamo possa venire impiegato nel prossimo futuro anche nel nostro continente: mentre il contesto politico cinese permette un utilizzo dell'intelligenza artificiale, anche nel processo, molto invadente e che per la nostra impostazione sarebbe quasi del tutto inaccettabile⁷⁶, la giustizia algoritmica negli Stati Uniti, ad oggi, ha assunto delle forme non così distanti da impieghi dell'IA in Europa, seppure, come visto, in altri contesti.

Inoltre, già dai primi approcci di ricerca nel panorama giudiziario statunitense è emersa una copiosa giurisprudenza, quasi mai indagata in Italia, sulla legittimità dell'impiego degli algoritmi nel sistema giudiziario penale e la compatibilità di essi con la *due process clause*. È proprio dall'analisi e dalla comprensione della giurisprudenza americana che verranno calate le problematiche sorte nell'ordinamento oltreoceano nel nostro sistema costituzionale, mettendone in luce differenze e similitudini.

Un'ultima ragione ha a che vedere con l'approccio iniziale della ricerca. L'interesse sul tema della giustizia algoritmica è sorto quando anche in Europa è giunta la notizia che un imputato nel Wisconsin – il sig. *Loomis* – era stato condannato sulla base di un algoritmo predittivo. La diffusione della notizia in termini sensazionalistici negli Stati Uniti prima, e in Europa poi, ha condotto

⁷⁶ Si cfr. T. KERIKMÄE - O. HAMULÁK - T. GÁBRIŠ, *Frontiers in AI Judiciary: A Contribution to Legal Futurology*, in *Acta Baltica Historiae et Philosophiae Scientiarum*, 11, 2, 2023.

immediatamente la letteratura a immaginarsi futuri apocalittici di giudici *robot* e decisioni automatizzate.

Solo qualche tempo dopo gli studi, anche in Italia, hanno iniziato a concentrarsi sulle reali problematiche che gli algoritmi ponevano sul piano delle garanzie costituzionali, senza quello strato di irrealtà che aveva circondato inizialmente la vicenda.

Spesso la letteratura costituzionalistica si è arrestata al commento e all'analisi di quel singolo caso giudiziario, indubbiamente interessante ma tutt'altro che isolato. Quello che è emerso nella ricerca, infatti, è che negli Stati Uniti l'impiego di algoritmi predittivi nel panorama giudiziario penale non rappresentava affatto una novità, e anzi già da molti anni algoritmi molto semplici venivano utilizzati per determinare indici di rischio di recidiva degli imputati.

Come facilmente immaginabile, ciò ha dato origine a una relevantissima giurisprudenza, chiamata a sindacare la compatibilità delle previsioni algoritmiche con l'impianto costituzionale americano. Prima di *Loomis* il giusto processo non era invocato dalle difese degli imputati, che concentravano piuttosto le contestazioni sull'abuso di discrezionalità del giudice che si verificava dinanzi a un algoritmo spesso incomprensibile e che catturava la decisione sulla determinazione della pena. Da *Loomis* in avanti si è cominciato a dubitare della legittimità delle previsioni algoritmiche anche in riferimento alle garanzie rilevanti del giusto processo, quali la *disclosure* e l'accuratezza delle prove, il diritto a un trattamento sanzionatorio individualizzato e l'obbligo di motivazione del giudice.

La comparazione con la giustizia algoritmica statunitense, soprattutto attraverso l'analisi di casi significativi nel campo del processo penale, ha permesso di saggiare la compatibilità della predizione algoritmica impiegata dai giudici con le garanzie del *due process of law*. Le soluzioni – spesso non convincenti – adottate dai giudici statunitensi consentono di riflettere quindi sulla compatibilità di siffatte predizioni con il giusto processo visto nella chiave integrata europea, che se è vero che condivide alcuni aspetti con il *due process* statunitense, spesso se ne distanzia richiedendo come vedremo garanzie più elevate.

L'incursione dell'ordinamento statunitense nel lavoro servirà quindi non a fornire risposte definitive – perché ancora non sappiamo quale sarà il peso dell'IA nella giustizia in Europa e in Italia – ma sarà utile piuttosto per darci delle prime indicazioni di ciò che si può accogliere di questa tecnologia e ciò che invece non sembra poter allinearsi alle garanzie costituzionali.

5. Obiettivi e limiti della ricerca

In ultimo, occorre chiarire quali sono le finalità e gli obiettivi della ricerca svolta, mettendo necessariamente in luce anche alcuni limiti che oggi non sembrano potersi superare se si vuole affrontare la giustizia algoritmica secondo la prospettiva costituzionalistica qui presentata.

5.1. Verso il giusto processo algoritmico

Le difficoltà di affrontare un tema così ricco di elementi anche estranei all'ambito di competenza del giurista non ci impediscono di fissare alcuni punti fermi che sembrano emergere dall'indagine, che sostanziano peraltro l'obiettivo di una ricerca di diritto costituzionale.

Non occorre possedere poteri divinatori per prevedere che la società globale andrà verso l'impiego di tecnologie sempre più avanzate per svolgere attività fino a poco tempo fa riservate all'essere umano.

In questo senso la Corte costituzionale, assumendo quasi le vesti della Pizia in un tempo in cui la trasformazione tecnologica era ancora agli albori, ha affermato che «l'avvento sempre più massiccio delle nuove tecnologie, specie nel mondo del lavoro, imporrà modifiche ed innovazioni sempre più radicali e profonde di qualifiche, di categorie professionali, di tempi di lavoro, di utilizzazione del fattore umano insieme alle macchine, ed automatismi sempre più sofisticati»⁷⁷.

⁷⁷ Corte cost., sent. n. 248 del 1986, punto 6 del *Considerato in diritto*.

Tra gli ambiti della società, certamente il diritto, inteso come ambito disciplinare, per le sue caratteristiche tipicamente umane ha già mostrato segni di resistenza all'accettazione totale della dimensione tecnologica. Ancor di più il processo giurisdizionale, tipicamente finalizzato a risolvere controversie tra esseri umani o, nell'ambito penale, a irrogare una punizione a un essere umano, fa intrinsecamente fatica ad adottare soluzioni tecnologiche e automatizzate.

Non v'è dubbio tuttavia che, negli ultimi anni, anche grazie a una digitalizzazione sempre maggiore all'interno dei sistemi giurisdizionali, la distanza in termini di impiego tecnologico tra la società nel suo complesso e il processo si è assottigliata, seppur abbia ancora una certa ampiezza⁷⁸.

Se questa è la premessa il diritto costituzionale ha un ruolo non trascurabile nel nuovo scenario tecnologico che caratterizzerà il procedimento giurisdizionale. Nello specifico tale ruolo è quello di rileggere le garanzie costituzionali del giusto processo alla luce delle evoluzioni in atto, provando a comprendere se esse potrebbero già reggere l'urto delle decisioni algoritmiche o se, invece, dovrebbero essere riadattate o reinterpretate per ricomprendere al loro interno la giustizia algoritmica. Senza anticipare considerazioni che verranno svolte nel corso dell'indagine, di seguito alcune delle principali questioni che si sono presentate durante il lavoro di ricerca.

Quanto ai diritti delle parti in giudizio, di fronte a una decisione assunta anche tramite un algoritmo predittivo, come visto spesso opaco, come può venire garantito il diritto di difesa? Se tale algoritmo, come avviene negli Stati Uniti nel processo penale, diviene una prova di colpevolezza prodotta in giudizio dalla parte pubblica, come è possibile, se è possibile, garantire il diritto alla parità delle armi e al contraddittorio? Sempre in ambito probatorio, sarebbe tollerabile l'impiego di prove algoritmiche non pienamente conoscibili? In generale, l'impiego di algoritmi da parte del giudice potrebbe essere compatibile con la *fairness* del processo giudiziario?

⁷⁸ E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit., p. 47.

Con riguardo invece alle garanzie istituzionali del giudice, di fronte ad algoritmi progettati dai privati e che, come vedremo, assumono nel processo un peso decisivo, è possibile garantire i principi di indipendenza e autonomia del potere giudiziario? Se tali algoritmi, inoltre, non sono del tutto conoscibili (o comprensibili) dallo stesso giudice che li impiega, che garanzie si dovrebbero adottare per mantenere fermo l'obbligo del giudice di motivare i provvedimenti giurisdizionali?

In ultimo e come sfida finale, la giustizia algoritmica per come oggi si è configurata e per come sembra potersi configurare in futuro, riuscirebbe ancora a garantire una tutela giurisdizionale *effettiva*?

Queste sono alcune delle domande che una ricerca di diritto costituzionale sulla giustizia algoritmica ha l'obbligo di affrontare e che sostanziano anche gli obiettivi del presente lavoro. Tali non possono che essere quelli di a) verificare la tenuta delle garanzie processuali rispetto agli strumenti di giustizia algoritmica; b) porre dei limiti oggi invalicabili all'impiego di strumenti algoritmici in situazioni in cui questi sembrano non poter considerarsi compatibili con i diritti in gioco; c) disegnare in ultimo il *giusto processo algoritmico*.

Sono obiettivi ambiziosi, ma che oggi – come già indicato – trovano conforto nella legislazione positiva e, in particolare, nel Regolamento sull'intelligenza artificiale che non si è sottratto, seppur in modo non sempre convincente, al perseguimento di tali obiettivi. Il ruolo della dottrina è quindi quello di partire dal dato positivo e dall'approfondimento della giurisprudenza rilevante per tentare di vedere quali sono ad oggi le opportunità e i vantaggi della giustizia algoritmica e quali, per altro verso, i coni d'ombra, tenendo conto che la prospettiva temporalmente più vicina oggi appare dunque quella di un incontro tra *humanitas* e *techne*⁷⁹ che favorisca una graduale ottimizzazione dell'*HCI* (*Human-Computer Interaction*)⁸⁰ o meglio, nel lavoro qui presentato, *JCI* (*Judge-Computer Interaction*).

⁷⁹ E. VINCENTI, Il «problema» del giudice-robot, in A. CARLEO, *Decisione robotica*, Il Mulino, Bologna, 2019, p. 123.

⁸⁰ L. FLORIDI, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, cit., p. 10.

Gli obiettivi del lavoro devono tuttavia fare i conti con dei limiti intrinseci di una ricerca sul tema qui preso in esame (si v. par. successivo), i quali se non devono scoraggiare lo studioso dall'affrontarne le problematiche, sicuramente lo obbligano ad assumere posizioni caute e provvisorie.

5.2. Pluralità di applicazioni, evoluzione tecnologica e quesiti aperti

La lettura del presente capitolo introduttivo probabilmente ha già fatto emergere alcuni dei limiti da considerare nella ricerca, che proviamo brevemente a sistematizzare in questo paragrafo.

Innanzitutto è utile comprendere la geometria circolare che ha portato alla metodologia adottata nel lavoro e la necessità di guardare in modo approfondito al sistema statunitense. Al riguardo, come detto lo scoglio principale era quello di trovare un contesto in cui il diritto ha dovuto *effettivamente* porsi il problema della compatibilità costituzionale della giustizia algoritmica.

Considerate le esperienze ancora embrionali nel nostro ordinamento e in quelli a noi vicini, la scelta è ricaduta in un contesto in cui da tempo si fa esperienza di giustizia algoritmica. Lo studio del fenomeno statunitense ci ha permesso quindi di ritornare nel contesto ordinamentale europeo con un bagaglio giuridico importante e ci ha consentito di conservare alcuni insegnamenti che, data la vicinanza di tante garanzie del giusto processo, sono adattabili al contesto europeo e nazionale. Come si vedrà tali insegnamenti sono estendibili a dinamiche processuali non strettamente connesse all'ambito penale in cui negli Stati Uniti si è sviluppata la giustizia algoritmica.

Tale struttura circolare presenta però dei punti di fragilità che attualmente non sembrano colmabili. Come già detto non è possibile affrontare la giustizia algoritmica come un fenomeno unitario, poiché le forme che potrebbe assumere sono pressoché illimitate. L'esperienza americana ci consegna una singola sfumatura della giustizia algoritmica, l'impiego di strumenti di IA per prevedere l'indice di recidiva degli imputati nel processo penale, ma le conclusioni raggiunte in tale ambito sono difficilmente applicabili a esperienze diverse di giustizia

algoritmica (come la giustizia predittiva o l'impiego di assistenti come Chat GPT per la scrittura delle sentenze). Quindi, mentre l'esperienza americana ci consegna numerosi elementi su cui riflettere, le esperienze europee sono ancora in una fase in cui è complicato definirne aspetti problematici e opportunità. Ciò non toglie peraltro che molti elementi di criticità emersi nel contesto statunitense – tra i tanti la progettazione e i dati dei sistemi algoritmici, l'opacità algoritmica, l'affidabilità dei sistemi e la loro decisività nei procedimenti decisionali – possano considerarsi comuni a tante forme di giustizia algoritmica, e su questi il presente lavoro si concentrerà.

Il primo limite consiste dunque nell'aver a disposizione pochi elementi direttamente afferenti al tema qui trattato che possano segnare un percorso da seguire per lo studioso. Peraltro, come detto, questo parziale vuoto è stato di recente colmato con l'approvazione del Regolamento, la cui analisi è comunque fondamentale per comprendere la direzione che assumerà la giustizia algoritmica negli ordinamenti europei.

Un secondo limite, già brevemente segnalato, ha a che vedere con le caratteristiche della giustizia algoritmica, le cui applicazioni dipendono massicciamente da come si evolverà la tecnologia. Se fino a poco tempo fa strumenti come Chat GPT non potevano nemmeno essere concepiti, oggi ci sembrano strumenti “normali” con cui interagire. Se fino a poco tempo fa il processo si articolava in fascicoli, oralità, applicazioni analogiche, oggi ci appare del tutto “normale” il processo telematico. Se le decisioni della pubblica amministrazione si basavano esclusivamente su scelte assunte da esseri umani, oggi consideriamo – coi limiti che vedremo – “normale” assumere decisioni sulla base di procedure algoritmiche. Tutto ciò per dire questo: ci troviamo davanti a una grande incertezza conoscitiva sulle forme che acquisirà l'impiego di algoritmi nel processo, abbiamo solo pochi esempi e su questi abbiamo l'obbligo di costruire un impianto teorico e pratico il più possibile resistente al tempo che passa, secondo un principio che è stato definito di “precauzione costituzionale”⁸¹, che sembra essere

⁸¹ A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *Biolaw Journal – Rivista di BioDiritto*, 1/2019, pp. 86-87.

stato accolto anche dal nuovo Regolamento sull'IA. Può darsi che questo impianto reggerà per un periodo lungo, perché può darsi che gli algoritmi e l'intelligenza artificiale faranno più fatica ad addentrarsi nella dinamica processuale di quanta ne stiano facendo per influenzare il resto della società. Può, d'altro canto, accadere che quanto emerso dalla ricerca abbia in poco tempo la necessità di venire aggiornato, con la speranza tuttavia che quanto qui discusso possa comunque rappresentare un punto di partenza degno di considerazione per le riflessioni future.

Insomma, sintetizzando obiettivi e limiti della ricerca, la domanda da porci è se il diritto costituzionale avrà la forza di reggere l'urto della cd. «quarta rivoluzione»⁸² e del «cambio di paradigma» che impone l'intelligenza artificiale⁸³; di questi cambiamenti in atto non abbiamo tuttavia piena contezza e ancora non siamo in grado di conoscerne le reali e future implicazioni. Come si dice, ai posteri l'ardua sentenza.

6. Struttura del lavoro

Brevemente, il lavoro si articolerà in quattro parti distinte. Nella prima parte si darà conto dello stato dell'arte della giustizia algoritmica negli Stati Uniti, considerando principalmente l'impiego dei *risk assessment tools* nel procedimento penale (cap. I), e in Europa, dando rilievo agli esperimenti di giustizia predittiva in alcune giurisdizioni europee e italiane (cap. II). L'ultimo capitolo di questa prima parte (cap. III) sarà invece dedicato alle prospettive futuribili di giustizia algoritmica, considerato l'uso recente che è stato fatto di Chat GPT da parte di alcuni giudici.

La seconda parte sarà invece finalizzata a mettere in luce gli approcci regolatori all'intelligenza artificiale e della giustizia algoritmica: dopo aver ricostruito i diversi orientamenti delle amministrazioni presidenziali statunitensi e indagato gli atti normativi oggi rilevanti negli Stati Uniti (cap. IV) l'analisi si concentrerà sul

⁸² L. FLORIDI, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, cit.

⁸³ A. SIMONCINI - S. SUWEIS, *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, in *Rivista di filosofia del diritto*, 1/2019, pp. 87-106.

percorso di regolazione dell'IA e della giustizia algoritmica in Unione europea e nell'ambito del Consiglio d'Europa con l'obiettivo di capire se i risultati conseguiti in Europa dal punto di vista regolatorio sono in grado di guidare il sentiero della giustizia algoritmica (cap. V).

La terza parte rappresenta il cuore del lavoro e intende porre in risalto i numerosi intrecci tra giustizia algoritmica e giusto processo. Dopo aver ricostruito i nodi problematici oggi più rilevanti degli algoritmi nel settore della giustizia (cap. VI) e le garanzie del *due process of law* e del *fair trial* nella prospettiva europea multilivello (cap. VII), partendo dagli insegnamenti della giurisprudenza americana si indagheranno i diversi livelli di incidenza di risultati algoritmici sulla decisione del giudice e le conseguenti ricadute sulle garanzie istituzionali e procedurali del giusto processo (cap. VIII). Infine, l'ultima parte del lavoro proverà a trarre alcune conclusioni dalla ricerca svolta, concentrandosi principalmente sulla teorizzazione di un giusto processo algoritmico (cap. IX).

PARTE I

LA GIUSTIZIA ALGORITMICA TRA STATI UNITI ED EUROPA: STATO DELL'ARTE, ESPERIENZE PRATICHE E PROSPETTIVE FUTURE

Premessa: i differenti sviluppi della giustizia algoritmica negli Stati Uniti e in Europa

Come già segnalato nel capitolo introduttivo, la giustizia algoritmica per come definita nel presente lavoro è ancora lontana dall'essere una realtà nel panorama giudiziario europeo.

Benché la discussione sia – probabilmente per la fascinazione che generano tali tematiche – particolarmente viva, in Europa siamo ancora in una fase di studio delle possibili applicazioni nei sistemi giudiziari.

Come vedremo, infatti, negli ordinamenti europei la giustizia algoritmica ha assunto le vesti di quella che è definita, non senza ambiguità, “giustizia predittiva”, che definiremo compiutamente nel capitolo II, ma che in sintesi riguarda la possibilità di predire (*recte* prevedere) la futura decisione del giudice.

In campo giudiziario si tratta, come già visto, di sperimentazioni locali o studi accademici (spesso in collaborazione tra università e singoli tribunali), mentre è noto l'utilizzo già da parte di studi legali e aziende di *legal tech* di *software* predittivi in grado di prevedere la decisione del giudice sulla base di migliaia di decisioni assunte nel medesimo caso da un determinato tribunale⁸⁴.

Quindi ad oggi in Europa non sembra ci siano organi giudiziari che, per assumere decisioni, si avvalgono di strumenti algoritmici e predittivi, e nemmeno esistono procedimenti giudiziari affidati esclusivamente a strumenti artificiali⁸⁵.

⁸⁴ Esempi sono *Lex Machina* e *PremonitionAi*.

⁸⁵ Come nota Andrea Simoncini in ID, *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*, in *Rivista di diritto processuale*, 2, 2024, per anni in realtà è circolata la notizia che in Estonia era in fase di studio un progetto che perseguiva l'obiettivo di affidare a programmi di IA la risoluzione di controversie civili di modesta entità (*small claims*). L'obiettivo era realizzare un “robo-judge” che, sulla base di atti e documenti caricati dalle

Ciò, peraltro, non significa che le esperienze attualmente in atto in Europa non siano meritevoli di attenzione o che non permettano di riflettere sulla compatibilità di strumenti predittivi con le garanzie costituzionali. E ciò è dimostrato dalla grande produzione dottrinale europea sul tema della giustizia algoritmica e, soprattutto, dall'interesse mostrato dai legislatori, come vedremo nella parte II di questo lavoro.

Come segnalato, l'esperienza europea non è tuttavia sufficiente per ragionare di giustizia algoritmica e garanzie costituzionali: occorre consegnare al dibattito giuridico un punto di vista vissuto, reale, che ha prodotto delle conseguenze sul piano ordinamentale. In questo senso l'esperienza di giustizia algoritmica maturata negli Stati Uniti è un insegnamento irrinunciabile, un campo di indagine necessario per capire come le predizioni algoritmiche impiegate nel processo decisionale mutino i paradigmi della funzione giurisdizionale e impongano serie riflessioni sul rispetto delle garanzie processuali.

A differenza dell'ordinamento europeo, negli Stati Uniti vi è una forma compiuta di giustizia algoritmica, in atto da tempo e ancora oggi sotto l'occhio attento di giudici e accademici. Non riguarda la sperimentazione di *software* per prevedere le decisioni delle Corti⁸⁶ ma l'impiego in alcune giurisdizioni statali penali di strumenti di *risk assessment*, vale a dire *software* algoritmici in grado di prevedere il rischio di recidiva degli imputati. Essi sono utilizzati per scopi differenti, ma per quanto di interesse nel presente lavoro verrà preso in considerazione il loro impiego da parte del giudice nella fase di quantificazione della pena.

Prima di addentrarci *funditus* nell'analisi di tali strumenti, occorre una breve delucidazione per evitare confusioni: anche i *risk assessment tools* sono strumenti

parti nella piattaforma elettronica, fosse in grado di assumere una decisione. Tale notizia è stata smentita da parte dello stesso Ministero della Giustizia estone, in data 16 febbraio 2022, il quale ha affermato che «As there have been a lot of questions relating the topic of AI Judge, we have to explain that the article about Estonian project of designing a 'Robot/Judge' in Wired from 25th of March 2019, is misleading. There hasn't been that kind of project or even an ambition in the Estonian public sector. Estonian Ministry of Justice does not develop AI robot judge for small claims procedure nor general court procedures to replace the human judge. We are still searching for ICT means to make court's workload, including administrative burden more bearable», si v. il seguente indirizzo: <https://www.just.ee/en/news/estonia-does-not-develop-ai-judge>.

⁸⁶ Anche se ci sono stati degli studi per prevedere l'esito delle decisioni della Corte Suprema, si cfr. D. M. KATZ - M. J. BOMMARITO - J. BLACKMAN, *A general approach for predicting the behavior of the Supreme Court of the United States*, in *Plos One*, 4/2017.

in senso lato predittivi, e rientrano a pieno nel solco della giustizia algoritmica secondo la prospettiva costituzionalistica segnalata nel capitolo introduttivo in quanto hanno un peso e un'influenza sulla decisione del giudice. Essi tuttavia non sono ricompresi nella nozione di giustizia predittiva adottata in questo lavoro, che invece come visto riguarda *software* utilizzati per prevedere le decisioni giurisprudenziali, di cui si tratterà *infra* (cap. II).

LA GIUSTIZIA ALGORITMICA NEGLI STATI UNITI: L'IMPIEGO DI STRUMENTI PREDITTIVI NEL PROCEDIMENTO PENALE

SOMMARIO: 1. La giustizia algoritmica nel processo penale statunitense - 2. Il dibattito filosofico di fondo sull'impiego dei *tools* - 3. L'evoluzione storica della valutazione del rischio di recidiva - 4. Il funzionamento dei *risk assessment tools* - 5. L'impiego dei tools nel contesto giudiziario statunitense: in particolare il *sentencing* - 6. L'acquisizione da parte del giudice del risultato algoritmico

1. La giustizia algoritmica nel processo penale statunitense

Come detto, l'esperienza più significativa e più capace di stimolare considerazioni di portata costituzionale avviene ad oggi nel contesto giudiziario statunitense, dove algoritmi in grado di prevedere il rischio di recidiva di un imputato o di un condannato vengono di norma impiegati nel processo penale latamente inteso⁸⁷.

Al riguardo, negli ultimi dieci anni gli Stati Uniti sono stati i più partecipi testimoni di un'esplosione nell'utilizzo di algoritmi nella sfera pubblica⁸⁸, e il processo penale non è rimasto escluso da tale fenomeno. Ciò è stato possibile per il verificarsi di diversi fattori: la disponibilità da parte dell'autorità pubblica di una mole di dati – attraverso soprattutto l'uso diffuso di piattaforme digitali e internet – che prima non era nemmeno immaginabile; la progettazione di sistemi algoritmici

⁸⁷ La letteratura americana sull'impiego dei *risk assessment tools* è vastissima e verrà presa in considerazione singolarmente nel corso del lavoro. Qui, in generale sul tema, si cfr., tra gli altri, C. SLOBOGIN, *Principles of Risk Assessment: Sentencing and Policing*, in *Ohio State Journal of Criminal Law*, vol. 15, 2018, pp. 583- 596; J. VILLASENOR - V. FOGGO, *Artificial intelligence, due process, and criminal sentencing*, in *Mich. St. L. Rev.*, 2020, pp. 295- 354; R. BERK *et al.*, *Fairness in Criminal Justice Risk Assessments: The State of the Art*, in *Sociological Methods & Research*, 50 (1), pp. 3-44.

⁸⁸ Per un panoramica si cfr. R. CALO - D.K. CITRON, *The Automated Administrative State: A Crisis of Legitimacy*, in *Emory Law Journal*, 70, 4, 2021, pp. 799-845.

sempre più efficienti e sofisticati da parte soprattutto di grandi società tecnologiche; la maggiore familiarità diffusa riguardo a decisioni assunte da sistemi artificiali⁸⁹.

Non è un caso che il *report* di fine anno di dicembre 2023 sul lavoro delle Corti federali – redatto dal giudice John G. Roberts, Jr., *Chief Justice* della Corte Suprema degli Stati Uniti – abbia come assoluta protagonista una riflessione sulla rivoluzione tecnologica e sull'avvento dell'IA⁹⁰.

Il *report* è estremamente interessante sia perché ci mostra la prospettiva dell'organo al vertice della giustizia statunitense, sia perché offre degli spunti indicativi rispetto al tema che il presente lavoro sta tentando di chiarire, ossia lo spazio che l'intelligenza artificiale e gli algoritmi hanno acquisito e potrebbero acquisire in futuro nel giudiziario. Scrive il Giudice che «*in criminal cases, the use of AI in assessing flight risk, recidivism, and other largely discretionary decisions that involve predictions has generated concerns about due process, reliability, and potential bias. At least at present, studies show a persistent public perception of a “human-AI fairness gap,” reflecting the view that human adjudications, for all of their flaws, are fairer than whatever the machine spits out*».

Già solo queste parole introducono uno dei temi che il presente capitolo intende affrontare: l'impiego dell'intelligenza artificiale nel processo penale ha fatto sorgere forti dubbi sul rispetto delle garanzie del *due process of law* e sulla tenuta di siffatte garanzie di fronte all'impiego di strumenti algoritmici predittivi. Un obiettivo sarà dunque quello di indagare la coerenza delle decisioni dei giudici che si sono trovati a giudicare sul rispetto delle garanzie processuali delle parti soggette a previsioni algoritmiche, cercando di individuare – anche con l'aiuto della dottrina statunitense – le contraddizioni e le incongruenze che inevitabilmente circondano decisioni che per la prima volta si rivolgono all'interpretazione di “fatti” del tutto nuovi, con proprie complessità – soprattutto tecniche – che non possono venire trascurate e che per questo saranno oggetto anch'esse di un approfondimento.

⁸⁹ D. KEHL - P. GUO - S. KESSLER, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing*, in *Responsive Communities Initiative*, Berkman Klein Center for Internet & Society, Harvard Law School, 2017, p. 3

⁹⁰ 2023 *Year-End Report on the Federal Judiciary*, disponibile al seguente indirizzo: <https://www.supremecourt.gov/publicinfo/year-end/2023year-endreport.pdf>.

Anche il secondo aspetto che il presente lavoro intende indagare trova un primo appiglio nelle parole del giudice Roberts nel *report*. In esso si mantiene salda la prospettiva antropocentrica dell'applicazione del diritto, nella quale i giudici «*perform quintessentialy human functions*»: mentre i giudici di prime cure infatti sono chiamati a pronunciare la sentenza anche sulla base di elementi e informazioni difficilmente elaborabili da una macchina, i giudici dei gradi superiori hanno il compito di verificare se le corti inferiori hanno esercitato in modo abusivo il proprio potere discrezionale, uno *standard* di verifica che per sua natura «*involves fact-specific gray areas*», anch'esse difficilmente categorizzabili dall'intelligenza artificiale. La seconda direttrice che verrà seguita riguarda dunque il modo in cui la discrezionalità giudiziaria viene plasmata e influenzata dall'impiego di algoritmi predittivi: anche in questo caso l'indagine attingerà dalle decisioni delle corti statunitensi con l'obiettivo ambizioso di capire se oggi il giudice abbia tutti gli strumenti adatti per governare un fenomeno di tale portata o se, al contrario, il suo ruolo venga parzialmente minacciato dalla predizione algoritmica.

Le conclusioni del *report*, infine, riflettono una personale convinzione del giudice redattore, peraltro ad oggi pienamente condivisibile: «*I predict that human judges will be around for a while. But with equal confidence I predict that judicial work—particularly at the trial level—will be significantly affected by AI. Those changes will involve not only how judges go about doing their job, but also how they understand the role that AI plays in the cases that come before them*».

Proprio perché non v'è dubbio che l'intelligenza artificiale provocherà cambiamenti epocali sul modo di esercitare la giurisdizione, siamo noi in qualità di giuristi a non dover farci sopravanzare dal cambiamento in atto ma a tentare di accompagnarlo man mano che esso si farà sempre più ingombrante, negli Stati Uniti come anche in Europa e in Italia. Se oggi, come si chiarirà, le decisioni dei giudici statunitensi spesso non convincono e appaiono in contraddizione tra loro – e ciò è comprensibile perché, come detto, si tratta di un fenomeno senza precedenti e con delle complessità sue proprie – lo sforzo della dottrina costituzionalistica deve riporsi in quello che è il suo ruolo per natura, quello di tentare di orientare la futura

giurisprudenza verso un'applicazione del diritto che sia più coerente con le garanzie costituzionali.

Ciò premesso, nei successivi paragrafi si tenterà di analizzare compiutamente il fenomeno dell'impiego di *risk assessment tools* nel *sentencing* per preparare il terreno alla successiva analisi della giurisprudenza americana che esploreremo nella terza parte del presente lavoro.

2. Il dibattito filosofico di fondo sull'impiego dei *tools*

L'impiego della valutazione del rischio di recidiva ha assunto un ruolo sempre più prominente all'interno del sistema penale statunitense ed è anche il frutto di un lungo e complesso dibattito filosofico all'interno degli Stati Uniti. Prima degli anni '70 la valutazione individuale della pericolosità del condannato era un elemento centrale per la sentenza, la quale doveva tenere in considerazione non solo la gravità del reato ma anche le caratteristiche dell'imputato, e avere quindi come fine la riabilitazione dello stesso⁹¹. Ciò comportava un significativo ampliamento della discrezionalità giudiziaria nella scelta della sentenza da applicare⁹².

Le preoccupazioni legate all'ampliamento della discrezionalità giudiziaria e il prevalere di dottrine giudiziarie più propense a valorizzare il fine retributivo della pena anziché la riabilitazione del condannato condussero a elidere gran parte della valutazione sul comportamento successivo dello stesso e a mitigare il potere giudiziario del giudice⁹³.

Tale trasformazione fu favorita anche dall'uscita di un saggio del giudice Marvin Frankel che criticava aspramente l'eccessiva discrezionalità dei giudici americani e che ebbe una notevole influenza sulle dottrine giudiziarie dell'epoca⁹⁴. Secondo l'Autore, il potere arbitrario del giudice nel *sentencing* doveva essere fortemente limitato e le decisioni dei giudici, per poter essere considerate legittime, avrebbero

⁹¹ B. L. GARRETT - J. MONAHAN, *Judging Risk*, in *California Law Review*, vol. 108, 2020, p. 442.

⁹² D. KEHL - P. GUO - S. KESSLER, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing*, cit., p. 6.

⁹³ *Ibidem*.

⁹⁴ M. FRANKEL, *Criminal Sentences: Law Without Order*, Hill and Wang, New York, 1973.

dovuto basarsi su «criteri oggettivi e oggettivamente verificabili»⁹⁵. In particolare, proponeva financo la creazione di una tabella o un elenco di fattori che contenessero «ove possibile, una forma di valutazione numerica o comunque oggettiva» e una commissione che analizzasse le sentenze e formulasse delle linee guida standardizzate per ricondurre la discrezionalità giudiziaria su binari più prevedibili⁹⁶.

Il saggio di Frankel diede avvio a una serie di studi su larga scala relativi ai fattori che inevitabilmente influenzano la decisione del giudice, fattori che un recente studio definisce con il nome di “*Noise*” (Rumore)⁹⁷, e che comprendono, tra gli altri, il momento della giornata in cui il giudice assume la propria decisione; la vittoria o la sconfitta della squadra di *football* locale preferita; il giorno del compleanno del singolo condannato o dello stesso giudice⁹⁸.

Le sollecitazioni di Frankel e il nuovo clima culturale creatosi portarono all’approvazione, nel 1984, del *Sentencing Reform Act*. Tale riforma istituì la *US Sentencing Commission*⁹⁹, con il precipuo compito di approvare linee guida vincolanti per i giudici nel procedimento di *sentencing*.

L’approvazione di linee guida vincolanti per il giudice e il venire meno di qualunque considerazione della persona del condannato (la sua pericolosità sociale, la sua idoneità a ricommettere un reato, la sua condizione fisica e mentale) a favore di una visione prettamente retributiva della pena comportarono l’inizio – e la fine appare ancora lontana¹⁰⁰ – dell’incarcerazione di massa (*mass incarceration*) e dell’inasprimento della punizione penale¹⁰¹. Per mitigare gli effetti di questi

⁹⁵ *Ivi*, p. 11.

⁹⁶ *Ivi*, p. 114.

⁹⁷ D. KAHNEMAN - O. SIBONY - C. R. SUNSTEIN, *Noise: A Flaw in Human Judgement*, trad. it. *Rumore: un difetto del ragionamento umano*, Milano, 2021.

⁹⁸ *Ivi*, p. 23.

⁹⁹ Cfr. *An Overview of the United States Sentencing Commission, United States Sentencing Commission*, 2011, disponibile al seguente indirizzo: https://www.ussc.gov/sites/default/files/pdf/about/overview/USSC_Overview.pdf.

¹⁰⁰ Gli ultimi dati mostrano che gli Stati Uniti sono la nazione che nel mondo ha la percentuale più alta di persone private della libertà personale, si v. <https://www.prisonpolicy.org/reports/pie2023.html>. Un report del 2023 certifica che circa cinque milioni di persone sono in qualche modo all’interno del circuito penale statunitense mentre circa due milioni sono dentro al carcere, si v. <https://www.sentencingproject.org/reports/mass-incarceration-trends/>.

¹⁰¹ D. KEHL - P. GUO - S. KESSLER, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing*, cit., p. 7.

fenomeni venne revocato il carattere vincolante delle linee guida dalla Corte Suprema nel 2005¹⁰² e ricominciò a prendere piede nella narrativa giudiziaria l'importanza di valutare le circostanze individuali di ciascun soggetto che entra nel circuito penale¹⁰³. Tale nuovo paradigma fu anche favorito dallo sviluppo dei primi strumenti digitali di valutazione del rischio, i *risk assessment tools*.

3. L'evoluzione storica della valutazione del rischio di recidiva

La valutazione del rischio di recidiva ha vissuto diverse fasi e la dottrina è concorde nel riconoscere almeno quattro generazioni di strumenti di valutazione¹⁰⁴.

La prima generazione di strumenti di valutazione del rischio di recidiva, apparsa negli Stati Uniti nella prima metà del secolo scorso¹⁰⁵, non contemplava sistemi artificiali ma coinvolgeva esperti clinici che si affidavano a interviste e a revisioni di documenti per raccogliere informazioni rilevanti per valutare il rischio commissione del reato. Queste valutazioni si basavano principalmente sull'esperienza e sulla conoscenza specialistica dell'esperto, piuttosto che su una raccolta di dati strutturata e standardizzata.

A partire dagli anni '70 vennero introdotti metodi artificiali per "misurare" il rischio, con l'utilizzo di macchine in grado di analizzare una serie di fattori sulla cui base produrre la propria analisi di rischio. La seconda generazione di *tools* comprende, quindi, strumenti meccanici che si basano su fattori statici (es. il sesso, l'età al momento dell'arresto, i precedenti penali: sono i cd. *historical factors*) che tuttavia risultarono presto inadatti in quanto consideravano solo fattori non

¹⁰² *United States v. Booker*, 543 U.S. 220 (2005).

¹⁰³ C.S. YANG, *Have Inter-Judge Sentencing Disparities Increased in an Advisory Guidelines Regime? Evidence from Booker*, in *SSRN Electronic Journal*, 2013, p. 1333.

¹⁰⁴ Si cfr. diffusamente D. KEHL - P. GUO - S. KESSLER, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing*, cit., pp. 9 ss. e M. HAMILTON, *Risk and Needs Assessment: Constitutional and Ethical Challenges*, in *American Criminal Law Review*, vol. 52, 2015, pp. 236 ss.

¹⁰⁵ D. KEHL - P. GUO - S. KESSLER, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessment in Sentencing*, cit., p. 8.

“controllabili” dal soggetto sottoposto all’analisi¹⁰⁶. Per ovviare al problema iniziarono a venire implementati nelle macchine anche fattori dinamici (terza generazione¹⁰⁷) quali l’impiego lavorativo, dipendenze da sostanze, l’età vista come fattore che muta nel tempo e così via (sono i cd. *criminogenic factors*), che permettevano di focalizzare l’attenzione non solo sul versante retributivo ma anche su quello rieducativo del reo.

I *tools* di quarta generazione, invece, utilizzano le tecniche dei precedenti sistemi (quindi l’analisi di fattori sia statici sia dinamici) ma migliorano nell’*accuracy* adottando, tramite procedure algoritmiche, un approccio più completo rispetto ai singoli fattori di rischio che, caso per caso, vengono in rilievo per misurare il rischio di recidiva di un singolo soggetto¹⁰⁸.

Se alcuni fanno rientrare i sistemi ad oggi più evoluti – come COMPAS (si v. *infra*) nella quarta generazione¹⁰⁹, altri propongono di considerare un’ultima e distinta fase di evoluzione¹¹⁰ – attualmente in corso – nella quale i *tools* cominciano a essere implementati attraverso tecniche di intelligenza artificiale, e in particolare di *machine learning*: detti sistemi sollevano inevitabilmente, quanto al loro impiego da parte dei giudici, interrogativi autonomi e specifici dei quali si dirà nel prosieguo.

4. Il funzionamento dei *risk assessment tools*

La presenza di decine di strumenti di *risk assessment* impiegati in ambito giudiziario nelle giurisdizioni statali degli Stati Uniti non consente di descrivere analiticamente il funzionamento di ognuno di essi, compito che peraltro sarebbe anche eccentrico rispetto alle finalità della ricerca.

¹⁰⁶ Un esempio di *tool* di seconda generazione è Static-99, il più utilizzato per il *sexual recidivism*, che contiene dieci fattori statici, cinque dei quali ricollegabili alla storia criminale del reo, si cfr. M. HAMILTON, *Risk-needs assessment: consituional and ethical challenges*, cit., p. 237.

¹⁰⁷ Esempi di *tools* di terza generazione sono HCR-20, lo strumento più utilizzato per la previsione della recidiva violenta, e Level of Service Inventory-Revised (LSI-R), che incontreremo nel corso del lavoro; si cfr. ancora, M. HAMILTON, *Risk-needs assessment: consituional and ethical challenges*, cit., p. 238.

¹⁰⁸ D. KEHL - P. GUO - S. KESSLER, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessment in Sentencing*, cit., p. 9.

¹⁰⁹ M. HAMILTON, *Risk-needs assessment: consituional and ethical challenges*, cit., p. 239.

¹¹⁰ B. L. GARRETT - J. MONAHAN, J. , *Judging Risk*, cit. p. 451.

Ciò che invece può risultare utile è tentare di comprendere – perlomeno a grandi linee – il funzionamento degli strumenti ad oggi più evoluti, sia perché il passare del tempo saranno tali strumenti (della quarta o della quinta generazione, a seconda della classificazione che si accoglie) a venire impiegati in misura prevalente sia in quanto sono gli strumenti che rispetto alle finalità della presente indagine presentano maggiori criticità e spunti di riflessione per ragionare di problematiche giuridiche quali il rispetto del diritto al giusto processo o il perimetro della discrezionalità giudiziaria¹¹¹.

Uno dei *risk tools* più evoluti, e indubbiamente il più studiato¹¹², è COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), strumento prodotto dalla società Equivant (già Northpointe) e utilizzato in diverse giurisdizioni statali, ad esempio Florida Wisconsin, Michigan, New York e New Mexico¹¹³.

Per capirne il funzionamento occorre muovere dalla guida realizzata dalla società Equivant, produttrice del *software*¹¹⁴: essa non ha una valenza meramente descrittiva e la sua importanza pratica si evince anche dal fatto che almeno una volta è stata utilizzata da un giudice per respingere le doglianze di un imputato che lamentava la violazione del *due process* per mancanza di trasparenza¹¹⁵.

Una – seppure embrionale – comprensione del funzionamento di COMPAS ha anche lo scopo di rideterminare il perimetro del discorso sull'intelligenza artificiale nel processo penale, che molto spesso ha travalicato i confini della realtà sfiorando discorsi avveniristici lontani da ciò che la dinamica processuale americana offre.

Occorre subito rilevare che non vi è alcuna indicazione sul personale che ha progettato il *software*, e dunque non è possibile ricavare il *background* di competenze alla base della progettazione di COMPAS.

¹¹¹ Un'analisi del funzionamento di COMPAS come caso di studio si ritrova in K. FREEMAN, *Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights*, cit.

¹¹² K. B. FORREST, *When machines can be judge, jury, and executioner: justice in the age of artificial intelligence*, World Scientific Publishing Co. Pte. Ltd., Singapore, 2021, p. 65.

¹¹³ M. BRENNER *et al*, *Constitutional Dimensions of Predictive Algorithms in Criminal Justice*, in *Harvard Civil Rights*, 55, 2020, p. 270.

¹¹⁴ La Guida COMPAS è disponibile al seguente link: <https://www.equivant.com/wp-content/uploads/Practitioners-Guide-to-COMPAS-Core-040419.pdf>.

¹¹⁵ *State v. Loomis*, 881 N.W.2d 749 (Wis 2016), §§ 54 e 69. Si v. *infra* al cap. VIII, § 3.

Inoltre, nella Guida si ammette che il sistema è frutto di un compromesso tra il tentativo di comprendere il maggior numero di fattori criminogeni e la facilità di utilizzo dello stesso. In altre parole, per garantire l'efficacia ed efficienza dello strumento si è scelto quali fattori criminogeni considerare nell'analisi e quali, invece, sacrificare¹¹⁶.

Il funzionamento del *software* COMPAS si divide in due "indici" o "scale": le scale di rischio ("*risk scales*") e le scale di bisogno ("*need scales*"). Le prime hanno uno scopo predittivo (sono cioè progettate per indicare una previsione del rischio di recidiva), le seconde invece hanno il fine di misurare le tendenze criminogene del soggetto e identificare meglio il programma di intervento¹¹⁷. Ciascuna scala COMPAS è stata costruita sulla base di una serie di fattori comportamentali e psicologici di grande rilevanza per il calcolo della recidiva¹¹⁸. I dati che l'algoritmo analizza sono autonomamente raccolti dai dipartimenti che ne fanno utilizzo, anche mediante la somministrazione al reo di un questionario¹¹⁹.

Le scale di rischio che COMPAS è in grado di analizzare sono diverse, ma ai nostri fini è sufficiente considerare in generale come queste sono costruite. Gli *input* della scala includono: il *background* criminoso (arresti precedenti, condanne passate ecc.), la situazione professionale o familiare, l'eventuale consumo di droghe, l'età al momento della valutazione COMPAS e l'età del primo arresto. Tutti fattori che, secondo la teoria criminologica accolta da COMPAS, sono decisivi per la previsione della recidiva¹²⁰. L'analisi algoritmica di tali fattori viene svolta con tecniche di *machine learning*¹²¹ e produce un indice di recidiva da 1 a 10, da 1 a 4 basso (*Low*), da 5 a 7 Medio (*Medium*), da 8 a 10 Alto (*High*).

¹¹⁶ Guida COMPAS, p. 2.

¹¹⁷ *Ibidem*, p. 12 .

¹¹⁸ Occorre rimarcare il fatto che i fattori criminogeni utilizzati sono il frutto di una scelta della società produttrice del software, che ha accolto alcune teorie criminologiche che tuttavia non hanno unanime riconoscimento, si cfr. in argomento K. B. FORREST, *When machines can be judge, jury, and executioner: justice in the age of artificial intelligence*, cit. pp. 69-70.

¹¹⁹ Consultabile al seguente indirizzo: <https://www.documentcloud.org/documents/2702103-Sample-Risk-Assessment-COMPAS-CORE>.

¹²⁰ Guida COMPAS, pp. 31 ss.

¹²¹ *Ibidem*, p. 13, in cui si legge: «*Machine learning methods are highly flexible and are usually preferred in applications where there is a complex relation- ship between predictors and outcomes*».

A causa dell'utilizzo di fattori standardizzati (cioè frutto di una elaborazione statistica di dati riferiti a gruppi di persone) non è da sottovalutare lo scenario in cui il sistema produca dei risultati controintuitivi: la guida stessa specifica che l'algoritmo potrebbe determinare un rischio alto di recidiva per un soggetto che non ha commesso alcun reato (perché, in ipotesi, è giovane e vive in un contesto socioeconomico classificato come pericoloso) e un rischio basso invece per un altro che ha alle spalle una complessa storia criminale. Ciò dipende dal fatto che, come chiarito dalla Guida stessa, «[r]isk assessment is about predicting group behavior (identifying groups of higher risk offenders) – it is not about prediction at the individual level. Your risk score is estimated based on known outcomes of groups of offenders who have similar characteristics»¹²². Non è dunque – e nemmeno potrebbe essere – una valutazione del rischio di recidiva individuale quanto piuttosto una valutazione del rischio del gruppo al quale il soggetto è stato identificato come appartenente¹²³.

In questo caso, specifica la Guida, spetterà al giudice considerare adeguatamente le circostanze del caso e sovertire (*override*) il risultato di COMPAS.

Si legge nella Guida: «we would expect staff to disagree with an actuarial risk assessment (e.g. COMPAS) in about 10% of the cases due to mitigating or aggravating circumstances which the computer is not sensitive to. In those cases staff should be encouraged to use their professional judgment and override the computed risk as appropriate - documenting it in COMPAS with the Override Reason - for monitoring by supervisory staff»¹²⁴. Ora, attraverso i dati relativi ai fattori di rischio l'algoritmo di COMPAS, mediante procedure di *pattern recognition*¹²⁵, classifica automaticamente ogni individuo in una “tipologia” che meglio si adatta al profilo del soggetto analizzato. Le tipologie di COMPAS sono

¹²² *Ibidem*, p. 35.

¹²³ Si vedrà nel prosieguo che una delle contestazioni che verranno mosse ai tools sarà proprio la mancanza di una valutazione di rischio individualizzata.

¹²⁴ Guida COMPAS, p. 35.

¹²⁵ Il *pattern recognition* è una tecnica di *machine learning* che consiste nell'identificazione di modelli partendo da dati grezzi, al fine di identificare la corretta classificazione per i nuovi dati. Per un approfondimento si cfr. la seguente pagina web: <https://www.mygreatlearning.com/blog/pattern-recognition-machine-learning/>.

otto per gli uomini¹²⁶ e otto per le donne¹²⁷. Sulla base dello *standard* tipologico assegnato all'individuo, gli operatori basano il proprio *case plan* (e il giudice penale la determinazione dell'ammontare della pena).

Ultima caratteristica di COMPAS da mettere in luce per comprendere come funziona l'algoritmo al suo interno è il *dataset* utilizzato per la progettazione e l'addestramento (*training*) dell'algoritmo. In che modo, cioè, il sistema, elaborando gli *input* che provengono dai dati relativi a ciascun soggetto, individua una tipologia standard da assegnare all'imputato e il livello di rischio di recidiva. È sicuramente indicativo che solo poche righe della Guida siano dedicate al *dataset*, che è ciò che nutre l'algoritmo e ne determina il funzionamento: come si vedrà nel prosieguo del lavoro (si v. *infra*, cap. VI §§ 1-2), uno dei problemi che pongono siffatti sistemi è l'accessibilità dei dati, e il disvelamento del *dataset* sarebbe già un primo passo per assicurare una maggiore trasparenza dell'algoritmo stesso.

Ad ogni buon conto, nella Guida si sottolinea che la “base di conoscenza” dell'algoritmo è costituita da un'indagine a campione su trenta mila persone svolta negli Stati Uniti tra il 2004 e 2005 su soggetti sia condannati e in carcere, sia non ancora condannati, sia in fase di esecuzione della pena al di fuori della struttura carceraria. All'interno del primo campione i progettisti, con modalità non rinvenibili e dunque non scrutinabili, hanno selezionato un campione ridotto di 7381 soggetti, che rappresentano il *dataset* di COMPAS; pertanto, le inferenze che compie l'algoritmo si fondano sostanzialmente sulla caratterizzazione criminale di

¹²⁶ In particolare, le otto tipologie in cui si può venire classificati sono: 1. “*Chronic drug abusers – most non-violent*”; 2. “*Low risk situational – fighting/domestic violence caution*”; 3. “*Chronic alcohol problems – DUI, domestic violence*”; 4. “*Socially marginalized – poor, uneducated, stressed, habitual offenders*”; 5. “*Criminally versatile – young marginalized persons often gang affiliated*”; 6. “*Socially isolated long term substance abuse – multiple minor and mostly non-violent offenses*”; 7. “*Serious versatile high risk individuals*”; 8. “*Low risk situational accidental category*”. Si v. Guida COMPAS, pp. 54 ss.

¹²⁷ In particolare: 1. “*Drug problems and anti-social sub-cultural influences – some with relationship conflicts*”; 2. “*Family disorganization and inadequate parenting – residential instability and minor non-violent offences*”; 3. “*Chronic substance abusers – women with higher social resources than other groups*”; 4. “*Marginalized poor and isolated older women – economic survival crimes*”; 5. “*Young antisocial poorly educated women with some violent offences and early delinquency onset*”; 6. “*Chronic long term criminal history A – multiple co-occurring social and psychological risk factors*”; 7. “*Chronic long term criminal history B – multiple co-occurring problems and high risk*”; 8. “*Late starters with multiple strengths and fewer risk factors – minor non-violent offence history*”. Si v. Guida COMPAS, pp. 59 ss.

tale gruppo campione. Benché la Guida non fornisca altre spiegazioni, i limiti dell'utilizzo di una simile base conoscitiva sono evidenti. In primo luogo, i dati si riferiscono a un campione di circa trentamila soggetti che erano nel circuito criminale tra il 2004 e il 2005 e non sembra ci sia stato alcun aggiornamento, di tal che COMPAS determinerà il rischio di recidiva di un individuo basandosi sui livelli di recidiva esistenti nei primi anni 2000¹²⁸. Inoltre, non sembra ci sia una differenziazione territoriale tra i dati raccolti (contestazione sollevata dinanzi alle Corti), onde una rappresentazione distorta di quelli che possono essere i diversi indici di recidiva dei diversi Stati o delle diverse regioni.

Ultimo aspetto da considerare sul funzionamento di algoritmi come COMPAS, anche se solo lateralmente verrà toccato nella presente ricerca, riguarda l'effetto discriminatorio che l'algoritmo *necessariamente* produce¹²⁹ e gli infiniti studi sulla sua *accuracy* predittiva¹³⁰.

5. L'impiego dei tools nel contesto giudiziario statunitense: in particolare il *sentencing*

¹²⁸ Si pensi, ad esempio, a una giurisdizione statale che utilizza COMPAS per fare una predizione sull'attuale rischio di un imputato afroamericano, utilizzando dati di un periodo in cui (nella specie all'inizio degli anni 2000) l'arresto per consumo di marijuana era particolarmente elevato per le minoranze etniche. Si cfr. K. B. FORREST, *When machines can be judge, jury, and executioner: justice in the age of artificial intelligence*, cit., p. 77.

¹²⁹ È facilmente intuibile che quando i dati sono raccolti all'interno di un sistema giurisdizionale discriminatorio, l'algoritmo riprodurrà le medesime discriminazioni di quel sistema (in informatica si usa l'efficace espressione *garbage in, garbage out*). Per un approfondimento si cfr. C. ORWAT, *Risks of Discrimination through the Use of Algorithms*, in *Federal Anti-Discrimination Agency*, 2020 e S. A. FRIEDLER - C. SCHEIDEGGER - S. VENKATASUBRAMANIAN, *On the (im)possibility of fairness*, in *ArXiv*, 23 settembre 2016.

¹³⁰ Come è evidente, si tratta di problematiche che, sebbene assai interessanti, solo lateralmente toccano il tema indagato nel presente contributo e che quindi non saranno oggetto di specifico approfondimento. Sulla tendenza dei *tools* a riprodurre discriminazioni esistenti nonché sui difetti di affidabilità, si cfr., *ex multis*, J. KLEINBERG - J. LUDWIG - S. MULLAINATHANY - C. R. SUNSTEIN, *Discrimination in the age of algorithms*, in *Journal of Legal Analysis*, vol. 10, 2018, pp. 113- 174; J. ANGWIN - J. LARSON - S. MATTU - L. KIRCHNER, *Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks*, Propublica, 23 maggio 2016; A. Z. HUQ, *Racial equity in algorithmic criminal justice*, in *Duke Law Journal*, vol. 68 (6), 2019, pp. 1043-1134; T. Z. ZARSKY, *An Analytic Challenge: Discrimination Theory in the Age of Predictive Analytics*, in *A Journal of Law and Policy for the Information Society*, vol. 14 (1), 2017, pp. 11-35.

La premessa sul funzionamento di siffatti algoritmi ci permette ora di soffermarci sulle modalità con cui essi vengono impiegati nel contesto giudiziario statunitense e gli ambiti maggiormente toccati dal loro utilizzo.

Tradizionalmente i *risk assessment tools* venivano usati nei programmi riabilitativi all'interno delle carceri per predisporre il programma di intervento più adeguato alle caratteristiche dell'imputato¹³¹ nonché nelle decisioni afferenti alla fase esecutiva della pena, quali le determinazioni sulla libertà condizionale (*parole decisions*), le condizioni per la concessione o la revoca della libertà vigilata (*supervision*) o i programmi di reinserimento¹³². Tale impiego è senz'altro ancora quello più frequente.

A queste applicazioni più risalenti si affiancano oggi altri utilizzi, entrati ormai nell'ordinarietà del processo penale.

Un uso molto diffuso riguarda la fase pre-processuale (*pre-trial*) con finalità cautelari. A seguito della privazione della libertà personale e prima della celebrazione del processo tali strumenti vengono impiegati per valutare la probabilità che un imputato si presenti in udienza; del pari, essi mirano a prevedere il rischio per la comunità nel caso in cui l'imputato venga rilasciato¹³³. È facile intuire che l'obiettivo ultimo dell'impiego dei *tools* nella fase cautelare è quello di ridurre il numero di persone all'interno del carcere prima di una sentenza di condanna¹³⁴.

Diversa è l'applicazione dei *tools* nel *sentencing*, su cui questo lavoro concentrerà la propria attenzione per due ordini di ragioni: innanzitutto è intorno

¹³¹ D. KEHL - P. GUO - S. KESSLER, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessment in Sentencing*, cit., p. 10.

¹³² M. HAMILTON, *Risk-needs assessment: consituitional and ethical challenges*, cit., pp. 234-235.

¹³³ M. BRENNER et al, *Constitutional Dimensions of Predictive Algorithms in Criminal Justice*, cit. p. 270.

¹³⁴ Lo strumento più utilizzato nella fase pre-processuale è il *Public Safety Assessment (PSA)*, impiegato in 29 giurisdizioni americane. Il PSA non cerca di identificare trattamenti riabilitativi per gli imputati, ma piuttosto è progettato per aiutare a prendere decisioni riguardo all'opportunità o meno che un individuo sia detenuto o rilasciato prima del processo. Lo strumento effettua una valutazione del rischio basata sui fattori di rischio statici (v. *supra*) e tale classificazione viene utilizzata per determinare se una persona è a basso rischio e può quindi essere rilasciata senza rischi per la pubblica sicurezza. Per degli studi su questo algoritmo si cfr. D. KEHL - P. GUO - S. KESSLER, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessment in Sentencing*, cit., p. 10.

all'utilizzo degli strumenti algoritmici in questa processuale che si è sviluppata gran parte della giurisprudenza rilevante, sia con riguardo all'eventuale *abuse of discretion* da parte del giudice, sia rispetto alla violazione delle garanzie del giusto processo. Inoltre, è proprio nel *sentencing* che il giudice è direttamente chiamato a considerare e interpretare il risultato algoritmico per condannare l'individuo a un determinato ammontare di pena, quindi assumendo una decisione foriera di implicazioni proprie di una sentenza di condanna.

Al riguardo, come giustamente notato, mentre gli elementi che un giudice deve valutare nella fase processuale sono ben definiti e lo strumento di previsione del rischio può senz'altro venire in aiuto per determinare l'eventuale messa in libertà dell'imputato prima della celebrazione del processo, una sentenza di condanna comprende un insieme ben più ampio di considerazioni¹³⁵. Nello specifico un giudice deve decidere le modalità di esecuzione della pena (se in carcere o al di fuori di esso); una volta decise le modalità, dovrà determinare, all'interno di una forbice edittale, la lunghezza della pena sulla base di obiettivi che spaziano, come già visto, dal profilo retributivo al fine rieducativo passando per la salvaguardia della pubblica sicurezza. Insomma, mentre nel *pre-trial* la decisione del giudice è meramente binaria (è preferibile che l'imputato stia in carcere o no in attesa del processo?) e non devono entrare valutazioni né di stampo retributivo né riabilitativo/rieducativo, il giudizio nel *sentencing* è ben più complesso e l'impiego di sistemi algoritmici pone problemi particolarmente spinosi¹³⁶.

Spiegate le ragioni per cui l'interesse della presente ricerca ricade principalmente sull'applicazione di algoritmi alla fase di determinazione della pena, occorre chiedersi se l'inserimento di tali strumenti in questa fase sia stata

¹³⁵ D. KEHL - P. GUO - S. KESSLER, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessment in Sentencing*, cit., p. 13.

¹³⁶ M. BRENNER *et al*, *Constitutional Dimensions of Predictive Algorithms in Criminal Justice*, cit., p. 277 in cui si dice «Some critics have identified the focus on recidivism in sentencing as posing particular concerns. While the decision to detain a defendant pending trial is binary, the sentencing decision is not binary. Sentencing requires the judge to decide the severity of the sentence. Since there is little evidence that a longer sentence will impact a defendant's recidivism, some scholars have questioned whether a higher likelihood of recidivism should necessarily result in a longer sentence».

accompagnata da una regolamentazione legislativa o si sia stata imposta per prassi all'interno del processo giudiziario.

A tal proposito, esistono alcune indicazioni para-normative e linee-guida che prendono in considerazione l'impiego dei *tools* nel *sentencing* nelle giurisdizioni statali.

In uno studio apparso nel 2017 all'esito di un progetto condotto dal *National Centre for State Courts (NCSC)* si indicava che almeno diciotto stati avessero introdotto, tramite modifiche statutarie o nuove *policy* amministrative e giudiziarie, l'utilizzo dei *tools* nel *sentencing*¹³⁷ (ma probabilmente almeno metà degli Stati fanno uso di strumenti di *risk assessment* in tale fase processuale¹³⁸). In tale studio vengono espressi nove principi (*guiding principles*) che dovrebbero guidare le Corti a un utilizzo corretto dei *risk tools*.

È utile qui dare conto di alcuni di siffatti principi per poi vedere, nell'analisi della casistica rilevante, se tali principi sono stati rispettati dalle Corti.

I primi due principi, come vedremo ricchi di contraddizioni, prevedono che la valutazione del rischio

should not be used as an aggravating or mitigating factor in determining the severity of an offender's sanction

ma deve essere solo

one factor to consider in determining whether an offender can be supervised safely and effectively in the community.

Particolarmente rilevante il quarto principio che pone l'accento sul *training* degli attori in gioco prevedendo che

Education regarding the nature and use of risk and needs assessment information is critical for all stakeholders (e.g., judge, defense attorney, prosecutor, probation officer, victim advocate)

¹³⁷ NATIONAL CENTRE FOR STATE COURTS, *Use of risk and needs assessment information in state sentencing proceedings*, settembre 2017, p. 3.

¹³⁸ B. L. GARRETT - J. MONAHAN, *Judging Risk*, cit., p. 453.

Infine, è utile richiamare il settimo principio e il nono principio, i quali richiedono che le giurisdizioni

should select instruments that fit their assessment needs and that have been properly validated for use with their offender populations

nonchè

should routinely review data related to the process and outcomes of using risk and needs assessment information and revise the system as appropriate to enhance effectiveness.

Un altro strumento giuridico che ha contribuito a modellare e favorire l'impiego dei *risk tools* all'interno delle giurisdizioni statali è l'inclusione nella proposta finale del 2017 del *Model Penal Code* dell'*American Law Institute* di una sezione dedicata agli strumenti di valutazione del rischio di recidiva¹³⁹.

Nel documento di legge che le commissioni legislative statali «*shall develop actuarial instruments or processes, supported by current and ongoing recidivism research, that will estimate the relative risks that individual offenders pose to public safety through their future criminal conduct*»

Inoltre, «*[w]hen these instruments or processes prove sufficiently reliable, the commission may incorporate them into the sentencing guidelines*»¹⁴⁰. Infine, il documento precisa che l'utilizzo di tali sistemi debba in ultimo risiedere «*in the discretion of the trial judge*»¹⁴¹.

Se queste sono le indicazioni di massima, all'interno delle giurisdizioni statali la regolamentazione dell'impiego di algoritmi predittivi è particolarmente

¹³⁹ Il *Model Penal Code* è un progetto dell'*American Law Institute* – un'organizzazione indipendente i cui lavori hanno un'estrema influenza sulle *policy* degli Stati Uniti, che mira a fornire un modello di codice penale per favorire la standardizzazione della legislazione all'interno delle giurisdizioni statali. Il primo progetto fu pubblicato nel 1962, aggiornato poi nel 2017.

¹⁴⁰ *Model Penal Code*, final draft, 2017, p. 386, disponibile al seguente link: https://robinainstitute.umn.edu/sites/robinainstitute.umn.edu/files/2022-02/mpcs_proposed_final_draft.pdf.

¹⁴¹ *Ibidem*, p. 387.

frammentata e solo in alcuni casi gli statuti statali contengono una regolazione puntuale¹⁴².

Solo per fare qualche esempio, in Arizona non si indica quale uso il giudice debba fare dei *tool* ma si prevede che il *Presentence Report* (si v. *infra*, § successivo) «*shall also contain case information related to criminogenic risk and needs as documented by the standardized risk assessment and other file and collateral information*»¹⁴³.

Invece, in Kentucky si prevede espressamente che il *sentencing judge* debba considerare (*shall consider*) «*the results of a defendant's risk and needs assessment included in the presentence investigation*» and «*[t]he likely impact of a potential sentence on the reduction of the defendant's potential future criminal behavior*»¹⁴⁴.

In termini simili si presenta anche lo statuto dell'Oklaoma che impone che la sentenza «*shall be based upon the supervision and intervention report from the risk and needs assessment*»¹⁴⁵.

In termini non obbligatori ma facoltativi si esprime invece lo statuto del Delaware, il quale prevede che la Corte possa (*may*) chiedere un *presentence report* che «*should include administration of an objective risk and needs assessment instrument*»¹⁴⁶.

In alcuni Stati invece – come l'Indiana¹⁴⁷, il Wisconsin¹⁴⁸ o il West Virginia¹⁴⁹ – sono le Corti supreme che, nel silenzio degli statuti, hanno autorizzato l'utilizzo dei *tools* da parte del giudice nel *sentencing*.

¹⁴² Lo stato della Virginia è stato il primo, nel 1994, a implementare un *tool* nel *sentencing*, B. L. GARRETT - J. MONAHAN, *Judging Risk*, cit., p. 442.

¹⁴³ *Ariz. Code of Judicial Admin.* § 6-201.01.

¹⁴⁴ *Ky. Rev. Stat. Ann.* § 532.007(3)(a)-(b).

¹⁴⁵ *Ok Stat*, 22 § 988.19 (2022).

¹⁴⁶ *Del. Code Ann.* tit. 11, § 4331.

¹⁴⁷ Si veda *infra* (capitolo VIII § 2) il caso *Malenchik*.

¹⁴⁸ Si veda *infra* (capitolo VIII § 3) il caso *Loomis*. Il codice penale del Wisconsin si limita a prevedere che «*the community services target offenders at a medium or high risk for revocation or recidivism as determined by valid, reliable, and objective risk assessment instruments that the department has approved*», Sez. 301.068.

¹⁴⁹ In una *concurring opinion* di una sentenza del 2015, *State v. Rogers*, No. 14-0373, 2015 WL 869323, si specifica, citando un *memorandum* del 2013, che sebbene i dipartimenti debbano preparare una valutazione del rischio di recidiva, «*circuit judges do not have to use the results of the LS/CMI in their sentencing decisions, emphasizing that the use of the information in an LS/CMI assessment is 'entirely left to [the circuit judges'] discretion*».

6. L'acquisizione da parte del giudice del risultato algoritmico

Chiarito il contesto regolatorio dei *tools* occorre in ultimo precisare con quali modalità, in sede processuale, il giudice prende conoscenza del risultato prodotto da uno strumento algoritmico di valutazione del rischio.

Una volta che il reo viene dichiarato colpevole dalla giuria, di norma il giudice richiede il *Pre-sentence Investigation Report* (PSI o PSIR), un rapporto predisposto solitamente da un funzionario della Corte contenente dati e informazioni inerenti al soggetto condannato, che il giudice è chiamato a valutare esercitando il massimo grado di discrezionalità (*great o board discretion*)¹⁵⁰.

Al riguardo, nella fase di *sentencing* i limiti normativi rispetto alla valutazione discrezionale delle prove e degli altri elementi che il giudice compie prima della determinazione finale sulla pena da applicare sono estremamente esigui: se la fase più propriamente processuale, e cioè la parte di trial il cui esito conclusivo sfocia nella dichiarazione di colpevolezza (*plea of guilty*), è informato al rispetto di rigidi confini nella valutazione delle prove, nella fase di determinazione della pena il giudice non è rigidamente vincolato ai fatti rilevanti nel giudizio, ma la sua valutazione deve considerare, nell'ottica di reinserimento del condannato nella società, anche elementi specifici della personalità dello stesso¹⁵¹.

In questo contesto processuale – negli Stati che hanno implementato l'impiego dei *tools* nel *sentencing* – all'interno del PSI vengono riportati anche i risultati dell'indagine realizzata attraverso il sistema algoritmico¹⁵², il quale normalmente indica, come visto, un indice di recidiva che varia tra alto, medio e basso e di cui il giudice potrà tenere conto per determinare l'ammontare della pena, le misure

¹⁵⁰ La giurisprudenza, sia a livello federale sia a livello statale, è assolutamente costante. Si v., di recente, in ambito federale, Supreme Court Of The United States, *Concepcion v. United States*, 27 giugno 2022.

¹⁵¹ D. KEHL - P. GUO - S. KESSLER, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing*, cit., pp. 15 ss.

¹⁵² Al sito di NATIONAL CENTER OF STATE COURTS, che da più di cinquant'anni raccoglie dati delle Corti dei diversi Stati, è possibile visionare un esempio di una sezione del PSI contenente i risultati del *risk assessment tool* al seguente indirizzo: https://www.ncsc.org/data/assets/pdf_file/0018/27171/ran-appendix-b.pdf.

trattamentali più idonee per il condannato e l'eventuale esecuzione della pena al di fuori del carcere¹⁵³.

¹⁵³ Per uno studio su come i giudici interpretano i risultati nei *risk assessment tools*, ancorché nei casi in cui essi sono utilizzati nel pre-trial, si cfr. M. DEMICHELE - M. COMFORT - K. BARRICK - P. BAUMGARTNER, *The Intuitive-Override Model: Nudging Judges Toward Pretrial Risk Assessment Instruments*, in *Federal Probation Journal*, vol. 85, n. 2, 2021.

LA GIUSTIZIA ALGORITMICA NEGLI ORDINAMENTI EUROPEI: IL PARADIGMA DELLA GIUSTIZIA ‘PREDITTIVA’

SOMMARIO: 1. Giustizia algoritmica e giustizia predittiva - 1.1. Il fondamento teorico della giustizia predittiva in Italia e in Europa - 1.2. Il funzionamento dei moderni sistemi di giustizia predittiva - 2. Studi e sperimentazioni di giustizia predittiva in Europa - 2.1. Il *machine learning* e il *natural language processing* per prevedere le decisioni della Corte europea dei diritti dell'uomo - 2.2. Il caso francese e gli *open data* - 2.3. Esperimenti di giustizia predittiva in alcune giurisdizioni italiane

1. Giustizia algoritmica e giustizia predittiva

Come visto nella premessa a questa parte del lavoro, in Europa non vi sono ancora esperienze di giustizia algoritmica presenti nei sistemi giudiziari, rimanendo esse ancorate a uno stadio ancora di sperimentazione e studio. Ciò detto, è utile dare conto dei progetti più rilevanti emersi negli ultimi anni nell'ordinamento europeo, non prima però di aver definito compiutamente la direzione della giustizia algoritmica oggi in Europa, vale a dire la cd. *giustizia predittiva*.

Per come definita in questo lavoro, la giustizia algoritmica nella prospettiva costituzionalistica qui adottata comprende al suo interno la giustizia predittiva, ossia la progettazione di strumenti predittivi in grado di prevedere la futura decisione del giudice.

Tale inserimento, tuttavia, potrebbe avvenire allorché avessimo contezza di ordinamenti che prevedano che un giudice possa o debba avvalersi di tali sistemi per raggiungere lui stesso una decisione del caso sottoposto al suo esame, ipotesi che ancora non si è realizzata nella prassi giudiziaria. Come già indicato, è la stessa Carta etica sull'uso dell'intelligenza artificiale a precisare che la giustizia predittiva è ancora relegata a un ambito privato e commerciale e non è impiegata dai tribunali:

«in quanto la maggior parte delle iniziative provengono dal settore privato e sono raramente integrate nelle politiche pubbliche»¹⁵⁴.

Della giustizia predittiva si sono fornite miriadi di definizioni e tale locuzione mantiene tutt'ora un certo grado di ambiguità che non aiuta a comprenderne i contorni. Tale ambiguità è ben espressa nella Carta etica in cui si precisa che «[l]'espressione giustizia predittiva dovrebbe essere abbandonata in quanto è ambigua e ingannevole. Tali strumenti sono basati su metodi di analisi della giurisprudenza che utilizzano metodi statistici che non riproducono in alcun modo il ragionamento giuridico, ma possono cercare di descriverlo»¹⁵⁵.

Il termine predizione è quindi inappropriato: come sottolineato da Rieling, già magistrata ad Amsterdam e una delle più acute studiose in Europa del rapporto tra tecnologia e giustizia¹⁵⁶, il risultato del sistema algoritmico non è una predizione del futuro – che richiama a ispirazioni premonitriche mistiche o chiaroveggenza – quanto piuttosto la previsione (in inglese *forecast*) – del possibile esito di una controversia sulla base di un'analisi statistica fondata su precisi parametri (soprattutto testuali, come si vedrà) di precedenti decisioni¹⁵⁷.

Ora, come già sottolineato, dato il successo che ha assunto tale termine negli studi accademici europei verrà utilizzato anche nel presente lavoro, benché – come già accennato – in un'accezione ristretta. In questo senso ci aiuta la definizione impiegata da Garapon e Lassègue nel celebre testo *Justice digitale*, in cui si scrive che la giustizia predittiva si riferisce alla capacità attribuita alle macchine di analizzare rapidamente il diritto rilevante per un caso e di anticipare la probabilità delle decisioni che potrebbero essere assunte¹⁵⁸.

¹⁵⁴ Carta etica, cit., p. 15

¹⁵⁵ Carta etica, cit. pp. 38.

¹⁵⁶ Si v. in particolare il grande lavoro di ricerca confluito in D. REILING, *Technology for Justice. How Information Technology Can Support Judicial Reform*, Leiden 2009.

¹⁵⁷ D. REILING, *Courts and Artificial Intelligence*, in *International Journal for Court Administration*, 8, 2020.

¹⁵⁸ Si cfr. A. GARAPON - J. LASSÈGUE, *Justice digitale*, cit, p. 219: «La justice prédictive désigne stricto sensu la capacité prêtée aux machines de mobiliser rapidement en langage naturel le droit pertinent pour traiter une affaire, de le mettre en contexte en fonction de ses caractéristiques propres (lieu, personnalité des juges, des cabinets d'avocats, etc.) et d'anticiper la probabilité des décisions qui pourraient intervenir. L'expression est devenue générique en renvoyant à toutes les innovations numériques dans le domaine du droit».

Se per alcuni autori, soprattutto italiani, giustizia predittiva comprende qualunque ipotesi di strumenti predittivi impiegati nel processo, e quindi anche in ipotesi i *tools* impiegati nel processo penale statunitense¹⁵⁹, nel presente lavoro si adotta quindi un'accezione ristretta, per diverse ragioni.

Innanzitutto è la stessa Carta etica che definisce la giustizia predittiva un'«analisi di una grande quantità di decisioni giudiziarie mediante tecnologie di intelligenza artificiale al fine di formulare previsioni sull'esito di alcune tipologie di controversie specialistiche»¹⁶⁰. Inoltre, se da una parte in Europa la letteratura accademica è più propensa a definire la giustizia predittiva nel senso ristretto che abbiamo indicato¹⁶¹, dall'altra nella letteratura statunitense non vi è alcun autore che si occupa dei temi legati alla giustizia algoritmica che definisca gli strumenti di previsione del rischio di recidiva come strumenti di giustizia predittiva.

Un'ultima precisazione: come abbiamo visto strumenti di giustizia predittiva sono già impiegati dal settore privato (studi legali, imprese assicuratrici) per fornire assistenza ai clienti e determinare le probabilità di successo di una controversia. Essendo questo un lavoro che si concentra sulla giustizia algoritmica (per come l'abbiamo definita) anche i ragionamenti sulla giustizia predittiva riguardano l'ipotesi (ancora non verificatasi) che tali sistemi vengano impiegati direttamente dal giudice come ausilio per assumere la propria decisione.

¹⁵⁹ Si cfr. per esempio A. SIMONCINI, *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*, cit., in cui l'A. distingue tra giustizia predittiva descrittiva e giustizia predittiva prescrittiva.

¹⁶⁰ Carta etica, cit., p. 47.

¹⁶¹ Oltre a Garapon e Lassègue, citati supra, si cfr. per esempio in Olanda F. BEX - H. PRAKKEN, *Can predictive justice improve the predictability and consistency of judicial decision-making?*, in E. SCHWEIGHOFER (a cura di), *Legal knowledge and Information Systems*, Amsterdam, 2021, pp. 207-213; in Germania si distingue gli strumenti di giustizia predittiva dai *tools* impiegati nel processo penale J. ULENAERS, *The Impact of Artificial Intelligence on the Right to a Fair Trial: Towards a Robot Judge?*, in *Asian J Law Econ*, 11 (2) 2020. Si v. anche, nel medesimo senso in Francia Z. NOURI - W. B. SALAH - N. A. OMRANE, *Artificial Intelligence and Administrative Justice: An Analysis of Predictive Justice in France*, in *Hasanuddin Law Review*, 10 no. 2, 2024, pp. 119-143. Anche in Italia alcuni preferiscono adottare la locuzione nella sua forma ristretta. Si cfr. R. BICHI, *Intelligenza digitale, giurimetria, giustizia predittiva e algoritmo decisorio. Machina sapiens e il controllo sulla giurisdizione*, in U. RUFFOLO (a cura di), *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Milano, 2020, pp. 423-447; S. ZOLEA, *The european Courts Faced with the unknowns of predictive Justice*, in A. DUMISTRESCU (a cura di), *Justice, Community and Freedom: IVR World Congress Proceedings*, 2022 pp. 436-445.

1.1. Il fondamento teorico della giustizia predittiva in Italia e in Europa

Come abbiamo sintetizzato il dibattito filosofico che ha accompagnato l'impiego di strumenti di previsione del rischio di recidiva nel processo penale statunitense, è utile richiamare a grandi linee il pensiero sui cui si regge l'idea di progettare sistemi di giustizia predittiva in grado di prevedere l'esito delle controversie.

È interessante notare che, rispetto al dibattito filosofico presente negli Stati Uniti intorno all'impiego dei *tools* predittivi (*supra* capitolo I § 2), l'obiettivo di fondo non sia così diverso. Come oltreoceano l'idea di dotare il giudice di strumenti "oggettivi" di decisione mirava a limitarne l'eccessiva discrezionalità e garantire una certa uniformità delle decisioni, così in Europa il fondamento teorico della giustizia predittiva affonda le radici nelle rinnovate esigenze di prevedibilità e certezza del diritto di fronte all'emergere di fenomeni legati alla cd. crisi della legalità¹⁶², quali l'avvento delle costituzioni, l'incidenza del cd. diritto giurisprudenziale (anche di natura sovranazionale, conseguenza dell'integrazione europea), la crisi della politica e la proliferazione delle fonti¹⁶³.

In un periodo storico definito in Italia "epoca della post-modernità"¹⁶⁴ o "età della giurisdizione"¹⁶⁵, l'ipotesi di impiegare strumenti algoritmici e predittivi per favorire la prevedibilità della giustizia appare a molti una valida soluzione.

Non solo in Italia, ma anche negli altri ordinamenti europei, incisi anch'essi dalle dinamiche appena menzionate, l'impianto dottrinale della giustizia predittiva ruota

¹⁶² Fenomeno approfondito ormai da tempo: si cfr. almeno, in termini critici, IRTI N., *Un diritto incalcolabile*, Giappichelli, Torino, 2016; ID., *L'età della decodificazione*, Giuffrè, Milano, 1999; v. anche ID., *La crisi della fattispecie*, in *Riv. Dir. Proc.*, 69, 1, 2014, pp. 36-44. In senso opposto, in termini positivi, P. GROSSI, *Ritorno al diritto*, Editori Laterza, Roma-Bari, 2015; più recentemente ID., *Storicità versus prevedibilità: sui caratteri di un diritto pos-moderno*, in *Questione giustizia*, trimestrale n. 4, 2018, pp. 17-23.

¹⁶³ Una sintesi esaustiva del dibattito italiano è rinvenibile nel numero 4 del 2018 di *Questione Giustizia* il cui primo obiettivo è intitolato *Una giustizia (im)prevedibile* e nei saggi raccolti in A. CARLEO (a cura di), *Calcolabilità giuridica*, Il Mulino, Bologna, 2017.

¹⁶⁴ P. GROSSI, *Storicità versus prevedibilità: sui caratteri di un diritto pos-moderno*, cit., p. 18. Sulla connessione tra queste dinamiche e l'ipotesi di automatizzare la giustizia si cfr. E. LONGO, *Predizione algoritmica e giustizia: le prospettive costituzionali*, in C. PANZERA - A. RAUTI (a cura di), *Attualità di diritto pubblico*, II, Editoriale Scientifica, 2022, pp. 87-125, spec. pp. 112-115

¹⁶⁵ A. DI PORTO, *Calcolo giuridico secondo la legge nell'età della giurisdizione, il ritorno del testo normativo*, in A. CARLEO (a cura di), *Calcolabilità giuridica*, Il Mulino, Bologna, 2017, p. 119.

intorno alla ricerca di maggiore prevedibilità delle decisioni giudiziarie¹⁶⁶. In questo senso, la stessa Carta Etica, sintetizzando tale dottrina, precisa che «[j]udicial decision processing by artificial intelligence, according to their developers, is likely, in civil, commercial and administrative matters, to help improve the predictability of the application of the law and consistency of court decisions»¹⁶⁷.

1.2. Il funzionamento dei moderni sistemi di giustizia predittiva

Come ci è apparsa utile un'analisi, per quanto sintetica, del funzionamento dei *tools* impiegati negli Stati Uniti, è altresì funzionale all'indagine comprendere come funzionano oggi gli strumenti di giustizia predittiva. Fino agli anni '90 la predizione in ambito giuridico veniva condotta con i cd. *sistemi esperti*¹⁶⁸, chiamati così per la loro capacità di ragionare e risolvere problemi come fossero esperti umani¹⁶⁹. Essi rientrano nei sistemi c.d. basati sulla conoscenza (*knowledge based system*), in cui i dati di *input* sono rappresentati dalle regole presenti in una determinata materia espresse in un linguaggio informatico¹⁷⁰ che la macchina elabora per giungere alla

¹⁶⁶ Si v., *ex multis*, F. BEX - H. PRAKKEN, *Can predictive justice improve the predictability and consistency of judicial decision-making?*, in E. SCHWEIGHOFER (a cura di), *Legal knowledge and Information Systems*, Amsterdam, 2021, pp. 207-213; B. BONDERO, *Justice prédictive: la fin de l'aléa judiciaire?*, in *Recueil Dalloz*, n. 10, 2017, pp. 532 ss.; Y. GAUDEMET, *La justice à l'heure des algorithmes. À propos de justice prédictive*, in *Revue du droit public et de la science politique en France et à l'étranger*, n. 3, 2018, pp. 651 ss.; G. YALCIN - E. THEMELI - E. STAMHUIS - S. PHILIPSEN - S. PUNTONI, *Perceptions of Justice By Algorithms*, in *Artificial Intelligence and Law*, 31 (2), 2023, pp. 269-292; A. GARAPON, *Les enjeux de la justice prédictive*, in *La semaine juridique*, 1, 2017, pp. 47-52.

¹⁶⁷ Carta Etica, cit., p. 5.

¹⁶⁸ La Carta etica definisce così un sistema esperto: «uno strumento in grado di riprodurre i meccanismi cognitivi di un esperto in un particolare campo. Più precisamente si tratta di un software in grado di rispondere a delle domande grazie a un ragionamento fondato su fatti e regole noti. Consiste in tre parti: una base di fatti; una base di regole; un motore inferenziale. Il motore inferenziale è in grado di utilizzare i fatti e le regole per produrre nuovi fatti fino a pervenire alla risposta alla domanda specialistica posta. La maggior parte dei sistemi esperti esistenti è fondato sui meccanismi della logica formale (logica aristotelica) e utilizza il ragionamento deduttivo» (traduzione nostra).

¹⁶⁹ A. SANTOSUOSSO - G. SARTOR, *La giustizia predittiva: una visione realistica*, cit., p. 1766.

¹⁷⁰ Cfr. *op. cit.* in cui gli autori portano un esempio utile per chiarire il funzionamento: «Si pensi, ad esempio, alle regole che stabiliscono che, per avere diritto a certe prestazioni previdenziali, sono richieste certe condizioni personali (età, stato di famiglia, ecc.) e certe condizioni patrimoniali (reddito, proprietà, ecc.), ulteriormente specificate in altre norme. Sistemi basati sulla conoscenza

risoluzione del problema attraverso un ragionamento sillogistico-deduttivo (secondo la formula *if-then*)¹⁷¹.

Tali sistemi non hanno prodotto significativi risultati nella previsione delle decisioni giurisprudenziali per diverse ragioni, in primo luogo per la difficoltà di raccogliere e immettere tutte le regole necessarie affinché il sistema raggiunga la conclusione richiesta. A questo si aggiunga che un esperto umano potrà ben avere alcune difficoltà a rendere le regole in modo chiaro, data la quantità di situazioni che si possono presentare in un unico problema¹⁷².

In secondo luogo, è complicato tradurre in un linguaggio informatico univoco fatto di regole e concetti la complessità del diritto e delle “regole” che un giudice segue per raggiungere una decisione di una controversia (non solo le norme, ma i precedenti, la coscienza sociale ecc...) ¹⁷³. Al tal riguardo nella Carta etica si scrive che «di fronte alla complessità generale dei sistemi giuridici, i vecchi sistemi esperti informatici, una volta annidate 200-300 regole logiche, raggiungevano rapidamente i loro limiti. La divisione del diritto in regole di produzione si è rivelata insufficiente a fornire una rappresentazione valida del corpus di conoscenze e di metodi che guidano un giurista»¹⁷⁴.

È dall'avvento dell'IA e dell'apprendimento automatico (*machine learning*), in cui la base di conoscenza non è fornita dall'uomo ma viene ricostruita dalla macchina sulla base dei dati inseriti nel sistema, che la giustizia predittiva ha tratto nuova linfa. Benché l'approfondimento del funzionamento e delle differenti tecniche di *machine learning* esuli dalle finalità del presente lavoro, una minima comprensione dell'ingegneria algoritmica alla base dei sistemi moderni di giustizia predittiva può aiutare a darne una visione realistica¹⁷⁵.

possono affrontare in modo soddisfacente il compito di applicare diligentemente le regole rilevanti per il caso in esame, una volta che il caso stesso sia stato formulato usando il linguaggio impiegato nelle regole»

¹⁷¹ G. SARTOR, *Intelligenza artificiale e diritto. Un'introduzione*, cit., p. 16.

¹⁷² K. WARWICK, *Artificial intelligence - The basics*, trad. it. di C. BARATTIERI DI SAN PIETRO - G. MAUGERI, Bologna, 2015, pp. 65 ss.

¹⁷³ Si cfr. ancora A. SANTOSUOSSO - G. SARTOR, *La giustizia predittiva: una visione realistica*, cit. p. 1767. Si v. a questo riguardo il problema dei “linguaggi” indicato nel capitolo introduttivo, § 1.

¹⁷⁴ Carta etica, cit., § 75 (traduzione nostra).

¹⁷⁵ A. SANTOSUOSSO - G. SARTOR, *La giustizia predittiva: una visione realistica*, cit. p. 1768.

Per capire la differenza tra sistema esperto e sistema di *machine learning* si consideri il seguente esempio banale: un ipotetico sistema per il riconoscimento di animali. Nel sistema esperto verrebbero codificate tutte le regole secondo cui un certo animale dovrebbe essere identificato come cane (quattro zampe, forma del muso, grandezza) e non come un altro animale. In un sistema di *machine learning* invece, l'algoritmo viene addestrato fornendo alla macchina¹⁷⁶ una quantità immensa di immagini di cani: una volta che alla macchina viene presentata una nuova immagine, se correttamente addestrata, essa sarà in grado di riconoscere il cane. Nell'ipotesi di giustizia predittiva, mentre in un sistema esperto – come visto – si dovrebbero codificare in un linguaggio informatico norme, precedenti e tutta la base di conoscenza che possiede un giudice per decidere un caso, nei sistemi di *machine learning* i dati consisteranno dei casi precedenti e dall'esito del giudizio. Il sistema, sulla base di parametri di volta in volta stabiliti (possono essere parametri testuali – come le occorrenze nella motivazione di determinate parole, o parametri extratestuali – come le caratteristiche del giudice), determina l'esito della controversia, con i limiti che a breve vedremo.

Tutti i sistemi di apprendimento automatico sono costituiti da due componenti fondamentali: l'algoritmo di apprendimento (*learning algorithm*) e il modello (*model*). L'algoritmo di apprendimento utilizza un set di dati di addestramento per costruire il modello. Generalizzando (astruendo) dai dati di addestramento il modello riesce a gestire nuovi casi che hanno delle somiglianze con quelli di addestramento. Se il nuovo caso assomiglia agli esempi precedenti, il sistema propone quindi una risposta simile a quella fornita per i casi analoghi¹⁷⁷.

I sistemi di apprendimento automatico possono utilizzare diversi metodi di apprendimento. Questi metodi si differenziano non solo per le prestazioni predittive, ma anche per la capacità di fornire spiegazioni. In effetti, spesso esiste un *trade-off* tra i due obiettivi: i sistemi che forniscono le previsioni più accurate

¹⁷⁶ L'addestramento può avvenire tramite l'essere umano o anche in automatico, cioè prelevate direttamente dalla macchina dal mondo digitale (per esempio il cd. *scraping* del web aperto).

¹⁷⁷ F. GALLI - G. SARTOR, *AI Approaches to Predictive Justice: A Critical Assessment*, in *Humanities And Rights Global Network Journal*, 5, 2023, p. 176.

sono spesso meno capaci di giustificare le loro decisioni (di questo tratteremo diffusamente *infra*, cap. VI).

Ovviamente la progettazione di tali modelli di *machine learning* per la giustizia predittiva è stata favorita, negli ultimi anni, dalla *digitalizzazione* delle fonti del diritto (in particolare le banche dati sempre più fornite di giurisprudenza), senza la quale tali sistemi non avrebbero la base di addestramento necessaria per funzionare. Ciò in quanto, come è stato detto, «dato che le predizioni di un sistema di AI si basano sui dati, quanto maggiori saranno la quantità e la qualità dei dati e quanto migliore il livello di analisi dei dati stessi, tanto maggiore sarà l'accuratezza della predizione possibile»¹⁷⁸.

Rispetto al funzionamento della *predizione* delle decisioni giurisprudenziali, comunque, occorre fare una ultima precisazione. Mentre i *tools* predittivi impiegati nelle giurisdizioni statunitensi sono impiegati direttamente dai giudici e mirano a prevedere eventi futuri (nella specie il comportamento dell'imputato), gli strumenti ad oggi sperimentati di giustizia predittiva, come vedremo, non sono impiegati dai giudici e non sono in grado di prevedere l'esito di casi non ancora decisi, ma "predicono" il risultato di una decisione che è già stata assunta sulla base di parti della decisione stessa (per esempio determinano il dispositivo sulla base di occorrenze lessicali presenti nella motivazione oppure indicano la decisione sulla base della porzione dei fatti indicati nella stessa)¹⁷⁹. Quindi l'obiettivo di tali sistemi non è la riproduzione di un ragionamento giuridico, ma l'individuazione delle correlazioni tra i diversi parametri di una decisione per prevedere gli esiti delle altre¹⁸⁰. Insomma non si predice l'esito di un caso futuro ma si predice il risultato di un caso già deciso. In ogni caso, gli esperimenti di sistemi in grado di predire l'esito dei procedimenti giudiziari sono stati di diversi tipi: la predizione può fondarsi su fattori non rilevanti nel merito del caso (es. identità dei giudici); può fondarsi sul contenuto testuale della sentenza (e sono gli esperimenti più recenti,

¹⁷⁸ A. SANTOSUOSSO - G. SARTOR, *La giustizia predittiva: una visione realistica*, cit. p. 1775.

¹⁷⁹ A. SANTOSUOSSO - G. SARTOR, *La giustizia predittiva: una visione realistica*, cit. p. 1772.

¹⁸⁰ Carta etica, cit., p. 23.

come si vedrà); può fondarsi su fattori giuridici rilevanti nel caso, anche se non testuali (come le caratteristiche delle parti)¹⁸¹.

Per progettare sistemi che realmente prevedano la decisione del giudice, la base di addestramento non potrebbe limitarsi ai precedenti, ma dovrebbe altresì basarsi su altri dati, spesso non liberamente disponibili, come gli atti di parte o l'intera decisione di un grado precedente rispetto all'esito del grado di appello¹⁸².

2. Studi e sperimentazioni di giustizia predittiva in Europa

Come si è detto, l'impiego di tecniche di *machine learning* per realizzare modelli di previsione delle decisioni giudiziarie è stata oggetto di alcuni studi accademici nonché di sperimentazioni all'interno di alcune giurisdizioni (spesso in collaborazione con le istituzioni universitarie).

Nei seguenti paragrafi, seguendo l'approccio induttivo segnalato nel capitolo introduttivo, tenteremo di segnalare le esperienze a nostro avviso più rilevanti, quelle cioè che hanno avuto un'eco più vasta nel dibattito costituzionalistico italiano ed europeo.

2.1. Il machine learning e il natural language processing per prevedere le decisioni della Corte europea dei diritti dell'uomo

Due studi che hanno avuto grande risonanza nel dibattito accademico hanno sperimentato algoritmi di *machine learning* basati sul *natural language processing*¹⁸³ (elaborazione del linguaggio naturale) per prevedere l'esito delle decisioni della Corte EDU.

¹⁸¹ Si cfr. ancora A. SANTOSUOSSO - G. SARTOR, *La giustizia predittiva: una visione realistica*, cit., p. 1772. Si v. anche D.M. KATZ, *Quantitative Legal Prediction--or--How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry*, in *Emory L.J.*, n. 4, 2012, pp. 906-966; K. D. ASHLEY - S. BRÜNINGHAUS, *Automatically classifying case texts and predicting outcomes*, in *Artificial Intelligence and Law*, 17 (2), 2009, pp. 125-165.

¹⁸² A. SANTOSUOSSO - G. SARTOR, *La giustizia predittiva: una visione realistica*, cit. p. 1772.

¹⁸³ Le tecnologie che si basano sul *natural language processing* consentono di prendere il linguaggio umano e "comprenderne" il significato ed è alla base degli strumenti di IA generativa

I risultati del primo studio, diretto da un ricercatore inglese dell'University College di Londra, sono stati presentati in un articolo scientifico pubblicato nel 2016¹⁸⁴.

L'obiettivo dichiarato dello studio era quello di prevedere se un determinato articolo della Convenzione era stato violato sulla sola base di informazioni testuali estratte da un caso, che comprende parti specifiche relative ai fatti, al diritto applicabile e alle argomentazioni presentate dalle parti in causa¹⁸⁵. Nel *disclaimer* iniziale i ricercatori specificano che la predizione non si basava sui ricorsi e le memorie delle parti – materiale a cui non avevano accesso – ma sulle parti descrittive delle decisioni della Corte già pubblicate, utilizzate come *proxies* ritenendo che quelle parti testuali della sentenza fossero simili a quelle contenute negli atti delle parti¹⁸⁶. Il *dataset* utilizzato ricomprendeva casi relativi agli artt. 3, 6 e 8 della Convenzione, per ognuno dei quali erano state selezionate da HUDOC tutte le sentenze (in inglese), in numero pari tra quelle di violazione e non violazione, per un totale di 584 decisioni.

Il modello ha prodotto un ottimo risultato, essendo stato in grado di prevedere le decisioni della Corte con un'accuratezza media del 79% (75% per i casi basati sull'art. 3 della Convenzione, sul divieto di tortura; 84% per l'art. 6, sul diritto a un processo equo; 78% in relazione all'art. 8, sul diritto al rispetto della vita privata e familiare).

Nella discussione dei risultati gli autori hanno tratto alcune conclusioni: innanzitutto strumenti di questo tipo possono essere utilizzati dai giudici «per identificare rapidamente i casi ed estrarre modelli correlati a determinati risultati», o ancora «per sviluppare indicatori preventivi per diagnosticare potenziali

come Chat GPT, si cfr J. KAPLAN, *Generative Artificial Intelligence: What Everyone Needs to Know*, Oxford, 2024, pp. 30 ss.

¹⁸⁴ N. ALETRAS - D. TSARAPATSANIS - D. PREOTIUC-PIETRO - V. LAMPOS, *Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective*, in *PeerJ Computer Science*, 2, 2016.

¹⁸⁵ *Ivi*, p. 2.

¹⁸⁶ *Ivi*, p. 4. Su questo assunto non mancano le perplessità, considerando che la motivazione della sentenza (compresa la descrizione dei fatti) risente giocoforza della decisione (di violazione o no) decisa dalla Corte. Su questo si v. F. PASQUALE - G. CASHWELL, *Prediction, persuasion, and the jurisprudence of behaviourism*, in *Univ. of Toronto L.J.*, vol. 68 (suppl. 1), 2018, pp. 63 s.

violazioni di articoli specifici nei ricorsi presentati e dare priorità ai casi in cui la violazione sembra molto probabile»¹⁸⁷. Inoltre, è emerso dalle inferenze compiute dal modello che le informazioni testuali contenute nella sezione dei “fatti” di un caso sono il miglior indice predittore della decisione, e cioè hanno una performance predittiva particolarmente elevata¹⁸⁸, cosa che peraltro non sorprende considerata la natura casistica della giurisprudenza della Corte EDU¹⁸⁹.

Gli autori hanno altresì precisato che tale studio possa aprire la strada a lavori futuri, ma che i problemi di accesso ai dati rappresentano un ostacolo significativo per gli scienziati che desiderano lavorare su questo tipo di dati giuridici. I grandi archivi come HUDOC, facilmente e gratuitamente accessibili, sono solo banche dati di giurisprudenza e non è possibile lavorare su altri tipi di dati che consentirebbero di giungere davvero a delle “previsioni” della decisione della Corte¹⁹⁰.

In questo senso, come è stato scritto, nello studio condotto «non si tratta della previsione di un fatto futuro, ma piuttosto dell’inferenza di un’informazione (la decisione della controversia) su un evento passato (la controversia già decisa), sulla base di informazioni sullo stesso evento che non includono il dato inferito (per es., la parte fattuale della decisione)»¹⁹¹. Visto il *dataset* molto limitato, l’impiego di sentenze già pubblicate (e quindi la mancanza di un valore predittivo), e l’inferenza soltanto testuale, lo studio non sembra poter fornire dei risultati particolarmente entusiasmanti o comunque che consentano di immaginare un utilizzo pregnante di tale tecnologia nel processo.

¹⁸⁷ *Ivi*, p. 3.

¹⁸⁸ *Ivi*, p. 11.

¹⁸⁹ Per dei commenti allo studio si v., per la dottrina italiana, C. BARBARO, *Uso dell’intelligenza artificiale nei sistemi giudiziari: verso la definizione di principi etici condivisi a livello europeo?*, in *Quest. giust.*, 4/2018, pp. 189 ss. Nella letteratura straniera, in termini critici, si v. J. MORISON - A. HARKENS, *Re-engineering justice? Robot judges, computerised courts and (semi) automated legal decision-making*, in *Legal Studies*, vol. 39 (4), 2019; F. PASQUALE - G. CASHWELL, *Prediction, persuasion, and the jurisprudence of behaviourism*, cit., pp. 63 ss.

¹⁹⁰ N. ALETRAS *et. al*, cit., p. 16.

¹⁹¹ A. SANTOSUOSSO - G. SARTOR, *La giustizia predittiva: una visione realistica*, cit. p. 1772. Anche Erik Longo si esprime nel medesimo senso «questi sistemi non “predicono” nulla ma si limitano a individuare una correlazione statistica: quindi calcolano delle probabilità. Non c’è nulla di “predittivo”, quindi, almeno se usiamo tale termine nel senso di immagine o anticipo del futuro». Si cft. E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit. p. 283.

Sulla base dei risultati ottenuti dallo studio britannico, una seconda indagine, pubblicata nel 2020 da ricercatori dell'Università di Groningen (Olanda)¹⁹², ha applicato le medesime tecniche di *machine learning* per l'analisi "predittiva" delle sentenze della Corte europea dei diritti. Questo è stato realizzato aumentando il numero degli articoli della Convenzione previsti (nove invece di tre) e dei casi considerati per articolo (1942 sentenze invece che 584), e sottoponendo al *machine learning* parti del caso diverse dal primo studio per diminuire i *bias* (in particolare, escludendo la sezione "law", che a volte menzionava esplicitamente il dispositivo finale di violazione non violazione: il punteggio finale delle previsioni è stato simile all'altro studio (77% contro 79%). Anche in questo caso, comunque, lo studio mirava a identificare quali fattori testuali nella motivazione hanno performance predittive migliori sul dispositivo della sentenza¹⁹³.

2.2. Il caso francese e gli open data

Anche in Francia è vivo il dibattito sulla giustizia predittiva e sull'intelligenza artificiale applicata al processo. Oltralpe la digitalizzazione totale dei processi e il ricorso strutturale agli *open data* giudiziari, come visto requisito necessario per qualunque esperimento di giustizia predittiva, ha ormai un lungo corso e ha assunto da tempo una portata strutturale e definitiva¹⁹⁴. Una riforma approvata nel 2016 (*loi* 2016/1321) *pour une La République Numérique* ha previsto la messa a disposizione gratuita del pubblico di tutti i dati in possesso della PA e di tutte le decisioni dei giudici mentre una riforma successiva, la *Loi n. 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice* ha previsto delle norme sull'anonimizzazione delle decisioni¹⁹⁵ (compreso il divieto riutilizzare i dati

¹⁹² M. MEDVEDEV - M. VOLS - M. WIELING, *Using Machine Learning to Predict Decisions of the European Court of Human Rights*, in *Artificial intelligence and Law*, vol. 28, 2020, pp. 237-266.

¹⁹³ *Ibidem*.

¹⁹⁴ Sul processo di digitalizzazione v. ancora E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit., pp. 106-112.

¹⁹⁵ Sul tema dell'anonimizzazione negli ordinamenti europei e il bilanciamento con le esigenze di pubblicità del processo si cfr. F. DONATI, *Trasparenza della giustizia e anonimizzazione dei provvedimenti giudiziari*, in A. PAJNO - F. DONATI - A. PERRUCCI (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, volume 2, 2022, pp 433-452.

identificativi dei giudici e dei membri della cancelleria dei tribunali allo scopo di profilare gli stessi)¹⁹⁶.

Nell'ambito della spinta verso la digitalizzazione e l'apertura dei dati giudiziari cresce nel dibattito pubblico francese l'interesse per la giustizia predittiva, soprattutto grazie ad un esperimento molto concreto portato avanti in due Corti d'Appello francesi: su iniziativa del Ministero della giustizia, le Corti di appello di Rennes e di Douai hanno testato, nella primavera del 2017, un *software* di giustizia predittiva applicato a controversie di appello principalmente in materia di calcolo di indennizzi¹⁹⁷. Il *software* era prodotto da una società privata (*Predictice*) alla quale il Ministero della giustizia, a seguito della vittoria di un bando pubblico, ha messo a disposizione i dati giudiziari che sarebbero serviti come *dataset*. Anche se il sistema algoritmico era stato progettato per essere impiegato nell'ambito privato del *legal tech*, l'impiego da parte dei magistrati mirava, anche qui, a testare uno strumento di ausilio alla decisione in modo da ridurre, se necessario, l'eccessiva variabilità delle decisioni giudiziarie, in nome del principio di uguaglianza dei cittadini di fronte alla legge.

Anche se l'esperimento ha avuto una portata limitata (è stato impiegato solo per due mesi da una decina di magistrati), ci servirà per comprendere alcune problematiche di natura costituzionale della giustizia predittiva che affronteremo nella terza parte del lavoro.

In sintesi, ai magistrati è stato chiesto di offrirsi volontari per testare l'applicazione per il calcolo degli indennizzi in alcune cause civili. Nella primavera del 2017 è stato organizzato un incontro per presentare il sistema ai magistrati interessati e avviare la sperimentazione.

¹⁹⁶ Art. 33 della legge in commento. Tale legge è stata oggetto di scrutinio da parte del *Conseil Constitutionnelle* francese (*Conseil Constitutionnel*, Décision n° 2019-778 DC, 21 marzo 2019) il quale non ha rilevato alcuna incostituzionalità della disposizione sul divieto profilare i giudici in quanto essa era finalizzata ad assicurare il corretto esercizio della giurisdizione, il quale avrebbe potuto subire alterazioni e pressioni in caso di profilazione che indurrebbero il giudice a decidere in base alla sua precedente giurisprudenza. Inoltre secondo il *Conseil* tale divieto non vulnerava nemmeno i principi del giusto processo in quanto la disciplina garantisce un'adeguata protezione della parità delle armi in sede processuale. Sul tema, diffusamente, E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit. pp. 106-118.

¹⁹⁷ Tale esperimento è segnalato anche nella Carta etica, cit., pp. 30-31.

La valutazione dell'esperimento da parte degli attori del sistema giudiziario – il Ministero, il presidente della Corte d'appello di Rennes e i magistrati – è stato oggetto di grande dibattito, anche con la società di *legal tech* che aveva progettato il sistema¹⁹⁸.

Maggiore interesse ai nostri fini è stata la valutazione da parte dei magistrati che si sono trovati tra le mani lo strumento predittivo: al termine del periodo di prova, i giudici che si sono offerti volontari si sono detti piuttosto delusi dallo strumento; le prime operazioni effettuate non li hanno convinti del loro valore, rispetto agli altri sistemi che utilizzavano abitualmente per prendere decisioni (come le banche dati)¹⁹⁹. Nel comunicato pubblicato nell'ottobre del 2017²⁰⁰, il primo presidente della Corte d'Appello di Rennes affermava: «[l]a giustizia deve riflettere sui suoi metodi di lavoro. Se *Predictice* si inserisce indubbiamente nella modernità, merita allo stato attuale di essere migliorato. Questo parere è condiviso da tutti i giudici che hanno partecipato con trasparenza e indipendenza. Il *software* non apporta alcun valore aggiunto rispetto ad altri motori di ricerca che già consentono un'analisi molto soddisfacente della giurisprudenza delle Corti d'Appello e della Corte di Cassazione»²⁰¹.

Ancora, sui limiti del *software*, il magistrato afferma che esso «si interessa solo al dispositivo di una decisione giudiziaria. L'algoritmo non sa leggere tutte le sottigliezze della motivazione, soprattutto quando la decisione è complessa. Per esempio, in appello, il dispositivo può molto bene confermare una parte della decisione dei giudici di primo grado ma il *software* non saprà dire quali elementi

¹⁹⁸ C. LICOPPE - L. DUMOULIN, *Le Travail des juges et les algorithmes de traitement de la jurisprudence. Premières analyses d'une expérimentation de "justice prédictive" en France*, in *Droit et Société*, n. 103, 2019, p. 535-554.

¹⁹⁹ In generale sul fallimento dell'esperimento e sulle conseguenze in Francia sul dibattito riguardante la giustizia predittiva si cfr. L. DUMOULIN. *De quoi la 'justice prédictive' est-elle le nom? Algorithmes, décision et jugement*, Science politique, Université Paris-Saclay, 2022.

²⁰⁰ Reperibile al seguente indirizzo: <https://www.dalloz-actualite.fr/interview/1-utilisation-de-l-outil-predictice-decoit-cour-d-appel-de-rennes>

²⁰¹ Traduzione nostra. Qui il testo originale in francese: «*La justice se doit de réfléchir à ses méthodes de travail. Si Predictice s'inscrit indéniablement dans la modernité, il mérite en l'état d'être amélioré. Ce constat est partagé par l'ensemble des juges qui ont participé en toute transparence et indépendance. Le logiciel n'apporte aucune plus-value par rapport à d'autres moteurs de recherche qui permettent déjà une analyse très satisfaisante de la jurisprudence des cours d'appel et de la Cour de cassation*».

sono stati confermati e in che modo. Non bisogna dimenticare che una decisione giudiziaria è un'opera intellettuale completa e spesso complessa»²⁰².

Ancor più determinanti nella valutazione negativa del *software* da parte dei magistrati sono stati i difetti di ragionamento, che hanno condotto a esiti aberranti o inappropriati, dovuti alla confusione tra semplici occorrenze lessicali presenti nelle motivazioni dei giudici e i nessi causali dei loro ragionamenti. Ancora il primo presidente affermava: «[a] volte, i risultati possono essere addirittura assurdi. Prendiamo l'esempio del calcolo dell'indennità per il licenziamento senza una causa reale e grave. Gli importi sono calcolati dal giudice in mensilità. Ma lo stipendio di un dirigente non è uguale a quello di un operaio. Basta una sentenza del tribunale che riguarda un dirigente per stravolgere completamente l'analisi, a seconda del caso o del settore lavorativo, e il risultato proposto non dice nulla sul caso da risolvere. Sarebbe più opportuno privilegiare un approccio basato sul numero di mensilità riconosciute, piuttosto che sapere se il probabile risultato sarà di 4.000 o 8.000 euro»²⁰³.

Come intuibile, il fallimento ha dato seguito a un grande dibattito in dottrina e giurisprudenza sull'opportunità della giustizia predittiva impiegata dai magistrati²⁰⁴, in cui tra l'altro si è sostenuta inoltre l'importanza che «l'autorità giudiziaria stessa organizzi l'alimentazione delle banche dati risultanti dalla sua

²⁰² Qui il testo originale in francese: «*Le logiciel ne s'intéresse qu'au dispositif d'une décision de justice. L'algorithme ne sait pas lire toutes les subtilités de la motivation, surtout lorsque la décision est complexe. Or, par exemple en appel, le dispositif peut très bien confirmer une partie de la décision des premiers juges et le logiciel ne saura pas dire quels éléments sont confirmés et de quelle manière. Il ne faut pas perdre de vue qu'une décision de justice est une œuvre intellectuelle complète et souvent complexe.*».

²⁰³ Qui il testo originale in francese: «*Parfois, les résultats peuvent même être aberrants. Prenons l'exemple d'un calcul d'indemnités de licenciement sans cause réelle et sérieuse. Les montants sont calculés par le juge en mois de salaire. Or le salaire d'un cadre n'est pas le même qu'un ouvrier. Il suffit qu'une décision de justice concerne un cadre pour fausser complètement l'analyse, selon l'affaire ou le bassin d'emplois, et pour que le résultat proposé ne dise rien du cas à résoudre. Il serait plus pertinent de privilégier une approche selon le nombre de mois de salaire alloué, plutôt que savoir si le résultat probable sera de 4 000 ou 8 000 €*».

²⁰⁴ L'anno successivo all'esperimento, in un colloquio tra i rappresentanti dell'Ordine degli Avvocati, del Consiglio di Stato e della Corte di Cassazione, il primo presidente della Corte di cassazione hanno preso entrambi le distanze dalla nozione di giustizia predittiva. B. LOUVEL, *Communication à l'occasion du colloque pour le bicentenaire de l'ordre des avocats au Conseil d'État et à la Cour de Cassation*, in *Ordre des avocats au Conseil d'État et à la Cour de cassation, La Justice prédictive*, Paris, 2018, p. 15-18.

attività, e quindi la diffusione delle decisioni giudiziarie»²⁰⁵. È probabile quindi che l'esperimento abbia rafforzato la loro convinzione della necessità di mettere in atto sistemi che si adattino ai valori e ai metodi operativi del sistema giudiziario, e che siano di fatto i più adatti a farlo²⁰⁶.

Più di recente, sempre in Francia, il decreto n. 2020-356 del 27 marzo 2020 aveva affidato al Ministero della Giustizia francese il compito di creare un sistema algoritmico, denominato *Datajust*, per finalità predittive di nuovi casi nuovi risarcimento del danno alla persona²⁰⁷. Esso rappresentava il tentativo del Ministero della Giustizia di dotarsi di un proprio strumento di fronte all'aumento dell'offerta di servizi di risarcimento del danno alla persona già offerti da operatori privati²⁰⁸. Anche se risulta che il progetto, ampiamente criticato²⁰⁹, soprattutto per l'entità del compito e le limitate risorse stanziare dal Ministero della Giustizia, sia stato poi abbandonato²¹⁰, è utile delinearne comunque le caratteristiche, ancora presenti sul sito del Ministero della Giustizia francese²¹¹: il progetto mirava a sviluppare un algoritmo incaricato di estrarre automaticamente e utilizzare i dati contenuti nelle decisioni giudiziarie riguardanti il risarcimento dei danni corporei. In particolare, si trattava di censire gli importi richiesti e offerti dalle parti nei procedimenti, le valutazioni proposte nell'ambito delle procedure di risoluzione amichevole delle controversie e gli importi assegnati alle vittime dai tribunali. Attraverso l'elaborazione algoritmica di un *dataset* composto dalle decisioni giudiziarie emesse in appello tra il 1° gennaio 2017 e il 31 dicembre 2019 dalle giurisdizioni amministrative e dalle sezioni civili delle giurisdizioni giudiziarie nei contenziosi riguardanti il risarcimento dei danni alla persona, trasmessi dal Consiglio di Stato

²⁰⁵ B. LOUVEL, *ult. op. cit.*

²⁰⁶ Come sostenuto in L. DUMOULIN, *De quoi la 'justice prédictive' est-elle le nom? Algorithmes, décision et jugement*, cit., p. 224.

²⁰⁷ In dottrina sul progetto si v., in termini critici, M. BEAU, *Décret «DataJust»: stupeur et mécontentements*, in *Le Monde du Droit*, 6 aprile 2020 e F. G'SELL, *Les progrès à petits pas de la «justice prédictive» en France*, in *Era Forum*, 2020, pp. 299 ss.

²⁰⁸ L. DUMOULIN, *De quoi la 'justice prédictive' est-elle le nom ? Algorithmes, décision et jugement*, cit., p. 227.

²⁰⁹ Anche di recente in V. RIVOLLIER, *Datajust. Histoire d'un échec. Séminaire Nouvelles technologies et justice*, Centre internet et société, Marzo 2023.

²¹⁰ O. DUFOUR - A. CORVIAUX, *Datajust: Plutôt que de faire de la justice prédictive, il faut engager une démarche d'indexation et de tri des décisions*, in *Actu-Juridique*, 27 gennaio 2022.

²¹¹ Al seguente indirizzo: <https://www.justice.fr/donnees-personnelles/datajust>.

e dalla Corte di Cassazione al Ministero, l'obiettivo era "predire", utilizzando tecniche di intelligenza artificiale supervisionata, un "benchmark indicativo di risarcimento" destinato alla distribuzione al pubblico e ai giudici.

2.3. Esperimenti di giustizia predittiva in alcune giurisdizioni italiane

Per capire lo stato della giustizia algoritmica (e predittiva) in Italia, è utile muovere dalle ultime relazioni del CSM sullo stato della giustizia telematica nel nostro ordinamento (la prima nel 2021²¹², la seconda di recente adottata (24 luglio 2024²¹³). Già solo a un primo confronto *ictu oculi* delle due relazioni emerge la crescita dell'interesse verso la giustizia algoritmica da parte dell'organo di autogoverno della magistratura: mentre nella relazione del 2021 solo mezza pagina è dedicata all'"intelligenza artificiale e giustizia predittiva", in cui vengono appena accennati alcuni progetti in corso per la valorizzazione del patrimonio giurisprudenziale, nella recente relazione un intero capitolo è dedicato in generale alle banche dati di giurisprudenza di merito con un *focus* specifico sulla giustizia algoritmica e sulla giustizia predittiva²¹⁴. In quest'ultima è inoltre presente una sezione sugli interventi normativi europei e nazionali sul tema²¹⁵ che prenderemo in considerazione nella successiva parte del presente lavoro.

Muovere da quest'ultima relazione ci consente di comprendere la direzione che oggi sta assumendo la giustizia algoritmica nel nostro sistema giudiziario.

Come già detto, gli esperimenti di giustizia predittiva sono una conseguenza dell'implementazione e miglioramento delle banche dati giurisprudenziali, i cui dati sono necessari per il funzionamento dei *software* algoritmici. In questo senso nella relazione si dà conto di numerosi progetti di creazione di archivi di giurisprudenza

²¹² CSM, *Relazione sullo stato della Giustizia telematica* – anno 2021, delibera 21 dicembre 2021.

²¹³ CSM, *Relazione sullo stato della Giustizia telematica* - anno 2024, delibera 24 luglio 2024.

²¹⁴ *Ivi*, pp. 52-57.

²¹⁵ *Ivi*, pp. 65-69.

(di merito) sia territoriali²¹⁶ sia nazionali²¹⁷, portati avanti dai singoli uffici giudiziari, sui cui peraltro il CSM vuole avere l'ultima parola. Essi, infatti, devono «essere monitorati con attenzione dall'organo di governo autonomo, che ne valuterà di volta in volta le importanti ricadute organizzative e l'impatto che possano avere sul lavoro dei Magistrati»²¹⁸.

I progetti locali vanno di pari passo con la creazione di una banca dati della giurisprudenza di merito realizzata in attuazione degli obiettivi PNRR inseriti nella misura “*Riforma 1.8 – Digitalizzazione del Sistema Giudiziario*” e volta ad assicurare la consultazione telematica pubblica dei provvedimenti adottati dai Tribunali e dalle Corti di appello²¹⁹. Tale progetto, secondo il CSM, se «per un verso appare certamente funzionale alla costruzione di un pensiero attivo e consapevole, anche nell'ottica della prevedibilità delle decisioni e del controllo di coerenza da parte dei cittadini del servizio Giustizia, per altro verso, potrebbe anche favorire effetti potenzialmente negativi, facilitando, per mero produttivismo, la passiva ricezione di apparati motivazionali, con conseguente conformismo giudiziario ed impoverimento della stessa qualità della giurisdizione»²²⁰.

Un'ultima serie di progetti, di maggiore interesse per il presente lavoro, è realizzata in coordinamento tra uffici giudiziari e altre istituzioni (soprattutto università) e sono finalizzati a sperimentare l'utilizzo di sistemi di intelligenza artificiale «per acquisire la possibile proiezione giudiziaria del caso proposto, in un'ottica predittiva delle decisioni»²²¹. Anche se non menzionati nella relazione, di seguito segnaleremo i progetti in corso più rilevanti, riservandoci di riportare i

²¹⁶ Tra questi sono valutati con favori gli archivi di giurisprudenza ad uso interno del singolo ufficio, oggi gestite e implementate dagli addetti all'ufficio per il processo (UPP). Tali progetti, tuttavia, si scrive nella relazione, «parrebbero non aver raggiunto l'auspicato livello di utilità in funzione di una piena conoscenza delle decisioni del territorio, poiché si registra, in una valutazione unitaria delle diverse iniziative locali, uno scarso livello di interesse alla fruizione di queste raccolte giurisprudenziali da parte dei magistrati».

²¹⁷ Su cui si v. *infra*.

²¹⁸ Relazione 2024, p. 52.

²¹⁹ Questa comprende sia una banca dati riservata al personale giudiziario (BDR) che si serve anche di sistemi di intelligenza artificiale generativa per supportare le ricerche (generare in tempo reale una sintesi dei provvedimenti ad uso degli utenti o fornire informazioni e risposte alle domande in linguaggio naturale), sia, anche se in una versione con funzionalità più limitate, una banca dati pubblica (BDP) accessibile tramite SPID.

²²⁰ Relazione 2024, p. 54.

²²¹ *Ibidem*.

dubbi espressi dal CSM rispetto agli esperimenti di giustizia predittiva quando si tratterà dei rischi e opportunità costituzionalmente rilevanti di tali progetti.

Occorre comunque premettere che, per la maggior parte dei casi, «si tratta per lo più di iniziative sperimentali e di livello locale in materia di giustizia civile»²²² di cui quindi non si hanno informazioni complete.

Una delle prime iniziative è stata realizzata in collaborazione tra la Corte d'Appello di Brescia, il Tribunale ordinario di Brescia e l'Università degli Studi di Brescia, formalizzata in una convenzione “Giustizia Predittiva” del 6 aprile 2018²²³ e divenuta operativa l'11 novembre 2021²²⁴. L'obiettivo del progetto era di fornire “predizioni” sugli orientamenti in specifiche materie e la loro tempistica. Questo per favorire, si legge nella convenzione, la trasparenza delle decisioni, la circolarità della giurisprudenza tra I e II grado e il superamento di contrasti inconsapevoli²²⁵. Sul sito sono presenti due sezioni (diritto del lavoro e diritto delle imprese) da cui è possibile accedere a sottosezioni divise per argomenti (es. infortuni sul lavoro per la prima e annullabilità delle deliberazioni assembleari per la seconda). Il risultato del processo è una “predizione” (così la chiama il sito) in un determinato caso²²⁶.

Senza voler sminuire l'importanza di progetti di questo tipo per favorire la conoscenza e la circolarità della giurisprudenza, quello di Brescia appare un'esperienza ancora lontana da quello che abbiamo inteso con giustizia predittiva²²⁷ e, anche se non conosciamo nello specifico le modalità di estrazione

²²² E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit. p. 280.

²²³ Reperibile al seguente indirizzo: https://giustizia.brescia.it/allegatinews/A_18592.pdf

²²⁴ Di seguito il link al sito: <https://giustiziapredittiva.unibs.it/#collaboratori>

²²⁵ Fautore del progetto è stato il Presidente della Corte d'Appello di Brescia, Claudio Castelli, da tempo interessato ai temi della digitalizzazione e dell'IA nella giustizia. Si v. C. CASTELLI - D. PIANA, *Giustizia predittiva. La qualità della giustizia in due tempi*, in *Questione giustizia*, trimestrale n. 4, 2018. Sul progetto di Brescia si v. altresì F. GALLI - G. SARTOR, *AI Approaches to Predictive Justice: A Critical Assessment*, in *Humanities And Rights Global Network Journal*, 2023, 5, pp. 165-217; M. LIBERTINI - M. ROSARIA MAUGERI - E. VINCENTI, *Intelligenza artificiale e giurisdizione ordinaria. una ricognizione delle esperienze in corso*, cit., p. 498.

²²⁶ È di tutta evidenza che siamo lontani da quello che intendiamo per predizione. Per fare un esempio tratto dal sito, sotto la sezione Diritto del lavoro-Infortunio sul lavoro-Condotta del datore di lavoro, il sito produce la sua “predizione” in questi termini «SE il datore di lavoro tollera una condotta pericolosa del lavoratore senza vietarla ALLORA è responsabile del danno subito dal lavoratore».

²²⁷ Anche se spesso citato nella letteratura (si v. per esempio LIBERTINI - M. ROSARIA MAUGERI - E. VINCENTI, *Intelligenza artificiale e giurisdizione ordinaria. una ricognizione delle esperienze*

ed elaborazione dei dati giurisprudenziali, non pare nemmeno che siano state utilizzate tecniche di intelligenza artificiale²²⁸: è lo stesso Claudio Castelli, Presidente della Corte d'Appello bresciana e fautore del progetto, che in sede di presentazione dello stesso ha precisato che «l'intenzione a livello locale è di passare dalla fase artigianale in cui ci troviamo, ove la scelta delle sentenze viene operata dalle sezioni di Tribunale e Corte d'Appello e l'elaborazione di casi e abstract viene operata da ricercatori dell'Università, a una fase più evoluta in cui l'estrazione delle sentenze avviene automaticamente dalla banca dati (che contiene tutte le sentenze emesse in formato integrale) attraverso algoritmi e/o parole chiave la cui elaborazione avviene attraverso un processo di *machine learning*»²²⁹.

Progetto più ambizioso ma ancora in corso e con obiettivi a lungo termine è il progetto “*Predictive Justice*”, guidato dal Lider-Lab presso l'Istituto Dirpolis della Scuola Superiore Sant'Anna di Pisa, e che coinvolge i Tribunali di Genova e Pisa. Le informazioni sono ancora frammentate ma dal sito del progetto e dalle dichiarazioni dei protagonisti appare un esperimento particolarmente sofisticato. Gli obiettivi sono molteplici: dalla predisposizione di *software* per automatizzare la pseudonimizzazione dei testi, alla creazione di strumenti innovativi per interrogare i materiali legali attraverso l'annotazione automatica, fino alla costruzione di strumenti predittivi basati sulla scienza dei dati e sull'intelligenza artificiale, passando per il tentativo di offrire spiegazioni comprensibili sul funzionamento degli strumenti utilizzati e adattarli alle diverse esigenze degli utenti finali²³⁰. In particolare, il progetto si sta concentrando sull'annotazione semantica²³¹ di decisioni su danni alla persona e dell'assegno di separazione e divorzio.

in corso, cit., pp. 499 ss.), ancora più embrionale appare il progetto “prevedibilità delle decisioni” portato avanti dalla Corte d'Appello di Bari (https://ca-bari.giustizia.it/it/progetto_prevedibilita.page). Il progetto è iniziato nel 2016 su iniziativa del Presidente della Terza Sezione Civile e ha portato alla creazione manuale di riassunti della giurisprudenza consolidata su argomenti ricorrenti. Sebbene l'iniziativa possa fornire un contributo prezioso per le applicazioni di giustizia predittiva, finora non ha coinvolto l'uso di sistemi automatizzati di IA.

²²⁸ Come confermano anche F. GALLI - G. SARTOR, *AI Approaches to Predictive Justice: A Critical Assessment*, cit., p. 192.

²²⁹ C. CASTELLI, *La giustizia predittiva*, in *Lamagistratura.it*, 11 novembre 2021.

²³⁰ Si v. il sito web del progetto al seguente indirizzo: <https://www.predictivejurisprudence.eu/>

²³¹ Annotazione semantica significa identificare delle espressioni all'interno di una pronuncia che permettano di etichettarla così da poter poi allenare un algoritmo.

L'obiettivo a breve termine è l'identificazione automatizzata di statistiche per la gestione dei flussi di casi, ossia raccomandare ai gestori dei tribunali come allocare efficacemente i casi tra le varie sezioni²³². L'obiettivo dichiarato a lungo termine sarebbe quello di sviluppare modelli predittivi che possano assistere i giudici²³³ o aiutare a valutare i tempi o l'esito di un processo²³⁴.

Sempre in un contesto di collaborazione tra uffici giudiziari e Università, si segnala il progetto "Giurisprudenza Predittiva", promosso dalla Corte d'Appello di Venezia con il Centro di Studi Giuridici di Ca' Foscari insieme con Unioncamere del Veneto, l'Ordine degli Avvocati di Venezia e la società Deloitte (al cui interno vi è un dipartimento di intelligenza artificiale, deputato alla costruzione dell'algoritmo). Si è partiti dalla raccolta e classificazione in una banca dati di 800 sentenze in materia di licenziamento per giusta causa²³⁵ in un triennio per giungere alla costruzione di un algoritmo che permettesse al programma di IA di formulare una domanda in linguaggio naturale e ricercare elementi significativi di interesse (anche per concetti semantici)²³⁶.

Quanto alla giustizia tributaria, grazie ai fondi europei previsti nell'ambito del *Next Generation EU*²³⁷, nel quale si inserisce il PNRR²³⁸, il Ministero dell'Economia e delle Finanze e il Consiglio della Presidenza della Giustizia Tributaria hanno finanziato il progetto PRO.DI.GI.T, tra le altre cose, per la

²³² Riassegnando per esempio al medesimo giudice il contezioso che riguardi questioni simili e agevolare la decisione, anche con riferimento alle tempistiche, si v. M. LIBERTINI - M. ROSARIA MAUGERI - E. VINCENTI, *Intelligenza artificiale e giurisdizione ordinaria. una ricognizione delle esperienze in corso*, cit. p. 502.

²³³ È interessante che, sempre sul sito, rispetto alla progettazione di sistemi predittiva, si dica che «*Although we are convinced that some rare instances, being simple and repetitive, can benefit from certain levels of automation, we remain convinced that human decision making rather than being replaced should be assisted by technology*».

²³⁴ Si cfr. V. MAGLIONE, *Pisa allena l'algoritmo che «prevede» le sentenze*, in *Norme e Tributi Plus*, 2021.

²³⁵ C. CASTELLI, *Giustizia Predittiva: i progetti in corso in Italia*, in *Agenda Digitale*, 2 agosto 2023.

²³⁶ Era possibile porre al programma domande della specie «È possibile licenziare un dipendente per giusta causa in seguito a sue assenze ingiustificate?» e ottenere una risposta sulle sentenze all'interno del perimetro delle 800 sentenze, in particolare quali e quanti i rigetti e accoglimenti. Si cfr. G. MASELLA, «*Giurisprudenza predittiva*». *Risultati operativi e prospettive future*, in *L'esperienza Uni4Justice e le prospettive future Le ricerche del team di Ca' Foscari*, 2023, pp. 279-292.

²³⁷ https://commission.europa.eu/strategy-and-policy/recovery-plan-europe_it

²³⁸ <https://www.governo.it/sites/governo.it/files/PNRR.pdf>

giustizia predittiva in ambito fiscale²³⁹. Progetti più a breve termine riguardano la digitalizzazione della giustizia tributaria, l'implementazione di una banca dati nazionale di giurisprudenza di merito nonché la sommarizzazione delle sentenze tributarie tramite l'intelligenza artificiale (in particolare con i modelli LLM come Chat GPT)²⁴⁰. Sembra che a seguito dello sviluppo del database giurisprudenziale e dell'attività di sommarizzazione delle sentenze tramite tecniche di *machine learning*, l'intento è quello di creare un sofisticato strumento di giustizia predittiva per fornire al giudice uno strumento di supporto ed al pubblico uno strumento di previsione dello sviluppo del contenzioso²⁴¹.

Anche se le notizie non sono certe, sembra che l'obiettivo del MEF sia quello di addestrare l'algoritmo predittivo su un milione di sentenze tributarie²⁴² messe a disposizione di un *team* di informatici e avvocati di diverse università italiane, come l'Università di Bologna, per sviluppare tecniche di intelligenza artificiale²⁴³.

Infine, un ultimo progetto ambizioso e di cui ancora non si conoscono i futuri sviluppi, anch'esso finanziato con i fondi europei, è il *Next Generation UPP*, condotto dall'Università di Torino in partenariato con altri undici atenei italiani del nord-ovest e che coinvolge 35 uffici giudiziari nell'ambito dei distretti di Corte d'Appello di Brescia, Genova, Milano e Torino. Essi si inserisce nel più ampio

²³⁹ Nella presentazione del progetto (<https://www.finanze.gov.it/it/Progetti-europei/PRO.DI.GI.T/Presentazione/>) si legge: «Il PRODIGIT è un progetto finanziato con fondi a carico del Programma PON Governance e Capacità Istituzionale 2014-2020, Asse 7, Obiettivo specifico 7.2, Azione 7.2.1, e prevede una serie di attività per l'innovazione della Giustizia Tributaria, con il supporto della tecnologia digitale e della intelligenza artificiale. Il progetto, a cui partecipano il Dipartimento delle finanze (DF) e il Consiglio di Presidenza della giustizia tributaria (CPGT), ha ottenuto il finanziamento europeo in data 25 maggio 2022, con la stipula tra l'Agenzia per la coesione territoriale e il Dipartimento, quale Beneficiario, della Convenzione per la regolamentazione dei rapporti di attuazione, gestione e controllo relativi al PRODIGIT». In dottrina si v. I. FERRANTI, *Prodigit, come funziona il progetto per la giustizia tributaria digitale*, in *Agenda Digitale*, 10 Marzo 2023.

²⁴⁰ Sui risultati parziali di quest'ultimo progetto si cfr. G. PISANO *et. al.*, *La sommarizzazione delle sentenze tributarie di merito nel Progetto PRODIGIT*, in *Rivista di Scienze Giuridiche, Scienze Cognitive ed Intelligenza Artificiale*, vol. 17 (1), 2024.

²⁴¹ R. BERTI - F. ZUMERLE, *Prodigit: l'intelligenza artificiale che migliora la giustizia tributaria*, in *Agenda Digitale*, 24 luglio 2023.

²⁴² I. CIMMARUSTI, *Un milione di sentenze per l'algoritmo predittivo nel processo tributario*, in *Sole 24 Ore*, 3 ottobre 2022.

²⁴³ F. GALLI - G. SARTOR, *AI Approaches to Predictive Justice: A Critical Assessment*, cit., p. 193.

progetto di estensione e implementazione dell'Ufficio per il Processo²⁴⁴. Negli obiettivi del progetto si specifica che «coerentemente con quanto evidenziato dalla Commissione europea per l'efficienza della giustizia del Consiglio d'Europa (CEPEJ) nel documento *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment*, vuole potenziare il processo di digitalizzazione delle attività giudiziarie come base per l'applicazione di strumenti di Legal Analytics (LA) e creare un ambiente tecnico nel quale far confluire diverse discipline, tra cui data science, intelligenza artificiale (IA), machine learning (ML) e natural language processing (NLP)».

Nell'ambito di tale ampio progetto, un *team* coordinato dal Prof. Cavallo Perin dell'Università di Torino in collaborazione con il CSI Piemonte (Consorzio per il Sistema Informativo) ha avviato la sperimentazione del progetto “Laboratorio Sentenze”²⁴⁵. Tale sperimentazione ha preso le mosse dalla condivisione con il CSI Piemonte di sentenze in materia di servizi pubblici (sembra circa mille sentenze in materia di servizi pubblici rese dal Consiglio di Stato tra il 2016 e il 2023)²⁴⁶ e mira alla progettazione di un algoritmo di analisi testuale delle sentenze da inserire in un *machine learning*, che individui l'esito di precedenti giudizi relativi a casi simili a quello in ordine al quale il giudice è chiamato a decidere²⁴⁷.

Ora, come visto, i progetti in corso in Italia sono ancora lontani da quello che abbiamo definito “giustizia predittiva” e non paiono comunque proiettarsi ancora al di fuori di una dimensione territoriale e di singole sperimentazioni: insomma non sembrerebbe che nei prossimi anni la giustizia algoritmica si porrà in testa alle

²⁴⁴ Progetto unitario per la diffusione dell'Ufficio per il Processo e l'implementazione di modelli operativi innovativi negli Uffici giudiziari per lo smaltimento dell'arretrato, promosso dal Ministero della Giustizia nell'ambito del PON Governance e Capacità Istituzionale 2014-2020 e realizzato in sinergia con gli interventi previsti dal Piano Nazionale di Ripresa e Resilienza (PNRR) a sostegno della riforma della giustizia. Si cfr. <https://www.nextgenerationupp.unito.it/home>.

²⁴⁵ Il progetto è stato presentato in occasione di un convegno tenutosi presso l'Università di Torino il 19 ottobre 2023 (<https://aldricus.giustizia.it/2023/10/17/whats-next-laboratorio-di-sentenze-un-esperimento-di-intelligenza-artificiale/>).

²⁴⁶ V. MAGLIONE, *Atenei e giustizia predittiva, in campo l'intelligenza artificiale*, in *Il Sole* 24 ore, 30 ottobre 2023.

²⁴⁷ Si v. la descrizione del progetto alla pagina web rinvenibile al seguente indirizzo: <https://www.nextgenerationupp.unito.it/eventi-e-news/laboratorio-sentenze>.

priorità del Paese²⁴⁸. E, ciò nonostante, i singoli progetti già in corso in Italia e le esperienze menzionate degli altri paesi ci mostrano già quali potrebbero essere le domande che il costituzionalista deve porsi, e che saranno approfondite nel capitolo VIII. L'urgenza di affrontare tali problematiche è avvertita anche dal Procuratore Generale presso la Corte di Cassazione Luigi Salvato che nell'intervento sull'amministrazione della giustizia nell'anno 2023 ha affermato che «in nome della prevedibilità e della velocità si invocano ulteriori sviluppi, la giustizia predittiva, affidata all'intelligenza artificiale. Questa non va aprioristicamente rifiutata, occorre sfruttarne le potenzialità, ma dobbiamo essere consapevoli che è qualcosa di radicalmente diverso da ogni precedente scoperta dell'uomo. È una tecnologia che plasma e diffonde forme non umane di logica; gli algoritmi di machine learning non sempre sono trasparenti, spiegabili o interpretabili, soprattutto se utilizzano tecniche di deep learning. Alto è il rischio della lesione dei diritti fondamentali e dell'alterazione dell'essenza del processo; alta deve essere attenzione e prudenza nell'applicarla»²⁴⁹.

²⁴⁸ Ne è dimostrazione il fatto che nella recente Strategia Italiana per l'Intelligenza Artificiale 2024-2026 (<https://assets.innovazione.gov.it/1721376223-01-strategia-italiana-per-l-intelligenza-artificiale-2024-2026.pdf>), eccetto un cenno en passant del progetto PRODIGIT, non si fa alcun riferimento alla giustizia algoritmica.

²⁴⁹ Procura Generale della Corte di Cassazione, Luigi Salvato, *Intervento sull'amministrazione della giustizia nell'anno 2023*, 25 gennaio 2024.

LE PROSPETTIVE FUTURIBILI DELLA GIUSTIZIA ALGORITMICA

SOMMARIO: 1. L'avvento dell'IA generativa nel mercato dell'intelligenza artificiale - 2. L'impiego di strumenti di Generative AI da parte dei giudici - 3. Brevi riflessioni conclusive: esperienze diverse, diversi interrogativi

1. L'avvento dell'IA generativa nel mercato dell'intelligenza artificiale

L'approccio induttivo segnalato nel capitolo introduttivo ci obbliga a dare conto di una nuova possibile forma di giustizia algoritmica che, sebbene in modo ancora sporadico, ha già fatto ingresso in alcune aule giudiziarie. Si tratta della messa sul mercato della cd. *Generative Artificial Intelligence (GAI)*²⁵⁰, soprattutto quelli che si servono di *Large Language Models (LLM)*²⁵¹, di cui fa parte ChatGPT²⁵²,

²⁵⁰ Si v. la definizione fornita sul sito <https://generativeai.net/> (traduzione nostra): «l'Intelligenza Artificiale Generativa (GenAI) è un tipo di Intelligenza Artificiale capace di creare una vasta gamma di dati, come immagini, video, audio, testo e modelli 3D. Lo fa imparando modelli dai dati esistenti e utilizzando questa conoscenza per generare nuovi output unici. La GenAI è in grado di produrre contenuti altamente realistici e complessi che imitano la creatività umana, rendendola uno strumento prezioso per molti settori come il gaming, l'intrattenimento e il design di prodotti. Le recenti scoperte nel campo, come GPT (*Generative Pre-trained Transformer*) e *Midjourney*, hanno significativamente avanzato le capacità della GenAI. Questi progressi hanno aperto nuove possibilità per l'utilizzo della GenAI nella risoluzione di problemi complessi, nella creazione di arte e persino nell'assistenza alla ricerca scientifica».

Esempi tipici di sistemi di intelligenza artificiale generativa includono generatori di immagini (come *Midjourney* o *Stable Diffusion*), modelli linguistici di grandi dimensioni (come GPT-4, PaLM o Claude), strumenti per la generazione di codice (come *Copilot*) o strumenti per la generazione audio (come VALL-E o *resemble.ai*).

²⁵¹ I modelli linguistici grandi dimensioni (LLM) sono modelli computazionali che hanno la capacità di comprendere e generare il linguaggio umano. Gli LLM hanno la capacità trasformativa di prevedere la probabilità di sequenze di parole o generare nuovo testo basato su un *input* dato. Si v. Y. CHANG - X. WANG - J. WANG *et al.*, *A survey on evaluation of large language models*, in *ACM Transactions on Intelligent Systems and Technology*, 2018. Gli LLM sono un'evoluzione del NLP. Gli LLM di ultima generazione sono basati su una rete neurale artificiale chiamata *General Pre-trained Transformer (GPT)* che è allenata su database di dati molto grandi ed è in grado di cogliere sfumature del linguaggio e generare un testo coerente e consapevole. Si v. E. LONGO, *La "giustizia digitale" tra nuove forme di efficienza e garanzia dei diritti costituzionali*, cit., p. 189 e T. BROWN - B. MANN - N. RYDER *et al.*, *Language models are few-shot learners*, in *Advances in neural information processing systems*, 2020, pp. 1877-1901.

²⁵² ChatGPT è stato lanciato sul mercato nel novembre del 2022 da OpenAI. Cfr. OPENAI, *ChatGPT: Optimizing Language Models for Dialogue*, 30 novembre 2022, disponibile all'indirizzo <https://openai.com/blog/chatgpt>.

lanciata sul mercato il 30 novembre 2022. Essa è addestrata notoriamente su una vastissima gamma di dati raccolti dal web (tramite il cd. *scraping*) e, tramite tecniche avanzate di *machine learning* di elaborazione del linguaggio naturale²⁵³, il modello è in grado di generare *output* di diversi tipi, tra testo, immagini, codici ecc... Per esempio, nel caso di generazione di un testo il modello utilizza le informazioni apprese durante l'addestramento per comprendere il contesto e generare una risposta coerente, considerando le parole precedenti in sequenza e, sulla loro base, selezionando la parola più probabile rispetto all'addestramento ricevuto²⁵⁴.

I dettagli specifici del *dataset*, a discapito del nome della società produttrice, non sono conoscibili in quanto informazioni proprietarie di *OpenAI*²⁵⁵.

Non serve qui raccontare l'impatto che ha avuto l'avvento dei modelli come ChatGPT non solo sul *mercato* dell'intelligenza artificiale – ChatGPT ha innescato una corsa, da parte dei grandi attori tecnologici, allo sviluppo di sempre più sofisticati di sistemi LLM – ma anche sul *discorso* sull'intelligenza artificiale. Se prima era un argomento spesso ignorato dal grande pubblico o ritenuto – erroneamente – circoscritto a pochi determinati settori, avulsi dal quotidiano delle persone, con l'avvento dell'IA generativa l'intelligenza artificiale è arrivata “sulla bocca di tutti”.

Non solo, come vedremo, l'ascesa improvvisa di strumenti di IA generative *general purpose*²⁵⁶, termine che si riferisce a sistemi di intelligenza artificiale capaci di svolgere una vasta gamma di compiti senza essere limitati a uno scopo specifico o predeterminato, ha rallentato i tempi dell'approvazione dell'*AI Act*, in

²⁵³ F. GALLI - G. SARTOR, *AI Approaches to Predictive Justice: A Critical Assessment*, cit., p. 176.

²⁵⁴ M. ABOUFOUL, *Despite Their Feats, Large Language Models Still Haven't Contributed to Linguistics. A review of Chomsky's views on linguistics and LLMs*, in *Towards Data Science*, 5 dicembre 2022.

²⁵⁵ Sull'impiego dei dati da parte di OpenAI è nota la causa intentata dal New York Times contro OpenAI per la violazione del *copyright*, in quanto l'azienda tecnologica avrebbe utilizzato illegalmente i dati protetti della testata per allenare gli algoritmi. Si cfr. M.M. GRYNBAUM - R. MAC, *The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work*, in *New York Times*, 27 dicembre 2023. Della problematica della trasparenza algoritmica si tratterà *infra* al capitolo VI.

²⁵⁶ G. OLIVATO, *Paving the path towards general purpose AI systems regulation in the AI Act: an analysis of the Parliament's and Council's proposals*, 28 febbraio 2024.

quanto il legislatore europeo ha voluto disciplinare anche strumenti come ChatGPT (inserendo un autonomo capo)²⁵⁷.

2. L'impiego di strumenti di *Generative AI* da parte dei giudici

Anche se è possibile ricomprendere l'impiego di ChatGPT da parte dei giudici nell'ombrello della giustizia predittiva, a nostro avviso, ricordata l'accezione ristretta assunta nel presente lavoro, siamo di fronte a una diversa forma di giustizia algoritmica potenzialmente influente sul processo decisionale del giudice²⁵⁸: un conto è infatti fornire il giudice di uno strumento che, sulla base di una quantità immensa di precedenti, gli suggerisca l'orientamento prevalente e il possibile esito della controversia; altro è fornirgli uno strumento in grado di scrivere direttamente la decisione sulla base degli elementi di fatto e dei quesiti rivolti dal giudice al sistema algoritmico. In questo senso parte della dottrina, pur ricomprendendo la GAI nella giustizia predittiva²⁵⁹, ritiene che con l'IA generativa siamo di fronte a «un vero e proprio “salto evolutivo” nell'impiego delle tecniche di giustizia predittiva sin qui conosciute» in quanto «la macchina assorbe l'intero processo decisionale e genera direttamente il testo, vale a dire il mezzo attraverso cui si manifesta la volontà»²⁶⁰.

Tale utilizzo non è solo ipotetico ma si sono già verificati dei casi degni di nota in alcune giurisdizioni.

Quello più interessante si è svolto in un tribunale del lavoro di Cartagena, in Colombia, dove il 31 gennaio 2023 il giudice Juan Manuel Padilla ha emesso una

²⁵⁷ Si cfr. G. LO SAPIO, *L'Artificial Intelligence Act e la prova di resistenza per la legalità algoritmica*, in *Federalismi*, n. 16/2024, pp. 265-290. Sulla regolazione di strumenti di *generative AI* nel contesto europeo si v. C. NOVELLI - F. CASOLARI - P. HACKER - G. SPEDICATO - L. FLORIDI, *Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity*, Working Paper, 15 marzo 2024.

²⁵⁸ Secondo la definizione offerta in questo lavoro, in coerenza con l'*AI Act*.

²⁵⁹ Come fa per esempio A. SIMONCINI, *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*, cit. L'A. opportunamente distingue tra giustizia predittiva “descrittiva”, utilizzata per prevedere l'esito di decisioni giudiziarie, e giustizia predittiva “prescrittiva”, come i *tools* impiegati negli Stati Uniti o la scrittura di sentenze tramite l'ausilio di strumenti di IA generativa.

²⁶⁰ A. SIMONCINI, *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*, cit., p. 402.

sentenza di sette pagine²⁶¹ su un caso in cui era in gioco il diritto fondamentale alla salute di un bambino a cui era stato diagnosticato un disturbo dello spettro autistico. Si trattava di un caso di appello relativamente semplice in cui la questione legale chiave era se la richiesta di una compagnia di assicurazione sanitaria di un pagamento per autorizzare una procedura medica violasse il diritto fondamentale alla salute del bambino e il diritto a una vita dignitosa (*salud y vida digna*). Il giudice Padilla ha confermato la sentenza di primo grado che aveva riconosciuto il diritto del bambino e l'esonero dai pagamenti richiesti²⁶².

La particolarità di tale sentenza è che il giudice estensore, dopo aver riportato i fatti, le norme applicabili e la giurisprudenza costituzionale pertinente al caso, per motivare la propria decisione – in conformità, secondo il giudice, alla Ley 2213 de 2022²⁶³ il cui art. 2 prevede che «*se podrán utilizar las tecnologías de la información y de las comunicaciones, cuando se disponga de los mismos de manera idónea, en la gestión y trámite de los procesos judiciales y asuntos en curso, con el fin de facilitar y agilizar el acceso a la justicia*» – ha riportato la “conversazione” intercorsa con ChatGPT e le domande rivolte al programma²⁶⁴.

In realtà, le risposte di ChatGPT sono state molto generiche²⁶⁵ e sembrano essere solo una conferma della decisione già prospettata dal giudice nelle parti precedenti

²⁶¹ La sentenza è disponibile al seguente indirizzo: <https://forogpp.com/wp-content/uploads/2023/01/sentencia-tutela-segunda-instancia-rad.-13001410500420220045901.pdf>

²⁶² Sul caso si v., nella dottrina colombiana, J.D. GUTIÉRREZ, *ChatGPT in Colombian Courts: Why we need to have a conversation about the digital literacy of the judiciary*, in *VerfBlog*, 23 febbraio 2023, <https://verfassungsblog.de/colombian-chatgpt>. In Italia riferisce compiutamente del caso E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit, pp. 13 ss. nonché R. PERONA, *ChatGPT e decisione giudiziale: per un primo commento alla recente sentenza del Juzgado Primero Laboral di Cartagena de Indias (Colombia)*, in *Diritti Comparati*, 21 febbraio 2023.

²⁶³ Ley 2213 de 2022 «*Por medio de la cual se establece la vigencia permanente del decreto legislativo 806 de 2020 y se adoptan medidas para implementar las tecnologías de la información y las comunicaciones en las actuaciones judiciales, agilizar los procesos judiciales y flexibilizar la atención a los usuarios del servicio de justicia y se dictan otras disposiciones*».

²⁶⁴ Le domande rivolte sono le seguenti: «*¿Menor autista esta exonerado de pagar cuotas moderadoras en sus terapias?; ¿Las acciones de tutela en estos casos se deben conceder? ¿Exigir en estos casos la cuota moderadora es una barrera de acceso al servicio de salud?; ¿La jurisprudencia de la corte constitucional ha tomado decisiones favorables en casos similares?»*.

²⁶⁵ Per esempio, alla domanda se la Corte costituzionale avesse già assunto decisioni favorevoli in casi simili, ChatGPT non cita precedenti ma si limita a specificare che «*Sí, es cierto. La Corte Constitucional de Colombia ha tomado decisiones favorables en casos similares relacionados con la exoneración de cuotas moderadoras para personas con diagnóstico de autismo. La jurisprudencia ha reconocido la importancia de garantizar el acceso a servicios de salud esenciales*».

della decisione. È lo stesso giudice che, interrogandosi sull'utilità di ChatGPT, afferma che per la risoluzione del caso ha «aggiunto» (*adicionar*)²⁶⁶ le ragioni prodotte dal sistema di IA, impiegata come «iniziativa per accelerare la risoluzione dei casi» e che l'inclusione di tali sistemi di generatori di testo non mira a «*reemplazar la decisión del Juez. Lo que realmente buscamos es optimizar los tiempos empleados en redacción de sentencias, previa corroboración de la información suministrada por IA*»²⁶⁷.

Pochi giorni dopo un'altra giudice colombiano, del Tribunale Amministrativo di Magdalen ha emesso un'ordinanza²⁶⁸ in un caso che riguardava una richiesta di risarcimento di danno da parte di un'azienda che aveva in appalto un servizio della Polizia Nazionale colombiana e dove parte attrice aveva richiesto – richiesta accettata dalla Polizia convenuta – che la prima udienza si svolgesse in un'aula virtuale nel metaverso. Anche in questo caso nell'ordinanza sono stati trascritti i quesiti e le risposte di ChatGPT che miravano a rispondere a domande tecniche che hanno aiutato la giudice a decidere come condurre un'udienza giudiziaria nel metaverso²⁶⁹. Il 15 febbraio, l'udienza si è effettivamente svolta attraverso Horizon Workrooms di Meta ed è stata trasmessa in diretta su YouTube²⁷⁰.

para las personas con autismo, y ha sostenido que la exigencia de pagar cuotas moderadoras puede ser una barrera de acceso a estos servicios». Su questo v. ancora E. LONGO, Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale, cit. pp. 14-15.

²⁶⁶ Sull'uso, in questo caso, di ChatGPT come argomento *ad abundantiam*, si v. A. SIMONCINI, *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*, cit., p. 401.

²⁶⁷ Traduzione nostra: «L'obiettivo di includere questi testi generati dall'IA non è in alcun modo quello di sostituire la decisione del Giudice. Quello che realmente cerchiamo è ottimizzare i tempi impiegati nella redazione delle sentenze, previa corroborazione delle informazioni fornite dall'IA».

²⁶⁸ Disponibile al seguente indirizzo: <https://forogpp.com/wp-content/uploads/2023/02/2020-014-siett-vs-nacion-policia-nacional-solicitud-audiencia-en-el-metaverso-1.pdf>.

²⁶⁹ Queste le domande rivolte a ChatGPT: «¿Qué es un avatar?»; «¿Cuál es el método más efectivo para verificar la autenticidad de quienes se conectan a una reunión y/o audiencia virtual?»; «¿Método de verificación de autenticidad de un avatar en el metaverso?».

²⁷⁰ L'udienza è liberamente visionabile al seguente link: <https://www.youtube.com/watch?v=LXi2TX9OBmQ>. Relativamente a questo caso si cfr. ancora Sul punto sia consentito richiamare E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit., pp. 13 ss. e J.D. GUTIÉRREZ, *ChatGPT in Colombian Courts: Why we need to have a conversation about the digital literacy of the judiciary*, cit.

Mentre dottrina e stampa specializzata hanno segnalato altri casi simili avvenuti in Sudamerica²⁷¹, anche in Europa si ha notizia dei primi impieghi di ChatGPT da parte dei giudici per la redazione di parti della sentenza: un giudice della Corte d'Appello del Regno Unito ha ammesso di aver utilizzato ChatGPT per riassumere una parte del *diritto* di una sentenza e poi ha copiato e incollato la risposta²⁷². Se da un lato il giudice ha elogiato l'efficienza dello strumento, definito «*jolly useful*», dall'altro ha precisato che conosceva già la risposta ed era quindi in grado di giudicare l'*output* dell'algoritmo, sottolineando l'importanza di non affidarsi all'intelligenza artificiale in caso di argomenti poco conosciuti²⁷³.

3. Brevi riflessioni conclusive: esperienze diverse, diversi interrogativi

La finalità del capitolo è stata quella di presentare gli sviluppi attuali e ragionevolmente prevedibili della giustizia algoritmica, nell'ambito del perimetro circoscritto che abbiamo indicato nel capitolo introduttivo. Come abbiamo detto, il *modus procedendi* che ci pare più adeguato al fine di trattare da una prospettiva costituzionale un tema così variegato è quello di adottare un approccio empirico e induttivo, in modo da riuscire a «conoscere e distinguere... resistendo alla tentazione di pronunciare giudizi generalizzati o di cedere a logiche “faziose”, piuttosto che alla forza delle argomentazioni razionali»²⁷⁴.

Le diverse esperienze segnalate, dai *risk assessment tools* impiegati negli Stati Uniti, alle esperienze ancora embrionali di “giustizia predittiva” sino ai primi

²⁷¹ J.D. GUTIÉRREZ, *Judges and Magistrates in Peru and Mexico Have ChatGPT Fever*, in *Tech Policy Press*, 19 aprile 2023.

²⁷² La notizia è apparsa su diverse testate online: <https://www.astonbond.co.uk/british-judge-used-chat-gpt-for-case-ruling/> e <https://www.theguardian.com/technology/2023/sep/15/court-of-appeal-judge-praises-jolly-useful-chatgpt-after-asking-it-for-legal-summary>.

²⁷³ Sul punto una recente guida redatta dagli organi di vertice della giustizia dell'Inghilterra e Galles (COURTS AND TRIBUNALS JUDICIARY, *Artificial Intelligence (AI). Guidance for Judicial Office Holders*, 12 dicembre 2023) ha riconosciuto che i giudici possono utilizzare le IA (anche generative), nel loro lavoro preparatorio e doverlo necessariamente dichiarare, ma rispettando una serie di cautele e accorgimenti relativi all'accuratezza dei dati, la responsabilità dell'impiego e alla privacy. Su tale documenti si veda, diffusamente, E. LONGO, *La “giustizia digitale” tra nuove forme di efficienza e garanzia dei diritti costituzionali*, cit., pp. 197 ss.

²⁷⁴ A. SIMONCINI, *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*, cit., p. 419.

esperimenti “amatoriali” di utilizzo di intelligenza artificiale generativa da parte del giudice pongono tutte degli interrogativi di natura costituzionale e di compatibilità con i principi del giusto processo che meritano attenzione. La base di partenza è che sono tutti strumenti che rientrano (alcuni lo hanno già dimostrato, altri ne hanno le potenzialità) nell’ambito di quei sistemi algoritmici che sono in grado di avere un’influenza sul processo decisionale del giudice. Tale denominatore comune ci consentirà di trattare alcune delle problematiche sotto una medesima lente di ingrandimento poiché le caratteristiche di tali algoritmi e la relazione che creano con la decisione giudiziaria sono spesso assimilabili. Ciò nonostante, l’accurata distinzione e classificazione dei diversi sviluppi di giustizia algoritmica in Europa e negli Stati Uniti ci consentirà altresì di maturare consapevolezza delle *differenze* rispetto ai rischi e le opportunità che presentano i sistemi algoritmici, così da non rischiare di riavvolgere fili di diverso colore in un indistinto gomitolo di lana.

La necessaria differenziazione dei rischi che le diverse applicazioni algoritmiche pongono rispetto alle garanzie costituzionali, peraltro, è riconosciuta nella stessa Carta Etica, nella cui Appendice II si distinguono gli utilizzi dell’IA che sono «*to be encouraged*», quelli che richiederebbero «*considerable methodological precautions*», quelli che sono da considerare solo quando saranno disponibili «*additional scientific studies*» e, infine gli utilizzi «*to be considered with the most extreme reservations*»²⁷⁵. Per quanto qui interessa, i sistemi predittivi impiegati nel processo penale appartengono all’ultima categoria (utilizzi da esaminare con le più estreme riserve), in quanto hanno già mostrato potenzialità discriminatorie (vedremo che in realtà i problemi del loro impiego vanno molto oltre l’idoneità a discriminare gli imputati).

Le applicazioni di giustizia predittiva, invece, si muovono in due categorie diverse previste nella Carta Etica a seconda delle prescrizioni che regolano il loro utilizzo. Negli utilizzi che richiedono maggiori studi scientifici rientrano quelle applicazioni che mirano a predire il risultato di una controversia in quanto non sono «*any more efficient because they are still limited by bias in the data sample*

²⁷⁵ Carta Etica, cit. pp. 63-67.

that they have processed and have to start again from square one if a law is amended or if there is a reversal in case law». Qualora però un ordinamento richiedesse al giudice non solo di trovare ausilio negli strumenti di giustizia predittiva ma di conformarsi alla predizione dell'esito della controversia, saremmo di fronte a quello che la Carta chiama «*Quantity-base norm*» cioè una massa di precedenti che produce un esito che il giudice segue come fosse una norma giuridica. Questo approccio creerebbe secondo la CEPEJ pericoli della cristallizzazione della giurisprudenza e i suoi effetti potenzialmente negativi sull'imparzialità e l'indipendenza dei giudici.

Ecco, la differenziazione fornita dalla Carta costituisce già un buon punto di partenza per interrogarsi sui limiti costituzionali all'impiego dell'IA nel processo, che sono diversi a seconda delle applicazioni in gioco. Tali limiti verranno approfonditi diffusamente nella terza parte del lavoro. Tuttavia, prima di costituzionalizzarli, occorre aggiungere un ultimo mattoncino alla costruzione fino a qui presentata, imprescindibile per avere un quadro completo della giustizia algoritmica oggi. Si tratta, banalmente, del dato positivo: lo sviluppo della giustizia algoritmica, più in Europa che negli Stati Uniti – come vedremo – da qualche anno è sotto l'occhio attento delle istituzioni e dei legislatori e nel prossimo capitolo tenteremo di fornire un quadro completo delle strategie di regolazione che ad oggi circondano l'intelligenza artificiale e il suo impiego nel settore della giustizia.

PARTE II

REGOLARE LA GIUSTIZIA ALGORITMICA: I DIVERSI APPROCCI DI STATI UNITI ED EUROPA

Nota introduttiva: perché regolare l'intelligenza artificiale...

Il presente capitolo intende percorrere i diversi approcci strategici, regolatori, e pratico-applicativi della giustizia algoritmica negli Stati Uniti e nell'ordinamento europeo.

È appena il caso di chiarire che non è possibile isolare il fenomeno della giustizia algoritmica dal contesto generale dell'intelligenza artificiale, sicché prima di riflettere diffusamente sulla prima, occorre dare conto dell'impianto regolatorio del fenomeno visto nel suo complesso.

Ancor prima, peraltro, occorre fare luce sulle ragioni per le quali, a differenza degli Stati Uniti, nel nostro continente la produzione normativa (in senso ampio) verso il fenomeno dell'intelligenza artificiale è, ad oggi, la più estesa e completa al mondo.

E ciò risulta essere ancora più interessante se si considera che l'industria dell'intelligenza artificiale in Europa non rappresenta ancora, come già visto, un settore particolarmente sviluppato, soprattutto se rapportato con gli investimenti massicci che negli ultimi anni sono stati fatti negli Stati Uniti e in Cina²⁷⁶. Già è divenuto chiaro – e nel corso di queste pagine si comprenderanno le ragioni – che regolare l'intelligenza artificiale è impresa ardua, perché gli ostacoli che si frappongono sono difficili da superare: siamo di fronte a un fenomeno “nuovo”, di cui ancora non conosciamo le potenzialità; è un fenomeno che corre veloce, che si

²⁷⁶ I dati aggregati che registrano gli investimenti privati in intelligenza artificiale dal 2013 dicono che gli Stati Uniti sono in testa con 335,2 miliardi di dollari investiti, seguiti dalla Cina con 103,7 miliardi di dollari e dal Regno Unito con 22,3 miliardi di dollari, dati *Report* dell'Università di Standford, 2024, cit. *Chapter* 4.3.

evolve a grande rapidità; il diritto sconta una fisiologica lentezza, trovandosi giocoforza a dover inseguire l'evoluzione.

Inoltre, anche solo il livello della regolazione è un tema delicato: come giustamente sottolineato, al pari della regolazione di internet risulta complesso regolare la tecnologia – fenomeno naturalmente globale e con interconnessioni restie a confinarsi all'interno di confini domestici – sulla base di regole e normative nazionali²⁷⁷. Nel nostro continente si è scelto di regolare l'intelligenza artificiale a livello europeo ma con ambizioni globali, in Unione europea tramite l'adozione in ultimo dell'*AI Act* mentre nel Consiglio d'Europa con la stipula del primo trattato sull'intelligenza artificiale, aperto alla firma di tutti gli Stati interessati anche extraeuropei. Come vedremo, l'azione normativa europea, soprattutto sul versante eurounitario, si muove su binari delicati e di difficile intersezione, avendo a cuore sia il rafforzamento del ruolo competitivo dell'Unione in un campo strategico come quello dell'IA, sia l'esigenza di allineare l'IA ai valori europei e al rispetto dei diritti fondamentali.

Al fondo vi è un dibattito di ampio respiro – non solo pratico ma anche filosofico²⁷⁸ – non ancora sopito tra due differenti posizioni che sono così sintetizzabili: da una parte vi è chi sostiene che una regolazione rigida dell'IA in un'industria non ancora sviluppata avrebbe l'effetto di disincentivare gli investimenti e frenare l'evoluzione tecnologica e che inoltre una regolazione flessibile e maneggevole sarebbe più idonea a regolare un fenomeno che, per sua natura, è in costante e rapidissima evoluzione²⁷⁹; dall'altra vi è chi sostiene, invece, che un cambiamento tanto sfaccettato e potenzialmente dirompente deve, ancor prima di affermarsi, essere scortato da una cornice giuridica chiara che ne assicuri il rispetto dei diritti e dei valori condivisi²⁸⁰.

²⁷⁷ D. R. JOHNSON - D. POST, *Law and Borders: The Rise of Law in Cyberspace*, in *Stanford Law Review*, 48, 5, 1996, p. 1371.

²⁷⁸ L. FLORIDI, *The European Legislation on AI: A Brief Analysis of its Philosophical Approach*, in *Philos. Technol.*, 34, 2021, pp. 215-222.

²⁷⁹ N. GUNNINGHAM - J. REES, *Industry Self-Regulation: An Institutional Perspective*, in *Law & Policy*, 19, 1997, p. 366.

²⁸⁰ Per una ricostruzione di tali posizioni si v. J.F. WEAVER, *Regulation of artificial intelligence in the United States*, in W. BARFIELD - U. PAGALLO (a cura di), *Research Handbook on the Law of Artificial Intelligence*, 2018, pp. 162-178. G. MOBILIO, *L'intelligenza artificiale e i rischi di una*

Si tratta del cd. *dilemma di Collingridge*, che pone sulla bilancia, in delicato equilibrio, due distinte questioni. La prima è quella appena segnalata: la tecnologia segue un'evoluzione così veloce che è complicato prevederne gli effetti nel lungo periodo. Ciò conduce alla seconda questione, al centro del dilemma: se non si riescono a prevedere gli sviluppi non è possibile oggi mitigare i rischi di qualcosa che non si conosce, ma se non si interviene si rischia di non riuscire più a intervenire quando i rischi diventano attuali²⁸¹. In realtà, come fatto notare da alcuni autori²⁸², è spesso proprio la presenza di una cornice giuridica chiara che aumenta la fiducia e favorisce l'adozione della tecnologia, consentendo agli attori in gioco di agire in un ambiente controllato e dove i confini sono normativamente stabiliti.

Di più: come visto lo sviluppo della tecnologia e dell'intelligenza artificiale ad oggi è in mano soprattutto a soggetti privati che stanno assumendo un potere economico (e politico) senza precedenti²⁸³: è quindi doveroso chiedersi quale sia in questo contesto il ruolo della regolazione pubblica per limitare tale potere e allinearlo con il rispetto dei diritti fondamentali²⁸⁴.

È del tutto ovvio, quindi, che l'una e l'altra posizione dipendono massivamente dalla scelta del bilanciamento dei valori in gioco, che sono sostanzialmente le esigenze del mercato da una parte e il rispetto dei diritti fondamentali dall'altra²⁸⁵. La scelta di posizionamento rispetto a tali interessi conduce poi a strategie

“*disruption*” della regolamentazione giuridica, in *BioLaw Journal – Rivista di BioDiritto*, n. 2/2020, pp. 401-424.

²⁸¹ Il dilemma è stato teorizzato in D. COLLINGRIDGE, *The Social Control of Technology*, New York, 1980.

²⁸² J.F. WEAVER, *Regulation of artificial intelligence in the United States*, cit. pp. 155-212.

²⁸³ Si cfr. M. BETZU, *I poteri privati nella società digitale: oligopoli e antitrust*, in *Rivista di diritto pubblico*, 3, 2021, pp. 739-760.

²⁸⁴ Sul tema v. O. POLLICINO - G. DE GREGORIO, *Constitutional Law in the Algorithmic Society*, in H.-W. MICKLITZ - O. POLLICINO - A. REICHMAN - A. SIMONCINI - G. SARTOR - G. DE GREGORIO (a cura di), *Constitutional Challenges in the Algorithmic Society*, Cambridge University Press, 2022. Sulla regolazione tra pubblico e privato nel contesto di Internet si cfr., *ex multis*, A. IANNOTTI DELLA VALLE, *Le regole di internet tra poteri pubblici e privati. Tutela dei diritti e ruolo dell'antitrust in una prospettiva costituzionale*, Editoriale Scientifica, Napoli, 2023. Ancora sui rapporti tra pubblico e privato e sulla valorizzazione del ruolo costituzionale della tutela della concorrenza si cfr. M. BETZU, *I poteri privati nella società digitale: oligopoli e antitrust*, in *Rivista di diritto pubblico*, 3, 2021, pp. 739-760, nonché ID, *poteri pubblici e poteri privati nel mondo digitale*, in *Rivista Gruppo di Pisa*, 2, 2021, pp. 166-191.

²⁸⁵ Si cfr. C. SCHEPISI, *Le “dimensioni” della regolazione dell'intelligenza artificiale nella proposta di regolamento della Commissione*, in *I Post di AISDUE*, IV, 16, 28 marzo 2022, p. 331.

regolatorie diverse, che vanno dalla *self-regulation* alla *co-regulation* fino alla *hard-regulation*: come vedremo, gli Stati Uniti – non sorprendentemente²⁸⁶ – fino all’ultima Amministrazione Biden hanno adottato un totale approccio *laissez-faire* disinteressandosi totalmente alla regolazione dell’intelligenza artificiale, lasciando mano libera sia alle amministrazioni pubbliche sia alle imprese private. La situazione mentre si scrive non è sostanzialmente mutata, se non per dei primi timidi aneliti di co-regolamentazione (solo però verso l’utilizzo dell’IA da parte di agenzie pubbliche) attraverso un *Executive Order* del Presidente Biden (si v. *infra* § 2.2.).

La strada percorsa nell’Unione europea, come si vedrà nei paragrafi successivi è stata invece diversa: già dal 2019 la Commissione Europea ha maturato i primi documenti di *soft law* per orientare le strategie di *Governance* dell’intelligenza artificiale fino ad arrivare a una proposta di regolamento nell’aprile del 2021, diventata legge il 12 luglio 2024. Il Regolamento, nella ricerca di un contemperamento tra la protezione dei diritti e l’obiettivo della crescita economica e l’armonizzazione delle norme per la tenuta del mercato unico, ha adottato il cd. approccio basato sul rischio (*risk-based approach*), che consente sia la tutela dei diritti fondamentali sia la possibilità per i produttori dell’intelligenza artificiale di internalizzare come costo di produzione il rispetto della normativa e così di evitare eventuali violazioni della legge *ex post*²⁸⁷. Insomma, sembra che in Europa si sia adottato, a nostro avviso opportunamente, un approccio precauzionale, la cui opportunità era già stata sollevata dalla dottrina italiana che, prendendo in prestito un principio tipico del diritto dell’ambiente²⁸⁸, ha parlato di «precauzione costituzionale» per indicare segnatamente che l’incertezza conoscitiva cui il travolgente sviluppo tecnologico ci pone di fronte non può essere usata come

²⁸⁶ Sul dibattito tra la necessità di regolare o meno l’IA negli Stati Uniti si cfr. J.F. WEAVER, *Regulation of artificial intelligence in the United States*, cit.

²⁸⁷ Su questo tema si cfr. A. SIMONCINI, *Verso la regolamentazione della Intelligenza Artificiale. Dimensioni e governo*, in *BioLaw Journal – Rivista di BioDiritto*, n. 2/2021, pp. 411-417 e F. PIZZETTI, *La proposta di Regolamento sull’IA della Commissione Europea presentata il 21.4.2021 (COM (2021) 206 final) tra Mercato Unico e competizione digitale globale*, in *Diritto di Internet* n. 4/2021., pp. 591 ss.

²⁸⁸ Si cfr. *ex multis*, D. KRIEBEL et. al., *The Precautionary Principle in Environmental Science*, in *Environmental Health Perspectives* 109 (9), 2001, pp. 871-876.

ragione legittima per sottrarsi al complicato tentativo di tenere sotto controllo tale sviluppo attraverso gli strumenti che offre il diritto costituzionale²⁸⁹.

... e la giustizia algoritmica

Diverse sono le voci che animano il dibattito, brevemente delineato nel precedente paragrafo, sulla necessità di regolare l'intelligenza artificiale: voci che, per le differenze culturali e giuridiche che abbiamo detto, negli Stati Uniti faticano a farsi sentire, e che in Europa hanno avuto la meglio.

Comprensibilmente, visto l'ingresso ancora timido negli ordinamenti europei di algoritmi e dell'intelligenza artificiale nei sistemi giudiziari (di cui abbiamo ampiamente discusso *supra*, cap. II), molte meno sono state le istanze di regolazione della giustizia algoritmica, tendenza che è fortunatamente mutata quando, anche a causa degli sviluppi che essa stava avendo al di fuori dal contesto europeo (e in particolare negli Stati Uniti), il Consiglio d'Europa attraverso la CEPEJ ha pubblicato il primo strumento di regolazione (*soft*) organica sulla intelligenza artificiale nei sistemi giudiziari, la Carta Etica (già più volte menzionata e di cui si metteranno in luce i principi nel prosieguo del capitolo).

È proprio l'esperienza statunitense dell'impiego di algoritmi predittivi all'interno del procedimento penale che ha mostrato chiaramente come l'impiego degli algoritmi nell'amministrazione della giustizia, e soprattutto (nel *focus* di questo lavoro), nel processo decisionale del giudice, pongano numerosi interrogativi rispetto alla protezione dei diritti fondamentali e dei principi che regolano il giusto processo. In particolare, sono emersi dubbi sul ruolo che possono avere i privati nella progettazione degli algoritmi e nella raccolta dei dati; dubbi sulla possibilità delle parti di conoscere e contraddire il risultato algoritmico; dubbi sulla capacità del giudice di comprendere tale risultato e quindi a discostarsene.

²⁸⁹ Segnatamente, propone tale approccio Andrea Simoncini in A. SIMONCINI - S. SUWEIS, *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, in *Rivista di filosofia del diritto*, 2019, pp. 87-106, ripreso in A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, cit. pp. 63 ss.

Nel contesto statunitense, l'approccio liberista rispetto all'impiego degli algoritmi ha obbligato i giudici a confrontarsi con queste dinamiche privi di una cornice normativa, e a calare – talvolta con approssimazione, come si vedrà – i principi generali del giusto processo nella nuova realtà algoritmica.

In Europa, come già anticipato, si è invece avvertita l'esigenza di porre sul tavolo una serie di principi e regole che devono (e dovranno) accompagnare lo sviluppo della giustizia predittiva, dal rispetto dei diritti fondamentali, al principio di trasparenza fino al principio di umanità della decisione. Tale spinta regolativa ha seguito un lungo processo osmotico tra Unione europea e Consiglio d'Europa, con quest'ultimo che, come visto, ha aperto la strada a una regolazione dell'IA e la prima che ha elevato tale regolazione all'adozione di uno strumento vincolante, l'*AI Act*, che contiene diverse previsioni sulla giustizia algoritmica.

Per quanto ancora non sappiamo come si evolverà la materia, si può già dire che rispetto a quello che è accaduto negli Stati Uniti – che non sembra prossimo a un cambiamento – in Europa la strada della giustizia algoritmica sembra più protetta e arginata dai rischi che questa ha già mostrato di poter comportare con riguardo alle garanzie costituzionali legate al giusto processo.

Tali premesse sono necessarie per comprendere al meglio i diversi approcci alla regolazione dell'IA e della giustizia algoritmica, che saranno delineati nelle pagine seguenti. Un'ultima precisazione si rende necessaria: per ragioni di organicità del lavoro nei seguenti paragrafi si darà conto *in generale* delle previsioni rilevanti sulla giustizia algoritmica. Solo in un secondo momento, quando si saranno costituzionalizzate le problematiche che questa ha posto sul ruolo del giudice e suoi diritti fondamentali delle parti, tali previsioni verranno di volta in volta calate nel contesto di riferimento.

LA REGOLAZIONE DELL'IA E DELLA GIUSTIZIA ALGORITMICA NEGLI STATI UNITI

SOMMARIO: 1. La regolazione frammentata dell'intelligenza artificiale negli Stati Uniti - 2. La regolazione federale: gli approcci strategici della Casa Bianca - 2.1. La *self-regulation* come paradigma regolatorio - 2.2. Verso la co-regolamentazione - 3. La regolazione statale dei *risk assessment tools* (rinvio) - 4. Quale giustizia algoritmica nel contesto regolatorio statunitense?

1. La regolazione frammentata dell'intelligenza artificiale negli Stati Uniti

Come è emerso nei paragrafi precedenti, mentre l'Europa si è dotata di un impianto regolatorio rigido e uniforme nei confronti di tutti gli Stati membri, negli Stati Uniti le strategie sono state e sono tutt'ora incentrate su visioni piuttosto differenti²⁹⁰.

La comprensione delle dinamiche regolatorie seguite oltreoceano è rilevante per diverse ragioni. In primo luogo, è negli Stati Uniti, oltre che in Cina, che ad oggi si è sviluppata in modo massiccio l'industria dell'intelligenza artificiale, sicché conoscere le *policy* adottate in questo contesto potrebbe fornire qualche insegnamento anche per ragionare di buona regolazione quando il settore dell'IA si svilupperà in misura maggiore anche nel nostro continente.

Inoltre – come visto – è negli Stati Uniti che sono emerse le esperienze più rilevanti per quanto concerne l'impiego dell'intelligenza artificiale nel panorama giudiziario e capirne le radici anche dal punto di vista di regolatorio consentirà di avere un quadro più ampio e così di svolgere una riflessione di carattere costituzionale traslata nel contesto europeo.

Coerentemente con la natura e gli obiettivi di questo lavoro, maggior rilievo verrà ovviamente dato agli aspetti che riguardano, direttamente o indirettamente, lo

²⁹⁰ Ma si v. A. LALLI, *I non radicalmente diversi approcci alla regolazione dell'intelligenza artificiale di Europa, Stati Uniti e Cina*, in A. LALLI (a cura di), *La regolazione pubblica delle tecnologie digitali e dell'intelligenza artificiale*, Giappichelli, 2024, pp. 115-140.

scenario relativo all'IA nel giudiziario. Ciò in quanto le applicazioni di intelligenza artificiale sono una miriade e sovente, come si vedrà, l'approccio negli Stati Uniti è stato – a differenza dell'Europa – quello di regolare specificamente e settorialmente le diverse applicazioni di IA.

È utile qui fornire quindi brevemente le coordinate delle iniziative assunte a livello federale per comprendere in generale la visione degli Stati Uniti sulle strategie regolatorie rispetto all'intelligenza artificiale.

Per quanto concerne la politica federale di governance dell'IA, la Casa Bianca e una serie di agenzie federali hanno presentato diverse iniziative, leggi e politiche relative all'IA. Sebbene nel corso degli anni siano entrate in vigore anche numerose leggi statali sull'IA, le politiche federali rivestono come detto un'importanza fondamentale per comprendere la strategia nazionale del paese. In effetti, le basi della strategia del governo federale in materia di IA sono già state stabilite e forniscono indicazioni su come verranno affrontate le questioni legali e politiche poste da questa nuova tecnologia negli anni a venire.

Prima di vedere quindi quali sono stati gli approcci che si sono susseguiti è utile introdurre un ultimo elemento: nella politica americana concernente la regolazione dell'intelligenza artificiale c'è un grande assente, il Congresso.

O meglio, il Congresso è intervenuto in alcuni specifici settori (principalmente negli ambiti delle macchine a guida autonoma e della difesa), ma ad oggi – a differenza dell'approccio Europeo – manca una regolamentazione specifica e vincolante rispetto all'intelligenza artificiale²⁹¹.

L'unico intervento organico del Congresso è stato il *National Artificial Intelligence Initiative Act* del 2020, entrato in vigore nel 2021²⁹², legge bipartisan che, lungi dal regolare gli usi dell'IA, aveva l'obiettivo di preservare il primato

²⁹¹ Ci sono stati nel corso degli ultimi anni delle proposte legislative ma che non sono mai andate in porto. Si v. V. BERGENGRUEN, *AI Regulation Takes Baby Steps on Capitol Hill*, in *Time.com*, 14 settembre 2023. Sicuramente un'iniziativa rilevante a livello congressuale è stata il *Bipartisan Framework for U.S. AI Act*, che proponeva di istituire un organismo di supervisione indipendente con il compito di richiedere alle aziende che sviluppano modelli di IA *general purpose* o *high risk* di registrarsi direttamente presso un'autorità designata. Si v. la proposta al seguente link: <https://www.blumenthal.senate.gov/imo/media/doc/09072023bipartisanaiframework.pdf>.

²⁹² *National AI Initiative Act* (Division E, Sec. 5001), consultabile al seguente indirizzo: <https://www.congress.gov/116/crpt/hrpt617/CRPT-116hrpt617.pdf#page=1215>

degli US nello sviluppo dell'IA²⁹³ e impegnava il Governo a implementare la così denominata *National AI Initiative* indicando quali obiettivi: «(1) *ensure continued United States leadership in artificial intelligence research and development*; (2) *lead the world in the development and use of trustworthy artificial intelligence systems in the public and private sectors*; (3) *prepare the present and future United States workforce for the integration of artificial intelligence systems across all sectors of the economy and society*; and (4) *coordinate ongoing artificial intelligence research, development, and demonstration activities among the civilian agencies, the Department of Defense and the Intelligence Community to ensure that each informs the work of the others*».

Tale legge ha previsto inoltre l'istituzione di organismi per fornire un coordinamento a livello federale, tra cui un *National Artificial Intelligence Office* incaricato di sostenere la ricerca e lo sviluppo dell'IA, le iniziative educative, la pianificazione tra le diverse agenzie e la cooperazione internazionale²⁹⁴.

La mancanza di una legislazione federale ha comportato che i singoli Stati abbiano da qualche anno iniziato a regolamentare alcuni aspetti dell'intelligenza artificiale, approvando leggi, tra gli altri, sulla non discriminazione algoritmica, sulla sicurezza, sulla privacy e sulla protezione dei consumatori²⁹⁵.

Il quadro che emerge è una politica sull'IA particolarmente frammentata, divisa tra strumenti legislativi non vincolanti provenienti dal Governo federale e legislazioni statali su singoli aspetti legati all'IA. Ci si potrebbe interrogare a lungo sui motivi per cui negli Stati Uniti pare oggi quasi utopistica l'idea di approvare una legislazione organica sull'intelligenza artificiale a livello federale²⁹⁶. In estrema sintesi: la Costituzione americana divide il potere tra il governo federale e quelli

²⁹³ H. ROBERTS - J. COWLS - E. HINE, *et al*, *Achieving a 'Good AI Society': Comparing the Aims and Progress of the EU and the US*, in *Sci Eng Ethics* 27, 68, 2021, p. 12.

²⁹⁴ H. PACK, *Regulation of Artificial Intelligence in the United States*, in *Interactive Robotics: Legal, Ethical, Social and Economic Aspects*, Springer International Publishing, 2018, p. 234.

²⁹⁵ N.T. LEE - O. CHIJOKE, *Why states and localities are acting on AI*, in *Brookings.edu*, 15 dicembre 2023.

²⁹⁶ Appare emblematico l'intervento a una conferenza di un Senatore repubblicano, Ted Cruz, il quale ha affermato che il Congresso «*doesn't know what the hell it's doing*» quando si tratta di regolare l'intelligenza artificiale. Si v. <https://www.politico.com/news/2023/06/15/ai-ted-cruz-congress-00102116>.

statali, il che complica un approccio unificato alla regolamentazione dell'IA; il processo legislativo degli Stati Uniti rende difficile l'approvazione delle leggi, specialmente in un campo in rapida evoluzione come quello dell'IA²⁹⁷; l'industria tecnologica statunitense esercita una considerevole influenza sulle discussioni normative riguardanti l'IA e il Governo, che vuole mantenere il ruolo di *leadership* nel settore, monitora da vicino gli interessi del settore privato e ne permette una sostanziale auto-regolamentazione quando si tratta di intelligenza artificiale²⁹⁸.

2. La regolazione federale: gli approcci strategici della Casa Bianca

I primi strumenti federali di para-regolazione dell'intelligenza artificiale negli Stati sono stati elaborati non in sede congressuale ma dagli uffici presidenziali e si sono succeduti a partire dalla Presidenza Obama. La politica americana di regolazione dell'IA è infatti iniziata alla fine della presidenza di Barack Obama nel 2016, proseguita poi nelle Amministrazioni Trump e Biden. Questi approcci sono caratterizzati da diversi temi coerenti, tra cui la minimizzazione dell'intervento governativo e l'enfasi sul ruolo del libero mercato e l'alta considerazione per la capacità di innovazione e sviluppo²⁹⁹.

2.1. La self-regulation come paradigma regolatorio

²⁹⁷ In generale sulla contrazione sostanziale del potere legislativo federale si cfr. C. BASSU - M. BETZU - F. CLEMENTI - G. COINU, *Diritto costituzionale degli Stati Uniti d'America. Una introduzione*, Giappichelli, 2022, pp. 34-36.

²⁹⁸ Per approfondire in generale le difficoltà di regolazione dell'IA, si vedano le interessanti considerazioni espresse in I. BREMMER - M. SULEYMAN, *The AI Power Paradox: Can States Learn to Govern Artificial Intelligence—Before It's Too Late?*, in *Foreign Affairs*, 2023.

²⁹⁹ E. HINE - L. FLORIDI, *Artificial intelligence with American values and Chinese characteristics: a comparative analysis of American and Chinese governmental AI policies*, in *AI and Society*, 39, 2022, p. 260.

Nell'ottobre 2016 il *White House Office of Science and Technology Policy (OSTP)* ha pubblicato un primo *report* sull'IA, denominato “*Preparing for the future of Artificial Intelligence*”³⁰⁰.

La visione governativa che emerge dal *report* pone come obiettivo della regolazione dell'intelligenza artificiale l'innovazione e lo sviluppo della stessa nel libero mercato, sicché «*broad regulation of AI research or practice would be inadvisable at this time*»³⁰¹.

In tal senso, si legge ancora nel *report*, la auto-regolamentazione da parte delle aziende e le linee guida esistenti sono già sufficienti, e semmai da adattare, se necessario, allo sviluppo dell'IA³⁰². In questo primo approccio al tema, il Governo federale vede quindi come assai limitato il proprio ruolo come regolatore pubblico, anche se in alcune occasioni il *report* sembra mettere in guardia sul ruolo che deve avere la regolazione per garantire la sicurezza pubblica, affermando che «*agencies must remain mindful of the fundamental purposes and goals of regulation to safeguard the public good, while creating space for innovation and growth in AI*»³⁰³.

L'approccio iniziale del governo federale è stato da più parti giustamente criticato sia per l'assenza di una visione organica del fenomeno sia per la scelta di dare la priorità all'innovazione rispetto alle conseguenze sociali del fenomeno: si è detto per esempio che esso «*can best be summarized as trying to fit AI into the specific vision of US national priorities, instead of seeing the new features of AI as a good opportunity to revisit these priorities, both nationally and internationally*»³⁰⁴. Come infatti espresso in questo lavoro (si v. *supra* la nota introduttiva alla presente parte), l'intelligenza artificiale pone al regolatore delle sfide del tutto nuove che, come tali, richiedono una regolamentazione autonoma in

³⁰⁰ Si v. il Report pubblicato sul sito della Casa Bianca seguente indirizzo: https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

³⁰¹ *Ivi*, p. 17.

³⁰² *Ibidem*.

³⁰³ *Ibidem*.

³⁰⁴ C. CATH *et al.*, *Artificial Intelligence and the 'Good Society', the US, EU, and UK approach*, in *Sci Eng Ethics*, 2018.

grado di bilanciare la positiva spinta all'innovazione con gli interessi della comunità.

In ogni caso, per quanto qui interessa, nel *report* si trovano alcune indicazioni rilevanti per quanto concerne l'impiego dell'intelligenza artificiale nel contesto giudiziario. Consapevole dell'impiego in alcune giurisdizioni statali di algoritmi che aduano l'organo giudiziario ad assumere la decisione (i *risk assessment tools*, di cui abbiamo parlato *supra*, cap. I) il *report* evidenzia che «*in the criminal justice system, some of the biggest concerns with Big Data are the lack of data and the lack of quality data. AI needs good data. If the data is incomplete or biased, AI can exacerbate problems of bias. It is important that anyone using AI in the criminal justice context is aware of the limitations of current data*»³⁰⁵.

La risposta a questo problema dovrebbe essere, secondo la visione dell'ufficio presidenziale – e vedremo che tale soluzione, se intesa in senso limitato, non risolve tutti i problemi – quella di garantire «*greater transparency when AI tools are used for public purposes*», che deve essere perseguita «*not only on the data and algorithms used, but also on the potential to have some form of explanation for any AI-based determination*»³⁰⁶.

L'approccio cauto (*laissez-faire approach*, come è stato definito³⁰⁷) rispetto a una regolazione dell'IA inizialmente non mutò nemmeno con l'amministrazione Trump che anzi non sembrava volere elaborare alcun piano nazionale rispetto al fenomeno³⁰⁸.

Dopo aver ricevuto critiche da più fronti, la Casa Bianca in realtà fece parzialmente marcia indietro e pubblicò nel 2019 un *Executive Order* dal titolo – a ben vedere emblematico – “*Maintaining American Leadership in Artificial Intelligence*”³⁰⁹ focalizzato principalmente sullo sviluppo economico dell'intelligenza artificiale all'interno degli Stati Uniti. Si legge nel documento che

³⁰⁵ *Preparing for the future of AI*, p. 30.

³⁰⁶ *Ibidem*.

³⁰⁷ H. ROBERTS, J. COWLS, E. HINE, *et al*, *Achieving a 'Good AI Society': Comparing the Aims and Progress of the EU and the US*, cit., p. 11.

³⁰⁸ *Ivi*, p. 12.

³⁰⁹ Executive Order 13859 “*Maintaining American Leadership in Artificial Intelligence*”, 11 febbraio 2019.

«[i]t is the policy of the United States Government to sustain and enhance the scientific, technological, and economic leadership position of the United States in AI R&D [research and development] and deployment through a coordinated Federal Government»³¹⁰.

L'*executive order* diede avvio alla *American AI Initiative*, una strategia nazionale che avrebbe dovuto seguire cinque principi: 1) promuovere i progressi tecnologici nell'IA per favorire la scoperta scientifica, la competitività economica e la sicurezza nazionale; 2) sviluppare standard tecnici appropriati e ridurre gli ostacoli alla sperimentazione e all'impiego sicuro dell'IA; 3) formare i lavoratori americani con le competenze necessarie per prepararli ai lavori del futuro; 4) promuovere la fiducia del pubblico nelle tecnologie di IA e proteggere le libertà civili, la *privacy* e i valori americani; 5) promuovere un ambiente internazionale che favorisca la ricerca e apra i mercati per le industrie americane proteggendo al contempo il vantaggio tecnologico degli Stati Uniti, comprese le tecnologie dell'IA, dai concorrenti³¹¹.

Anche qui le indicazioni appaiono assai vaghe e sembrano andare nella direzione di lasciare quanto più possibile il mercato dell'intelligenza artificiale privo di vincoli giuridici particolarmente stringenti.

Nell'ambito della *AI Initiative* vennero approvati altri documenti non vincolanti relativi alla regolazione dell'IA. Tra essi spicca, per quello che qui interessa, un *Memorandum* del Novembre 2020 "*Guidance for Regulation of Artificial Intelligence Applications*"³¹².

L'approccio seguito nel *Memorandum*, definito da alcuni come «*fairly permissive*»³¹³ riflette nuovamente l'idea che la regolazione possa soprattutto essere di ostacolo alla crescita e all'innovazione del settore dell'intelligenza artificiale.

Il *Memorandum* si rivolge al potere regolatorio delle agenzie governative e afferma che «[f]ederal agencies must avoid regulatory or non-regulatory actions

³¹⁰ *Ivi*, p. 1.

³¹¹ *Ivi*, p. 2.

³¹² Il *Memorandum* è consultabile al seguente link: <https://www.whitehouse.gov/wp-content/uploads/2020/11/M-21-06.pdf>

³¹³ Si v. *Artificial Intelligence: U.S. Federal Consideration* al seguente indirizzo: <https://www.lexology.com/library/detail.aspx?g=7d055a38-f35c-4346-917b-92a4132f6aac>.

that needlessly hamper AI innovation and growth. Where permitted by law, when deciding whether and how to regulate in an area that may affect AI applications, agencies should assess the effect of the potential regulation on AI innovation and growth. While narrowly tailored and evidence based regulations that address specific and identifiable risks could provide an enabling environment for U.S. companies to maintain global competitiveness, agencies must avoid a precautionary approach that holds AI systems to an impossibly high standard such that society cannot enjoy their benefits and that could undermine America's position as the global leader in AI innovation»³¹⁴.

Se queste sono le coordinate generali, nel *Memorandum* si dà conto comunque – seppure in maniera molto generica – dei possibili rischi che l’IA potrebbe recare a «*privacy, individual rights, personal choice, civil liberties, public health, safety*»³¹⁵ e la scelta tra regolazione e non regolazione deve tenere conto del rischio che ogni applicazione di intelligenza artificiale pone rispetto a tali interessi, in modo quindi non dissimile all’approccio seguito in Unione europea. Si legge infatti che «*[i]t is not necessary to mitigate every foreseeable risk; in fact, a foundational principle of regulatory policy is that all activities involve tradeoffs. Instead, a risk-based approach should be used to determine which risks are acceptable and which risks present the possibility of unacceptable harm, or harm that has expected costs greater than expected benefits*»³¹⁶.

Nel *Memorandum*, infine, non si trovano indicazioni specifiche rispetto all’impiego di strumenti di IA nel giudiziario, ma solo coordinate generali sul dovere delle agenzie di considerare le problematiche di giustizia (*fairness*) e non discriminazione rispetto alle decisioni assunte dalle macchine, «*as well as whether the AI application at issue may reduce levels of unlawful, unfair, or otherwise unintended discrimination as compared to existing processes*»³¹⁷.

³¹⁴ *Memorandum* 2020, p. 2.

³¹⁵ *Ivi*, p. 3.

³¹⁶ *Ivi*, p. 4.

³¹⁷ *Ivi*, p. 6.

2.2. Verso la co-regolamentazione

Con l'amministrazione Biden, iniziata nel 2021, si assiste a un primo parziale cambio di rotta e il Governo statunitense inizia a voler regolamentare il libero mercato dell'intelligenza artificiale in modo più stringente rispetto all'amministrazione Trump e a costruire una visione più coerente e diversificata dell'IA, che non abbracci solo il versante economico e dello sviluppo commerciale ma che prenda in considerazione anche la dimensione sociale del fenomeno³¹⁸.

Un primo documento che mostra le rinnovate priorità del Governo federale statunitense è il *BluPrint for AI* denominato “*AI Bill of Rights making automated systems work for the american people*”³¹⁹. Esso enuncia cinque principi che dovrebbero guidare lo sviluppo e l'impiego dei sistemi automatici nell'era dell'intelligenza artificiale e proteggere il pubblico dai rischi connessi ad essa. Sono principi che, come si vedrà, troveranno espressione anche nella legislazione europea e che si rivolgono in generale ai sistemi di intelligenza artificiale impiegati in un numero indefinito di campi. Occorre comunque darne conto perché aiutano senz'altro a fornire una delimitazione giuridica anche ai sistemi di IA utilizzati nel campo giudiziario che, come visto, oggi negli Stati Uniti non ricevono una regolazione autonoma e specifica.

Il primo principio, *Safe and Effective Systems*, richiede che i sistemi automatizzati debbano essere sviluppati con la consultazione di diverse parti interessate e di esperti per identificare i potenziali impatti negativi sulla società nonché la presenza di valutazioni indipendenti che attestino la sicurezza degli stessi³²⁰.

³¹⁸ E. HINE - L. FLORIDI, *Artificial intelligence with American values and Chinese characteristics: a comparative analysis of American and Chinese governmental AI policies*, cit., p. 263.

³¹⁹ U.S. OFFICE OF SCIENCE AND TECHNOLOGY POLICY, *Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People*, ottobre 2022. Il documento è consultabile al seguente link: <https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf>

³²⁰ *Ivi*, p. 5.

Nel secondo principio, *Algorithmic Discrimination Protections*, si riconosce la potenzialità intrinseca degli algoritmi a riprodurre le discriminazioni esistenti e invita progettisti, sviluppatori e utilizzatori di sistemi automatizzati ad adottare misure attive per proteggere gli individui e le comunità dalla discriminazione algoritmica e per utilizzare e progettare i sistemi in modo equo³²¹.

Il terzo principio attiene alla *Data Privacy* e quindi innanzitutto all'utilizzazione di dati che siano strettamente necessari per il contesto in cui il sistema viene impiegato. Inoltre prevede la richiesta di consenso e autorizzazione per l'utilizzo dei dati personali, e forme più stringenti di protezione per quanto concerne l'utilizzo di dati in settori sensibili (*sensitive domains*), quelli cioè legati alla salute, al lavoro, all'istruzione alla giustizia penale e al settore finanziario³²².

Il principio del *Notice and Explanation* specifica che quando un sistema automatizzato è in uso, gli utenti devono essere informati con un linguaggio semplice che descriva il funzionamento generale del sistema e il ruolo svolto dall'automazione, i soggetti responsabili della progettazione del sistema e le spiegazioni dei risultati³²³.

Infine, il principio denominato *Human Alternatives, Consideration, And Fallback*, molto importante per il contesto del presente lavoro – come meglio si vedrà – prevede innanzitutto la possibilità di rinunciare (*opt-out*), quando ciò è previsto dalla legge, al trattamento automatizzato in favore di una decisione umana. Inoltre, il principio considera i rimedi che dovrebbero caratterizzare decisioni automatizzate. In questo senso dovrebbe venire sempre garantito l'accesso a un esame umano tempestivo e a un rimedio efficace. L'esame deve essere accessibile, equo, effettivo e accompagnato da un'adeguata formazione degli operatori. Infine, in settori sensibili come la *giustizia penale* tale principio deve essere applicato con maggiore rigore. Proprio a quest'ultimo riguardo, nelle specificazioni del principio in esame si mette in guardia sull'utilizzo di sistemi automatizzati nel *sentencing*. Si legge infatti che «*automated systems should not be allowed to directly intervene in*

³²¹ *Ivi*, p. 5.

³²² *Ivi*, p. 6.

³²³ *Ivi*, p. 6.

*high-risk situations, such as sentencing decisions or medical care, without human consideration»*³²⁴. Come già si è visto, in molte giurisdizioni statali tale monito è ancora inascoltato, data la diffusione dell'impiego di sistemi automatizzati nel processo di determinazione della pena, spesso accompagnato da autorizzazioni legislative o giurisprudenziali (si v. *supra*, cap. I § 5).

Il *Blueprint* è stato accolto positivamente da chi ritiene che una regolazione che abbia come stella polare non solo l'aspetto economico ma anche i diritti delle persone e negativamente da chi ha espresso preoccupazione sul fatto che principi così stringenti potrebbero far perdere il ruolo di *leadership* agli Stati Uniti rispetto ai concorrenti, soprattutto la Cina³²⁵.

Sebbene sia uno strumento non vincolante³²⁶, la sua portata non deve essere sottovalutata in quanto l'efficacia persuasiva di tali raccomandazioni sulla futura legislazione è ben nota e come vedremo esso ha determinato effettivamente un positivo cambio di rotta rispetto all'approccio federale alla regolazione dell'IA³²⁷. Al riguardo un numero non irrilevante di agenzie federali ha implementato le linee-guida indicate nel *Blueprint* e iniziato a promuovere *policy* per un uso responsabile dell'intelligenza artificiale³²⁸.

Un ulteriore passo verso una più massiccia regolazione pubblica dell'intelligenza artificiale è stato compiuto di recente con l'approvazione da parte dell'Amministrazione Biden, il 30 ottobre 2023, dell'*Executive Order “on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence”*³²⁹, che ha introdotto diverse direttive e termini alle agenzie federali per incoraggiare sì l'innovazione, ma riconoscendo altresì i rischi relativi al rispetto dei diritti fondamentali derivanti dall'IA. Trattandosi del primo strumento parzialmente

³²⁴ *Ivi*, p. 51.

³²⁵ Si cfr. E. STRICKLAND, *6 Reactions to the White House's AI Bill of Rights*, in *IEEE Spectrum*, 14 ottobre 2022.

³²⁶ E. HINE - L. FLORIDI, *The Blueprint for an AI Bill of Rights: In Search of Enaction, at Risk of Inaction, in Minds and Machines*, 30 novembre 2022.

³²⁷ E. PARK, *The AI Bill of Rights: A Step in the Right Direction*, in *Orange County Lawyer Magazine*, Vol. 65, 2, 2023, p. 27.

³²⁸ N. T. LEE - J. MALAMUD, *Opportunities and blind spots in the White House's blueprint for an AI Bill of Rights*, in *Brookings.com*, 19 dicembre 2022.

³²⁹ THE WHITE HOUSE, *Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, 30 ottobre 2023.

vincolante rispetto alla regolamentazione dell'IA, l'*Executive Order* – nonostante non sia esente da critiche – rappresenta senz'altro un punto di svolta nella politica americana.

Anche se con un approccio molto meno dettagliato rispetto al Regolamento europeo sull'IA dovuto al diverso strumento giuridico adottato – da una parte una legge vincolante, dall'altra un atto che non prevede una normativa organica ma piuttosto delle indicazioni (talora vincolanti, talora no) rivolte all'apparato governativo federale affinché si conformi ai principi stabiliti nel documento – anche l'*Executive Order* sembra condividere un approccio basato sul rischio. Tuttavia, mentre nell'*AI Act* europeo il contraltare all'utilizzo dell'IA è il rischio per la salvaguardia dei diritti fondamentali, nell'*Executive Order* sembra essere più che altro la sicurezza dei sistemi e la limitazione delle interferenze con la *national security* statunitense. Ciò peraltro non significa che sia assente un affondo sul rispetto dei diritti fondamentali, ma che il rispetto di essi non sembra nemmeno stavolta al centro dell'interesse dell'amministrazione americana.

Come già avvenuto per gli strumenti regolatori fino a qui analizzati, anche per l'*Executive Order* non sarà necessario svolgere una disamina profonda del documento, che esula dallo scopo del presente lavoro. È invece senz'altro utile fornire una panoramica generale delle previsioni *ivi* contenute e soffermarsi su quelle che più nel dettaglio affrontano il problema dell'intelligenza artificiale impiegata nel settore della giustizia.

Innanzitutto nell'*Order* vi è la previsione di otto principi cui le agenzie federali dovrebbero ispirarsi per adempiere alle azioni richieste nello stesso.

Il primo principio richiede che l'intelligenza artificiale sia «*safe and secure*» e cioè che siano previste delle procedure e delle regole in grado di «*mitigate risks from these systems before they are put to use*».

Il secondo principio mette nuovamente al centro l'innovazione e il ruolo degli Stati Uniti di locomotiva della progettazione di questa tecnologia, supportando l'adozione di *policy* in grado di attrarre i migliori talenti nel campo in modo tale che «*companies and technologies of the future are made in America*».

L'Order prosegue indicando nel terzo principio il supporto ai lavoratori americani, che dovranno fronteggiare un forte cambiamento in quanto poiché «*AI creates new jobs and industries, all workers need a seat at the table, including through collective bargaining, to ensure that they benefit from these opportunities*». L'obiettivo è dunque quello di «*support responsible uses of AI that improve workers' lives, positively augment human work, and help all people safely enjoy the gains and opportunities from technological innovation*».

Solo al quarto principio si prevede che le *policies* inerenti all'IA devono essere improntate al rispetto dell'«*equity*» e dei «*civil rights*» soprattutto nel campo delle assunzioni lavorative (*hiring*), nel settore sanitario (*healthcare*) e il diritto ad abitare (*housing*). In questi contesti – si legge nell'Order – è nota la suscettibilità dei sistemi artificiali di aggravare discriminazioni e pregiudizi esistenti, così intensificando le ingiustizie già esistenti. Qui si intravede anche la consapevolezza delle possibili storture nell'impiego di IA nel settore della giustizia in quanto si legge che «*it is necessary to hold those developing and deploying AI accountable to standards that protect against unlawful discrimination and abuse, including in the justice system and the Federal Government*».

Il quinto principio concerne la protezione dei consumatori nel nuovo contesto dell'intelligenza artificiale che deve essere ulteriormente rafforzata soprattutto in alcuni campi particolarmente critici – quali i servizi finanziari, l'istruzione, l'ambito immobiliare e i trasporti – «*where mistakes by or misuse of AI could harm patients, cost consumers or small businesses, or jeopardize safety or rights*».

Il sesto principio è incentrato sul rispetto della *privacy* e dei dati personali, diritti sempre a rischio quando si tratta di intelligenza artificiale e per la protezione dei quali il Governo federale «*will ensure that the collection, use, and retention of data is lawful, is secure, and mitigates privacy and confidentiality risks*»

Il settimo principio è un invito fatto dall'Amministrazione federale di riuscire a tenere sotto controllo i rischi che pone l'IA e migliorare la propria capacità di regolarne l'uso per garantire migliori risultati per i cittadini americani. Per raggiungere tali obiettivi occorre, secondo l'Order, «*modernize Federal Government information technology infrastructure, remove bureaucratic*

obstacles, and ensure that safe and rights-respecting AI is adopted, deployed, and used».

Infine, l'ottavo e ultimo principio sembra racchiudere tutti gli altri sette e, come si può immaginare, prevede che sia il Governo federale statunitense a guidare il progresso sociale, economico e tecnologico mondiale non soltanto attraverso l'espansione dell'industria dell'IA, ma anche implementando garanzie di responsabilità e sicurezza dei sistemi, da condividere con quanti più attori mondiali possibile, per assicurare «*that AI benefits the whole world, rather than exacerbating inequities, threatening human rights, and causing other harms*».

Come anticipato, per ognuno dei principi enunciati, l'*Executive Order* richiede alle singole agenzie federali e ai dipartimenti amministrativi, entro dei termini che variano da 30 ai 356 giorni, di sviluppare una serie di strumenti regolatori e strategie di intervento nei propri settori di competenza³³⁰.

Quanto al tema qui di interesse, nell'ambito del principio della non discriminazione e della protezione dei *Civil rights*, l'*Order* contiene un serie di previsioni interessanti che riguardano l'impiego dell'IA nel settore giudiziario.

Si prevede che l'*Attorney General*, la figura a capo del *Department of Justice*, al fine di assicurare «*fair and impartial justice for all, with respect to the use of AI in the criminal justice system*», debba presentare al Presidente, entro 365 giorni, un *report* sull'impiego dell'intelligenza artificiale nella giustizia penale, e in particolare l'IA impiegata nel *sentencing*, nella fase pre-processuale e più in generale nella predizione del rischio di recidiva condotta dai cd. *risk assessment tools*.

Nel suddetto *report* dovranno essere indicati gli ambiti dove l'IA potrebbe accrescere il livello di protezione delle libertà fondamentali, i limiti che invece devono continuare a esistere rispetto all'impiego di tale tecnologia nel processo e, aspetto come si vedrà molto importante, le modalità con cui promuovere la presenza, all'interno delle strutture pubbliche che hanno a che vedere con la

³³⁰ Gli ultimi dati disponibili della Casa Bianca rivelano che nel gennaio 2024 tutti gli obiettivi richiesti dall'*Order* sono stati rispettati dalle agenzie. Si v. il sito della Casa Bianca al seguente indirizzo: <https://www.whitehouse.gov/briefing-room/statements-releases/2024/04/29/>.

giustizia penale, di «*highly qualified and service-oriented officers and staff with relevant technical knowledge*». A quanto consta, mentre si scrive tale *report* non è stato ancora presentato.

All'*Executive Order* è seguito un *Memorandum* predisposto dall'*Office of Management and Budget* della Casa Bianca che integra e arricchisce le previsioni contenute nell'ordine esecutivo presidenziale³³¹.

Nel *Memorandum* vengono identificati, con un elenco esplicitamente definito non esaustivo, i sistemi di IA che si presume possano avere un impatto significativo sulla sicurezza e sui diritti (*safety-impacting* e *rights-impacting*).

È definito *rights-impacting* il sistema «*whose output serves as a principal basis for a decision or action concerning a specific individual or entity that has a legal, material, binding, or similarly significant effect*» su alcuni diritti degli individui, tra cui vengono individuati: «*[c]ivil rights, civil liberties, or privacy, including but not limited to freedom of speech, voting, human autonomy, and protections from discrimination, excessive punishment, and unlawful surveillance*»³³². Inoltre, ha un impatto sui diritti, secondo il *Memorandum*, il sistema che «*is used or expected to be used, in real-world conditions, to control or significantly influence the outcomes*» delle decisioni assunte in determinati contesti. Tra essi, per quanto qui interessa, vengono indicati i sistemi che producono previsioni di rischio di recidiva per gli individui e quelli impiegati per assumere decisioni nelle diverse fasi processuali (tra cui, ancora, il *sentencing*)³³³.

Ora, per tali sistemi il *Memorandum* richiede che le agenzie pubbliche, entro dicembre 2024, dovranno adottare alcune guarentigie (*minimum practices*) sia prima di poter impiegare tali sistemi sia durante il loro utilizzo. Tra le prime vengono ricomprese: una valutazione dell'impatto (*impact assessment*) dei sistemi impiegati, da aggiornare periodicamente; dei *test* delle *performances* dell'IA nel

³³¹ Office of Management and Budget, Memorandum For The Heads Of Executive Departments And Agencies “*Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence*”, 28 marzo 2024.

³³² *Ivi*, p. 29

³³³ *Ivi*, p. 32.

mondo reale; una valutazione di un organismo indipendente che certifichi che il sistema funzioni correttamente.

Le cautele che le agenzie devono assicurare nel ciclo di utilizzo dei sistemi di IA *rights-impacting* sono: un monitoraggio continuo sul corretto funzionamento del sistema; qualora dal monitoraggio emergano rischi di utilizzo prima sconosciuti, la mitigazione di essi tramite un controllo umano più stringente; garantire un *training* adeguato e specifico per coloro che sono chiamati ad interpretare gli *output* algoritmici; infine che quando un'IA viene utilizzata venga predisposta una documentazione in un linguaggio comprensibile del sistema che funga da pubblicità sia per coloro i quali impiegano l'IA sia per la generalità della popolazione³³⁴.

Sia l'*Executive Order* sia il *Memorandum* rappresentano indubbiamente un passo in avanti verso un approccio regolatorio pubblico del fenomeno dell'intelligenza artificiale a livello federale³³⁵.

Dopo anni di totale disinteresse nei confronti di una qualche forma di controllo pubblico sull'utilizzo di una tecnologia così problematica, la politica statunitense sembra voler mostrarsi in nuove vesti nel contesto internazionale, soprattutto ora che l'Unione europea ha approvato il primo strumento organico di regolazione vincolante dell'intelligenza artificiale.

In assenza, come visto, di legislazione da parte del Congresso, l'intervento governativo può infatti servire sia come spinta verso l'organo legislativo federale ad approvare delle serie leggi in materia, sia verso i singoli Stati ad allinearsi alle indicazioni che provengono dal Governo centrale³³⁶. Questi ultimi, infatti, sebbene negli anni abbiano approvato diversi *bill* in materia, presentano degli apparati normativi molto frammentati, che non si attagliano al contesto di un fenomeno che richiede una regolamentazione quanto più possibile uniforme³³⁷.

³³⁴ *Ivi*, pp. 17-21.

³³⁵ N.T. LEE, *How the White House Executive Order on AI ensures an effective governance regime*, in *Brookings.edu*, 28 marzo 2024.

³³⁶ N. T. LEE - O. CHIJOKE, *Why states and localities are acting on AI*, in *Brookings.edu*, 15 dicembre 2023.

³³⁷ S. FRIEDLER - J.HAVEN B - J. CHEN, *How the AI Executive Order and OMB memo introduce accountability for artificial intelligence*, in *Brookings.edu*, 16 novembre 2023.

Gli aspetti positivi non possono tuttavia oscurare le ancora numerosissime fragilità dell'attuale approccio di co-regolamentazione adottato negli Stati Uniti³³⁸.

Innanzitutto, gli *Executive Orders* non creano nuove leggi ma possono solo rivolgersi all'apparato governativo affinché si conformino alle previsioni indicate. Questo comporta che l'*Order* di cui si discute rientra ancora in una regolazione *soft* che manca di un regime di governance adeguato ed efficace, oltre che di meccanismi di monitoraggio, applicazione e sanzione.

Inoltre, la stabilità degli atti presidenziali è molto precaria in quanto al cambio di Amministrazione, che potrebbe avvenire con le prossime elezioni di novembre 2024, essi possono essere annullati o privati di efficacia – cosa peraltro appare molto probabile se consideriamo l'approccio che ha avuto l'Amministrazione Trump alla regolazione dell'IA.

Infine, proprio perché l'*Order* non crea nuove leggi, non sorgono obblighi nuovi per le imprese private che sviluppano e mettono sul mercato i sistemi di IA. Sebbene alcuni ritengano che in ogni caso l'impatto normativo del decreto presidenziale possa avere un'ampia portata anche nel settore privato³³⁹, essa rimane solo su un piano di influenza persuasiva esterna e non – come dovrebbe essere – su un piano di obblighi legislativi.

In definitiva, anche per anticipare il prossimo capitolo sulla regolazione in Europa, rispetto al Regolamento sull'intelligenza artificiale europeo, che prevede una normativa organica e vincolante, l'*Executive Order* delinea principalmente le linee guida che le agenzie federali devono seguire al fine di plasmare le pratiche del settore, ma non impone una regolamentazione ai soggetti privati. Inoltre, mentre come visto l'ordine esecutivo statunitense può essere modificato o revocato, soprattutto alla luce di cambiamenti elettorali, l'*AI Act* dell'Unione europea propone regole giuridicamente vincolanti in una struttura di *governance* duratura³⁴⁰.

³³⁸ In dottrina si v. M. WÖRSDÖRFER, *Biden's Executive Order on AI and the E.U.'s AI Act: A Comparative Computer-Ethical Analysis*, in *Philosophy & Technology*, 37(3), 2024.

³³⁹ S. FRIEDLER - J. HAVEN B - J. CHEN, *How the AI Executive Order and OMB memo introduce accountability for artificial intelligence*, cit.

³⁴⁰ B.C. LARSEN - S. KÜSPERT, *Regulating general-purpose AI: Areas of convergence and divergence across the EU and the US*, in *Brookings.edu*, 21 maggio 2024.

In termini di influenza sul panorama globale, infine, l'*Executive Order* sembra avere principalmente obiettivi di politica interna, mentre l'*AI Act* dell'UE, invece, parrebbe possedere la forza di rappresentare un precedente globale come già ha fatto il Regolamento generale sulla protezione dei dati (GDPR)³⁴¹.

3. La regolazione statale dei *risk assessment tools* (rinvio)

Conclusa la disamina degli approcci governativi federali alla regolazione dell'intelligenza artificiale e della giustizia algoritmica, necessari per comprendere in generale il panorama normativo ad oggi presente negli Stati Uniti, sarebbe del tutto eccentrico rispetto al *focus* di questo lavoro addentrarsi sulle iniziative legislative dei singoli Stati. Al contrario, comprendere come gli algoritmi predittivi sono entrati a far parte della quotidianità del processo penale ci è utile a inquadrare il rapporto tra “legislazione” federale e le modalità con cui gli stati hanno autorizzato l'impiego dei *tools* algoritmici nella giustizia penale.

Essendo un tema inscindibilmente connesso al discorso su come vengono impiegati tali sistemi, in quale fase processuale e con quali accortezze, si è deciso di parlarne già nel precedente capitolo (si v. *supra*, cap. I) a cui si rimanda.

4. Quale giustizia algoritmica nel contesto regolatorio statunitense?

Prima di procedere con l'analisi del contesto regolatorio europeo, l'abbondanza e iniziative messe in luce descrittivamente nel presente capitolo impone di dedicare un ultimo paragrafo all'individuazione dei principi e orientamenti che riguardano specificamente il contesto del presente lavoro. Ciò ci permette di fare ordine nella comprensione di quale giustizia algoritmica oggi si sta discutendo negli Stati Uniti.

Innanzitutto – e ciò era già chiaro dalle esperienze segnalate nel capitolo precedente – l'uso di algoritmi nei sistemi giudiziari è oggi limitato al contesto del procedimento penale: gli strumenti “normativi” che abbiamo preso in

³⁴¹ S. FELDSTEIN, *Evaluating Europe's push to enact AI regulations: how will this influence global norms?*, in *Democratization*, 31, 5, 2024.

considerazione che trattano della giustizia algoritmica limitano il proprio perimetro a quella che viene chiamata nei documenti “*criminal justice*”. Già questo è un segno di profonda differenza con le sperimentazioni di giustizia algoritmica presenti nel nostro continente (in cui gli algoritmi sono stati principalmente testati per essere impiegati nei processi civili o amministrativi. si v. *supra*, cap. II). Come visto, infatti, già la Carta Etica della CEPEJ ritiene che l’impiego dell’IA nel settore penale sia da considerare con estreme riserve, e l’*AI Act* – come si vedrà – rimarrà coerente con tale l’impostazione.

In ogni caso gli strumenti algoritmici nel *sentencing* sono considerati una situazione ad alto rischio per la sicurezza e per i diritti (*Memorandum* del marzo 2024) tanto che in un documento si dice che non dovrebbe esserne consentito l’impiego (*Blueprint*). Come vedremo, tale impostazione sembra ricalcare l’approccio basato sul rischio seguito dal Regolamento europeo sull’intelligenza artificiale, il quale definisce “ad alto rischio” quei sistemi che «hanno un impatto nocivo significativo sulla salute, la sicurezza e i diritti fondamentali delle persone nell’Unione»³⁴²

Quanto ai principi che regolano la giustizia algoritmica, gli strumenti automatizzati dovrebbero essere spiegabili e quindi trasparenti, anche se non viene esplicitata una soglia di trasparenza (principio *Notice and Explanation* contenuto nel *Blueprint*). Inoltre dovrebbe essere data la possibilità di richiedere una decisione umana e comunque, in caso di decisione automatizzata, la garanzia di un esame umano tempestivo e di rimedio efficace (principio *Human Alternatives, Consideration, And Fallback*).

Infine, in diverse iniziative viene sottolineata l’importanza della preparazione e la formazione di coloro i quali sono chiamati ad impiegare i sistemi algoritmici (*Executive Order 2023* e il successivo *Memorandum*).

Per completare il discorso rispetto alla visione della giustizia algoritmica oggi negli Stati Uniti, di recente (maggio 2024) un gruppo di lavoro bipartisan del Senato statunitense ha approvato una *roadmap* indicando una serie di raccomandazioni per

³⁴² *Considerando n. 46* del Regolamento.

governare l'intelligenza artificiale³⁴³. Tale documento pone ancora l'innovazione nel campo dell'IA in cima alle priorità e contiene raccomandazioni in realtà minime per l'introduzione di regolamentazioni forti intorno alla tecnologia, suggerendo invece ai comitati di esaminare ulteriormente se le normative esistenti possano essere applicate all'IA.

E tuttavia nel documento emergono alcune indicazioni importanti rispetto al tema qui in oggetto. Il gruppo di lavoro ha organizzato, con la partecipazione di più di 150 esperti, diversi *forum*, ognuno dei quali dedicato a uno specifico tema legato all'IA. Uno di essi ha trattato le aree maggiormente a rischio (*high risk*) in cui l'intelligenza artificiale viene impiegata, tra cui viene individuata la *criminal justice*. Nella *roadmap* vengono sintetizzate le conclusioni del *forum*, denominato *High impact uses of AI*.

In tali settori preme l'esigenza di trasparenza dei sistemi algoritmici³⁴⁴: la natura opaca di alcuni sistemi di IA potrebbe risultare giuridicamente «*unacceptable*»³⁴⁵. Per questo, il gruppo di lavoro incoraggia le commissioni legislative competenti a verificare la presenza di lacune nelle leggi applicabili al suddetto contesto a sviluppare delle iniziative legislative per colmare tali lacune. Tali iniziative dovrebbero quindi porre requisiti appropriati, caso per caso, sugli usi ad alto rischio dell'IA, come i requisiti di trasparenza, spiegabilità, test e valutazione³⁴⁶.

Come si vedrà, diversi principi indicati nei vari documenti segnalati corrispondono ai principi espressi nel contesto regolatorio europeo e nelle regole vincolanti previste nel nuovo Regolamento sull'intelligenza artificiale.

Sebbene dunque siamo ancora di fronte a strumenti di regolazione debole e a un contesto di impiego dell'IA circoscritto al procedimento penale, alcune delle problematiche generali dell'impiego di strumenti algoritmici nel campo della giustizia sembrano essere, almeno parzialmente, individuate anche oltreoceano negli approcci regolatori federali. Lo sguardo sull'intelligenza artificiale e sulla

³⁴³ The Bipartisan Senate AI Working Group, *Driving U.S. Innovation in Artificial Intelligence. A roadmap for artificial intelligence policy in the United States Senate*, disponibile al seguente indirizzo: https://www.schumer.senate.gov/imo/media/doc/Roadmap_Electronic1.32pm.pdf.

³⁴⁴ *Ivi*, p. 29.

³⁴⁵ *Ivi*, p. 10.

³⁴⁶ *Ibidem*.

giustizia algoritmica è quindi cambiato negli ultimi anni: da una visione del tutto priva di considerazioni sui possibili rischi dell'IA all'individuazione di settori *high risk*, di principi irrinunciabili, di barriere legislative necessarie.

Tuttavia, la non vincolatività delle previsioni che si sono qui evidenziate – oltre al fatto che si tratta strumenti normativi molto recenti e successivi all'impiego già strutturale dell'IA nel processo penale – ha comportato, e tuttora comporta, che nella folta casistica giurisprudenziale che si è formata in alcuni Stati rispetto alla compatibilità degli strumenti algoritmici impiegati nella giustizia penale con le regole del *due process*, i principi di trasparenza, di spiegabilità, di ricorso effettivo, di necessaria preparazione degli utilizzatori (tutti connaturati, come vedremo, al diritto a un processo giusto) sono stati costantemente messi da parte (si v. *infra*, cap. VIII, Sez. II).

LA REGOLAZIONE DELL'IA E DELLA GIUSTIZIA ALGORITMICA IN EUROPA

SOMMARIO: 1. La regolazione della giustizia algoritmica tra Unione europea e Consiglio d'Europa - 2. Il percorso di regolazione soft dell'Unione europea sino all'approvazione del regolamento sull'intelligenza artificiale - 2.1. I primi approcci strategici della Commissione europea - 2.2. Un primo strumenti di regolazione soft: gli "Orientamenti etici per un'IA affidabile" - 2.3. La crescente necessità di un quadro normativo per l'IA: il Libro Bianco sull'intelligenza artificiale - 3. L'*AI Act*: la regolazione *hard* (con elementi *soft*) dell'intelligenza artificiale - 3.1. La *RoadMap*: dalla Proposta della Commissione all'approvazione del Regolamento sull'intelligenza artificiale - 3.2. Profili generali del Regolamento - 4. La giustizia algoritmica nel prisma dell'*AI Act* - 4.1. La giustizia algoritmica nella Proposta della Commissione e il parere del Comitato economico sociale europeo - 4.2. L'idea di giustizia algoritmica nelle previsioni del Regolamento - 4.3. Gli strumenti di previsione del rischio di recidiva all'interno dell'*AI Act* - 4.4. L'implementazione dell'*AI Act* nel disegno di legge di delega italiano sull'intelligenza artificiale - 5. La recente strategia della Corte di Giustizia Europa sulla giustizia algoritmica - 6. La giustizia algoritmica nei lavori del Consiglio d'Europa - 6.1. Il trattato sull'intelligenza artificiale e le specifiche previsioni sulla giustizia algoritmica - 6.2. I lavori della CEPEJ e i principi della Carta Etica sull'impiego dell'IA nei sistemi giudiziari (e il suo seguito) - 7. Breve nota conclusiva: l'azione regolatoria europea come sentiero sicuro della giustizia algoritmica

1. La regolazione della giustizia algoritmica tra Unione europea e Consiglio d'Europa

Conclusa la disamina degli approcci regolatori negli Stati Uniti, occorre soffermarsi su quelli ad oggi assunti nel continente europeo, sia nell'ambito dell'Unione europea sia del Consiglio d'Europa.

Come vedremo, mentre il Consiglio d'Europa, tramite la Commissione europea per l'efficienza della giustizia (CEPEJ) ha nel tempo prodotto numerosi strumenti di *soft law* per regolare specificamente la giustizia algoritmica, l'Unione Europea, dal 2018 in avanti, ha cominciato a vedere l'intelligenza artificiale nel suo

complesso come un fenomeno che necessitava di una regolazione, quanto meno *soft*. Il percorso dell'Unione europea, con mutamenti di prospettive e di priorità, è arrivato infine ad approvare il primo Regolamento al mondo che regola in modo organico l'intelligenza artificiale. Sia nelle strategie di pararegolazione sia nell'articolato del Regolamento la giustizia algoritmica ha una sua autonomia e delle proprie regole, che saranno fondamentali a definire il perimetro dell'impiego degli algoritmi all'interno dei sistemi di giustizia.

2. Il percorso di regolazione *soft* dell'Unione europea sino all'approvazione del regolamento sull'intelligenza artificiale

2.1. I primi approcci strategici della Commissione europea

Come è naturale nel sistema istituzionale dell'Unione europea, prima attrice e protagonista indiscussa del processo regolatorio sull'intelligenza artificiale è stata la Commissione.

La prima iniziativa si colloca nel marzo 2018 quando l'organo esecutivo europeo, attraverso un comunicato stampa³⁴⁷, rendeva nota la volontà di istituire, attraverso una *call for applications*, un gruppo lavoro sull'intelligenza artificiale «*to gather expert input and rally a broad alliance of diverse stakeholders*». Il gruppo di esperti, si legge nel comunicato, avrebbe avuto il compito di assistere la Commissione europea nella costruzione di una «*broad and diverse community of stakeholders in a "European AI Alliance"*», supportare l'implementazione della «*European initiative on artificial intelligence*» e, soprattutto, presentare entro la fine dell'anno una bozza di linee guida per lo sviluppo e l'utilizzo etico dell'intelligenza artificiale sulla base dei diritti fondamentali dell'UE, in particolare tenendo in considerazione principi quali l'equità, la sicurezza, la trasparenza, il futuro del lavoro, la democrazia e, più in generale, l'impatto sui diritti scritti nella Carta dei diritti fondamentali.

³⁴⁷ Visionabile al seguente indirizzo: <https://ec.europa.eu/commission/presscorner/detail/en/>.

Nel mese seguente, nell'aprile 2018, la Commissione, in una comunicazione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni³⁴⁸, presentava una serie di misure «*to put artificial intelligence (AI) at the service of Europeans and boost Europe's competitiveness in this field*». In particolare, venivano proposti tre differenti approcci strategici, che manifestano l'attenzione mostrata sin da subito dall'Unione europea non solo verso il rispetto dei diritti fondamentali ma anche della necessità di investire maggiori risorse in un settore così competitivo.

Il primo approccio era volto ad aumentare il sostegno finanziario e incoraggiare gli investimenti nella ricerca tanto nel settore pubblico quanto nel settore privato, in ambiti considerati fondamentali come la sanità, la guida autonoma, l'agricoltura, l'attività industriale, l'energia, le tecnologie per l'Internet di prossima generazione, la sicurezza e la pubblica amministrazione (compresa la giustizia).

Il secondo indicava la necessità di prepararsi ai cambiamenti socioeconomici portati dall'intelligenza artificiale: in particolare, a causa dell'impatto dell'IA sul mondo del lavoro, venivano incoraggiati gli Stati membri a modernizzare i sistemi di istruzione e formazione (soprattutto per sviluppare competenze specifiche in materia di IA) e a sostenere le transizioni nel mercato del lavoro.

Il terzo approccio, sicuramente il più interessante ai nostri fini, mirava a fornire un quadro etico e legale adeguato, per far fronte soprattutto alle questioni legate alla «*liability or potentially biased decision-making*» dell'intelligenza artificiale. In questo senso la Commissione si impegnava a presentare entro la fine del 2018 degli orientamenti etici sullo sviluppo dell'IA, basati sulla Carta dei diritti fondamentali, tenendo conto di principi enucleati nell'art. 2 del Trattato sull'Unione europea³⁴⁹. Nella visione della Commissione, dunque, l'emersione di nuove tecnologie non significava necessariamente l'elaborazione di nuovi principi quanto piuttosto il loro

³⁴⁸ La Comunicazione è visionabile al seguente indirizzo: <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52018DC0237>.

³⁴⁹ Art. 2 TUE: «L'Unione si fonda sui valori del rispetto della dignità umana, della libertà, della democrazia, dell'uguaglianza, dello Stato di diritto e del rispetto dei diritti umani, compresi i diritti delle persone appartenenti a minoranze. Questi valori sono comuni agli Stati membri in una società caratterizzata dal pluralismo, dalla non discriminazione, dalla tolleranza, dalla giustizia, dalla solidarietà e dalla parità tra donne e uomini».

adattamento al nuovo contesto tecnologico, tenuto conto che «[l]’UE si basa su un quadro normativo solido ed equilibrato, che può definire i principi globali di un approccio sostenibile alla tecnologia in questione»³⁵⁰. Se caliamo questa impostazione nel solco della giustizia algoritmica, emergerà nel corso del lavoro che, in effetti, non sembra che la rivoluzione portata dall’intelligenza artificiale richiederà la formulazione di nuovi principi costituzionali (con una eccezione, come vedremo), quanto piuttosto un ripensamento e adattamento delle categorie costituzionali ben conosciute ai nuovi paradigmi tecnologici.

2.2. Un primo strumenti di regolazione soft: gli “Orientamenti etici per un’IA affidabile”

Finalmente nel giugno 2018 la Commissione, dando seguito alla comunicazione menzionata nel precedente paragrafo, comunicava l’istituzione di un *High Level Group on Artificial Intelligence*, composto da 52 esperti proveniente dall’accademia, dal mondo imprenditoriale e della società civile, con il compito di predisporre gli orientamenti etici per lo sviluppo di un’intelligenza artificiale affidabile.

Questi venivano infine pubblicati nell’aprile 2019 e rappresentano il primo documento di *soft law* approntato dall’Unione europea per la regolamentazione dell’intelligenza artificiale³⁵¹.

L’obiettivo degli orientamenti, indicato al principio del documento, è quello di promuovere un’IA affidabile, che deve basarsi su tre diverse componenti: legalità (l’IA deve ottemperare a tutte le leggi e ai regolamenti applicabili; eticità (l’IA deve assicurare l’adesione a principi e valori etici) e robustezza «dal punto di vista tecnico e sociale poiché, anche con le migliori intenzioni, i sistemi di IA possono causare danni non intenzionali»³⁵². Vengono poi indicati i destinatari delle linee

³⁵⁰ Comunicazione della Commissione, 25 aprile 2018.

³⁵¹ COMMISSIONE EUROPEA, Direzione generale delle Reti di comunicazione, dei contenuti e delle tecnologie, *Orientamenti etici per un’IA affidabile*, 2019 (il nome inglese è *Ethics Guidelines For Trustworthy AI*).

³⁵² *Ivi*, p. 6.

guida che sono «tutti i portatori di interessi nel campo dell'IA che progettano, sviluppano, distribuiscono, implementano, utilizzano l'IA o ne sono interessati, compresi, a titolo esemplificativo, imprese, organizzazioni, ricercatori, servizi pubblici, agenzie governative, istituzioni, organizzazioni della società civile, soggetti privati, lavoratori e consumatori»³⁵³.

Gli orientamenti forniscono una «base orizzontale» regolatoria, vale a dire delle indicazioni relative all'IA in generale, senza distinguere tra i possibili usi, da quelli più rischiosi a quelli meno rischiosi. Tuttavia, il gruppo di esperti è consapevole della necessità di dotarsi di una regolazione più specifica che distingua i rischi che gli *usi* – e non la tecnologia in sé – dell'intelligenza artificiale potrebbe comportare, posto che «situazioni differenti danno vita a sfide differenti. I sistemi di IA che offrono consigli musicali non suscitano le stesse preoccupazioni etiche dei sistemi di IA che propongono terapie mediche fondamentali».

Essendo orientamenti di natura generale, quindi, l'attuazione di «deve essere adattata all'applicazione di IA specifica», precisando quindi che dovrebbe essere valutata la necessità di un ulteriore approccio settoriale a integrazione del quadro di riferimento orizzontale. Queste indicazioni sono particolarmente rilevanti in quanto saranno accolte nell'impostazione generale del Regolamento sull'intelligenza artificiale che, come si vedrà, propone un approccio basato sul rischio che si concentra non sull'intelligenza artificiale in sé, come tecnologia astrattamente rischiosa, ma classificando dettagliatamente, e regolando, gli *usi* di tale tecnologia, in base al rischio che essi producono in relazione al rispetto dei diritti fondamentali.

Gli orientamenti si dividono in tre diversi capitoli, che si collocano su tre diversi livelli di astrazione, dal più astratto al più concreto. Nel primo capitolo “Basi per un'IA affidabile”, si delinea l'approccio basato sui diritti fondamentali e si identificano i principi etici che devono essere rispettati al fine di garantire l'eticità e la robustezza dell'intelligenza artificiale. A tal proposito si specifica che «il rispetto dei diritti fondamentali, nel quadro della democrazia e dello Stato di diritto,

³⁵³ *Ivi*, p. 7.

costituisce la base più incoraggiante per individuare principi e valori etici astratti che possono essere resi operativi nei sistemi di IA»³⁵⁴. Di tali diritti il documento ne menziona alcuni specificamente, come la dignità umana, la libertà individuale, l'uguaglianza e la non discriminazione. Tra essi compare altresì, e ai nostri fini è particolarmente rilevante perché mostra come l'Unione avesse già in mente i possibili utilizzi dell'IA nel campo della giustizia, il rispetto per la democrazia, della giustizia e dello Stato di diritto. Esso significa che «[i] sistemi di IA devono servire a mantenere e a promuovere i processi democratici e a rispettare la pluralità dei valori e delle scelte di vita degli individui. Essi non devono compromettere i processi democratici, la decisione umana o i sistemi di voto democratici». Inoltre, viene specificato che «nei sistemi di IA deve essere insito l'impegno a garantire di non operare con modalità che compromettano gli impegni di base su cui si fonda lo Stato di diritto, le leggi e i regolamenti obbligatori e a garantire il giusto processo e l'uguaglianza di fronte alla legge»³⁵⁵.

Le fondamenta su cui quindi costruire il castello regolatorio dell'IA sono rappresentate, secondo il Gruppo di esperti, dal rispetto dei diritti fondamentali. Compare inoltre per la prima volta nei documenti istituzionali dell'Unione europea la necessità che l'impiego dell'intelligenza artificiale non debba compromettere le garanzie processuali del *giusto processo* e la *decisione umana*, quindi anche quella del giudice. Vedremo che tali preoccupazioni confluiranno ampiamente nel Regolamento, e saranno anche il centro delle riflessioni che il presente lavoro intende proporre, proprio per la definizione di giustizia algoritmica che abbiamo accolto (strumenti algoritmici idonei a incidere sulla decisione del giudice).

Radicati nei diritti fondamentali, il documento individua quattro «imperativi etici» a cui gli operatori del settore dovrebbero attenersi: rispetto dell'autonomia umana, prevenzione dei danni, equità, esplicabilità. Gli ultimi due, che discendono dai diritti alla non discriminazione e alla solidarietà e, soprattutto, dai «diritti

³⁵⁴ *Ivi*, p. 11

³⁵⁵ *Ivi*, p. 12.

relativi alla giustizia (sanciti dall'articolo 47 CDFUE)»³⁵⁶, sono quelli ai nostri fini più interessanti.

Quanto all'equità, si esplicita infatti che la «dimensione procedurale dell'equità implica la capacità di impugnare le decisioni elaborate dai sistemi di IA e dagli esseri umani che li gestiscono e la possibilità di presentare un *ricorso efficace* contro di esse. A tal fine, l'organismo responsabile della decisione deve essere identificabile e i processi decisionali devono essere *spiegabili* (corsivi nostri)»³⁵⁷. Strettamente connessa all'equità è quindi l'esplicabilità e cioè che le «decisioni, per quanto possibile, devono poter essere spiegate a coloro che ne sono direttamente o indirettamente interessati. Senza tali informazioni, una decisione non può essere debitamente impugnata»³⁵⁸. E anche quando, in presenza di algoritmi i cui procedimenti rimangono opachi (cd. *black box*³⁵⁹), non è possibile spiegare del tutto l'*output* algoritmico, sarebbero «necessarie altre misure per garantire l'esplicabilità (ad esempio, la tracciabilità, la verificabilità e la comunicazione trasparente sulle capacità del sistema)».³⁶⁰ Come si può immaginare, la problematica della trasparenza delle decisioni algoritmiche e la loro spiegabilità ha interessato gli studiosi di ogni ambito e, nel contesto di decisioni algoritmiche assunte nel contesto giudiziario, assume un'importanza decisiva per poter considerare *costituzionalmente accettabile* soluzioni di questo tipo (si v. *infra*, cap. IX).

Nel secondo capitolo degli *Orientamenti*, i principi enunciati nel primo capitolo vengono tradotti in requisiti concreti per realizzare un'IA affidabile. I requisiti individuati, in un elenco definito esplicitamente non esaustivo, sono sette e sono rivolti «ai diversi portatori di interessi che partecipano al ciclo di vita dei sistemi di IA: sviluppatori, distributori, utenti finali e società in generale». Essi sono: 1) intervento e sorveglianza umani, 2) robustezza tecnica e sicurezza, 3) riservatezza

³⁵⁶ *Ivi*, p. 13. Anche l'autonomia umana ovviamente richiama necessariamente l'autonomia del giudice ma nel documento è intesa nel senso di assicurare che gli esseri umani non vengano manipolati dalle applicazioni di IA.

³⁵⁷ *Ivi*, p. 14.

³⁵⁸ *Ibidem*.

³⁵⁹ Se ne parlerà diffusamente *infra*, cap. VI § 2.

³⁶⁰ Commissione europea, *Orientamenti*, cit. p. 15.

e governance dei dati, 4) trasparenza, 5) diversità, non discriminazione ed equità, 6) benessere sociale e ambientale, e 7) *accountability*.

Le spiegazioni fornite nel documento rispetto ad alcuni di tali requisiti sono utili a delineare alcuni limiti e confini che dovrebbe avere l'impiego dell'IA nell'ambito della giustizia algoritmica e che, come si vedrà, non sono facili da implementare. Quanto all'intervento e alla sorveglianza umani, si specifica che «[g]li utenti dovrebbero essere in grado di adottare decisioni autonome e informate in merito ai sistemi di IA. Dovrebbero ricevere le conoscenze e gli strumenti per comprendere e interagire con i sistemi di IA a un livello soddisfacente e, ove possibile, essere in grado di valutare autonomamente o contestare il sistema in modo ragionevole³⁶¹. Tale requisito, come si vedrà, si lega indissolubilmente a uno degli elementi che ad oggi appaiono maggiormente ostativi all'impiego generalizzato dell'IA nel processo, vale a dire la quasi totale mancanza di formazione dei giudici sull'utilizzo dei sistemi automatizzati (si v. *infra* cap. IX).

Di enorme rilevanza ai nostri fini è il requisito della *trasparenza*, strettamente connesso al principio dell'esplicabilità indicato nel primo capitolo degli *Orientamenti*. Esso impone che «[i] set di dati e i processi che determinano la decisione del sistema di IA, compresi quelli di raccolta ed etichettatura dei dati, come pure gli algoritmi utilizzati, dovrebbero essere documentati secondo i migliori standard per consentire la tracciabilità e aumentare la trasparenza»³⁶² e, inoltre, che «[s]e un sistema di IA influisce considerevolmente sulla vita delle persone, dovrebbe sempre essere possibile richiedere una spiegazione adeguata del processo decisionale del sistema. Tale spiegazione dovrebbe essere tempestiva e adeguata alle competenze del portatore di interesse in questione»³⁶³.

Inoltre, come corollario al requisito della non discriminazione e equità, appare qui utile menzionare un elemento rilevante indicato nel documento, vale a dire la *partecipazione dei portatori di interessi*, secondo cui «è consigliabile consultare i portatori di interessi che possono essere direttamente o indirettamente interessati

³⁶¹ *Ivi*, p. 18.

³⁶² *Ivi*, p. 20.

³⁶³ *Ivi*, p. 21

dal sistema durante l'intero ciclo di vita di quest'ultimo». In questo senso, ci sembra di poter tradurre il principio anche nell'ambito della giustizia algoritmica: in numerose esperienze di impiego dell'IA in campo giudiziario i giudici si trovano tra le mani strumenti algoritmici progettati da organismi (quasi sempre privati) che spesso non possiedono nessuna competenza in materia di giustizia: il principio di coinvolgimento degli operatori interessati nella progettazione dei sistemi, che troveremo in altri strumenti normativi europei, garantirebbe invece una maggiore corrispondenza tra le funzionalità del sistema e il fine che dovrebbe perseguire, fine che senza dubbio non può essere definito da soggetti estranei al campo della giustizia.

Un ultimo requisito che appare qui utile menzionare è l'*accountability*, che comporta innanzitutto «la possibilità di valutare algoritmi, dati e processi di progettazione». Secondo gli *Orientamenti*, la verificabilità «non implica necessariamente che le informazioni sui modelli di business e sulla proprietà intellettuale relative al sistema di IA debbano essere sempre disponibili in modo aperto. La valutazione da parte di revisori interni ed esterni e la disponibilità delle relazioni di tale valutazione può contribuire all'affidabilità della tecnologia».

A questo ultimo riguardo, come spiegato *supra*, l'approccio del documento è di natura generale e orizzontale: probabilmente nel campo della giustizia, dove prevalgono esigenze di pubblicità, motivazione dei provvedimenti, diritto di difesa e diritto alla tutela giurisdizionale, sarebbe difficile immaginare una soluzione diversa da quella di una comprensione integrale dell'algoritmo impiegato (si v. diffusamente *infra*, cap. IX).

Nel terzo capitolo, infine, viene proposta una “bozza pilota” di *Lista di controllo* che le aziende produttrici di algoritmi dovrebbero prendere in considerazione per la valutazione dell'affidabilità del sistema algoritmico nelle fasi di sviluppo, distribuzione o utilizzo dell'IA e adattarla allo specifico caso d'uso del sistema³⁶⁴.

³⁶⁴ Tale *Lista* è stato rivista nel dicembre 2019 e poi di nuovo nel luglio 2020, a seguito di un processo di valutazione che ha coinvolto più di 350 stakeholder. Il feedback sulla *Lista* pilota è stato fornito in tre modi: un sondaggio online compilato dai partecipanti; la condivisione delle migliori pratiche su come realizzare un'IA affidabile attraverso la *European AI Alliance*; e una serie di interviste con gli operatori del settore. Si v. la versione definitiva della *Lista* al seguente link:

Secondo gli *Orientamenti*, partendo dai requisiti elencati nel capitolo II, le imprese dovrebbero porsi alcuni quesiti necessari prima di poter immettere nel mercato un sistema algoritmico.

Solo per fare qualche esempio chiarificatore, rispetto al requisito dell'intervento umano la *Lista di controllo* invita, nel caso di un sistema di IA che interagisce con il processo decisionale degli utenti finali umane, di farsi, tra le altre, la seguente domanda: «[i]n questi casi, sussiste il rischio che il sistema di IA influenzi l'autonomia umana interferendo in modo non intenzionale con il processo decisionale dell'utente finale?»³⁶⁵.

Ancora, quanto al requisito della trasparenza, «[è] stato valutato fino a che punto le decisioni prese dal sistema di IA e i loro risultati sono comprensibili?»³⁶⁶.

Nell'ultima parte del documento vengono infine individuati esempi in cui un'IA affidabile sarebbe idonea a rappresentare una grande opportunità per le sfide contemporanee della società ed esempi in cui invece l'utilizzo dell'IA desta grande preoccupazione³⁶⁷. Essendo delle categorie di rischio che verranno riprese nell'*AI Act*, l'analisi dei possibili utilizzi benefici o, al contrario, rischiosi verrà svolta in quella sede (si v. *infra* § 3). Ci limitiamo qui a dire che mentre nell'elenco – si badi, non esaustivo – degli *Orientamenti* non si fa cenno alcuno ai possibili rischi di impiego dell'IA nel campo della giustizia, cinque anni dopo nell'*AI Act*, come vedremo, gli impieghi nell'amministrazione della giustizia acquisiscono una propria autonomia normativa e una propria categoria di rischio.

Si ricollega a ciò il giudizio generale dei commentatori dopo l'approvazione degli *Orientamenti*: questi ultimi sono stati generalmente accolti con favore³⁶⁸ ma anche criticati da alcuni sono diversi punti di vista³⁶⁹, soprattutto per la debolezza

<https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>.

³⁶⁵ *Ivi*, p. 30.

³⁶⁶ *Ivi*, p. 35.

³⁶⁷ *Ivi*, pp. 39-43.

³⁶⁸ L. FLORIDI, *Establishing the Rules for Building Trustworthy AI*, in L. FLORIDI (a cura di) *Ethics, Governance, and Policies in Artificial Intelligence. Philosophical Studies Series*, vol 144. Springer, 2021, pp. 41-45.

³⁶⁹ Si cfr., *ex multis*, E. HICKMAN - M. PETRIN, *Trustworthy AI and Corporate Governance – The EU's Ethics Guidelines For Trustworthy Artificial Intelligence from a Company Law Perspective*, in *European Business Organization Law Review*, 2021, pp. 593-625.

dello strumento in sé e per la mancanza di indicazioni normative più precise a fronte dell'enunciazione di principi etici troppo generici.

Ora, senz'altro lo strumento sconta la non vincolatività ma occorre considerare che è intervenuto in un momento in cui ancora si discuteva animatamente sulla necessità di regolare rigidamente una materia di cui ancora non si conoscevano del tutto le implicazioni sulla società (per una ricognizione di tale dibattito si v. *supra*, alla nota introduttiva della presente parte) e, anche se non è stato particolarmente innovativo e dirompente, ha contribuito a creare un quadro di riferimento e un vocabolario concettuale comune in UE³⁷⁰, tanto che molte delle indicazioni contenute negli *Orientamenti* saranno riproposte del Regolamento. Inoltre, gli *Orientamenti*, sebbene non fossero finalizzati a dotare l'Unione di un quadro normativo adeguato, si sono inseriti comunque in una legislazione vincolante già applicabile a diversi impieghi dell'intelligenza artificiale (si pensi alla normativa sulla protezione dei dati personali o a quella sulla protezione dei consumatori di fronti a prodotti difettosi), contribuendo ulteriormente un'interpretazione della normativa *eticamente* orientata.

Infine, come detto, gli *Orientamenti* hanno effettivamente avuto un seguito normativo molto importante, che è culminato con l'approvazione dell'*AI Act*, che, come si vedrà, è stato influenzato non marginalmente dall'impostazione etica fornita dal gruppo di esperti scelti dalla Commissione europea.

2.3. La crescente necessità di un quadro normativo per l'IA: il Libro Bianco sull'intelligenza artificiale

Proprio dalle critiche mosse agli *Orientamenti* per la mancanza di un approccio normativo, e non solo etico, al tema dell'intelligenza artificiale, è nata una nuova iniziativa della Commissione europea che avrebbe influenzato massicciamente il successivo sviluppo regolatorio, l'approvazione del *Libro Bianco sull'intelligenza*

³⁷⁰ L. FLORIDI, *Establishing the Rules for Building Trustworthy AI*, cit., p. 42.

*artificiale*³⁷¹, in cui sostanzialmente veniva evidenziata l'inadeguatezza della vigente normativa a fronte dei problemi nuovi che l'IA ha posto sul tavolo.

Si legge infatti all'inizio del documento che la Commissione sostiene «un approccio normativo e orientato agli investimenti con il duplice obiettivo di promuovere l'adozione dell'IA e di affrontare i rischi associati a determinati utilizzi di questa nuova tecnologia. Lo scopo del presente libro bianco è definire le opzioni strategiche su come raggiungere tali obiettivi»³⁷². Gli elementi costitutivi del *Libro Bianco*, esplicitati nell'introduzione, sono i seguenti: la definizione di un quadro strategico che stabilisce misure per allineare gli sforzi a livello europeo, nazionale e regionale, tramite un partenariato tra il settore pubblico e privato «per conseguire un “ecosistema di eccellenza”» e la predisposizione di un futuro quadro normativo per l'IA in Europa, «che creerà un “ecosistema di fiducia” unico»³⁷³.

Quanto al primo elemento la Commissione indica la necessità di intensificare alcune azioni su più livelli, tra cui la collaborazione con gli Stati membri per un maggiore investimento del settore dell'IA; la creazione di più reti e sinergie tra diversi centri di ricerca europei sull'intelligenza artificiale migliorarne l'eccellenza, trattenere e attrarre i migliori ricercatori e sviluppare le migliori tecnologie; promuovere forme di partenariato pubblico-privato e promuovere l'impiego dell'IA nel settore pubblico, soprattutto in sanità e trasporti³⁷⁴.

Il secondo elemento costitutivo – la predisposizione di un quadro normativo – deriva esplicitamente dagli *Orientamenti*, Si legge infatti che «[i]n aggiunta a questa serie di orientamenti non vincolanti del gruppo di esperti ad alto livello, e in linea con gli orientamenti politici della presidente, un quadro normativo europeo chiaro rafforzerebbe la fiducia dei consumatori e delle imprese nell'IA e accelererebbe quindi l'adozione della tecnologia. Tale quadro normativo dovrebbe essere coerente con altre azioni volte a promuovere la capacità di innovazione e la competitività dell'Europa in questo settore»³⁷⁵.

³⁷¹ COMMISSIONE EUROPEA, *Libro bianco sull'intelligenza artificiale. Un approccio europeo all'eccellenza e alla fiducia*, 19 febbraio 2020.

³⁷² *Ivi*, p. 1.

³⁷³ *Ivi*, p. 3.

³⁷⁴ *Ivi*, pp. 5-10

³⁷⁵ *Ivi*, p. 11.

È inoltre interessante che fin da subito compaia esplicitamente – ed è la prima volta nei documenti europei che si occupano del tema – il ruolo critico che l'intelligenza artificiale potrebbe avere proprio nel campo giudiziario. La Commissione afferma infatti che il nuovo quadro normativo dovrebbe «garantire risultati ottimali sul piano sociale, ambientale ed economico nonché il rispetto della legislazione, dei principi e dei valori dell'UE. Ciò è particolarmente importante negli ambiti che interessano direttamente i diritti dei cittadini, ad esempio nel caso delle applicazioni di IA per le forze dell'ordine e la *magistratura*» (corsivo nostro)³⁷⁶.

Quindi per la prima volta, probabilmente anche per la maggiore risonanza che hanno raggiunto alcune applicazioni di intelligenza nel settore della giustizia nel mondo e in particolare negli Stati Uniti, l'Unione europea manifesta consapevolezza della *particolare importanza* del campo della giustizia algoritmica, il quale richiede la predisposizione di un quadro normativo adeguato.

Nell'enunciazione delle coordinate rispetto al quadro normativo, nel *Libro bianco* emergono molti elementi che troveranno conferma dell'*AI Act* – dalla classificazione dei sistemi in scale differenti di rischio, alla ricerca di una definizione ampia di intelligenza artificiale, a meccanismi di certificazione preventivi alla messa sul mercato – il che ci permette direttamente di analizzarli in quella sede.

Basti qui evidenziare le conclusioni del *White paper*, in cui la Commissione esalta il proprio ruolo nell'avviare, a partire da tale documento, «un'ampia consultazione della società civile, dell'industria e del mondo accademico negli Stati membri in merito a proposte concrete per un approccio europeo all'IA. Tali misure comprendono sia strumenti strategici per promuovere gli investimenti nella ricerca e nell'innovazione, favorire lo sviluppo delle competenze e sostenere l'adozione dell'IA da parte delle PMI, sia proposte relative a elementi essenziali di un futuro quadro normativo»³⁷⁷.

³⁷⁶ *Ibidem*.

³⁷⁷ *Ivi*, p. 28.

3. L'AI Act: la regolazione *hard* (con elementi *soft*) dell'intelligenza artificiale

3.1. La RoadMap: dalla Proposta della Commissione all'approvazione del Regolamento sull'intelligenza artificiale

Esula dai fini del presente lavoro sviscerare analiticamente ogni modifica e ogni fase della negoziazione che, dalla proposta della Commissione del 21 aprile 2021, ha portato di recente all'approvazione del "Regolamento del parlamento europeo e del consiglio che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n. 300/2008, (UE) n. 167/2013, (UE) n. 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (cd. regolamento sull'intelligenza artificiale o *AI Act*), pubblicato nella gazzetta ufficiale dell'Unione europea il 12 luglio 2024.

Tuttavia, appare utile dare conto brevemente della *roadmap* che si è di recente conclusa in quanto è stata una delle più lunghe e intricate negoziazioni della procedura legislativa europea.

Solo in seguito, quando si prenderà in considerazione la disciplina regolamentare rispetto alla giustizia algoritmica, verranno segnalate le importanti modifiche intervenute nelle negoziazioni tra Commissione, Parlamento e Consiglio che hanno sostanzialmente modificato il testo originario promosso nella Proposta.

Nell'aprile 2021 la Commissione ha pubblicato la propria Proposta di regolamento sull'intelligenza artificiale³⁷⁸, facendo crescere immediatamente l'interesse al tema sia nell'opinione pubblica sia nella letteratura accademica, che non ha mancato fin da subito di segnalare aspetti positivi e criticità del testo presentato³⁷⁹. Come già anticipato, l'analisi che si è fatta nei precedenti paragrafi

³⁷⁸ Commissione europea, *Proposta di regolamento del parlamento europeo e del consiglio che stabilisce regole armonizzate sull'intelligenza artificiale* (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'unione, 21 aprile 2021.

³⁷⁹ Si cfr. *ex multis*, per alcuni commenti usciti in seguito alla proposta nella letteratura italiana ed europea: L. FLORIDI, *The European Legislation on AI: A Brief Analysis of its Philosophical Approach*, cit., pp. 215-222; F. PIZZETTI, *La proposta di Regolamento sull'IA della Commissione Europea presentata il 21.4.2021 (COM (2021) 206 final) tra Mercato Unico e competizione digitale*

del percorso “normativo” antecedente alla Proposta è stato diretto a comprendere che il Regolamento non è “caduto dal cielo” ma invece si è inserito in un contesto di riflessioni durate anni: non a caso nel Regolamento stesso si afferma l’importanza di «ricordare gli orientamenti etici per un’IA affidabile del 2019 elaborati dall’AI HLEG indipendente nominato dalla Commissione» (*considerando n. 27*).

In ogni caso, a seguito dell’acquisizione dei pareri del Comitato economico e sociale europeo³⁸⁰, della Banca Centrale europea³⁸¹ e del Comitato delle regioni³⁸², dapprima il Consiglio ha adottato il proprio approccio generale (*General approach*) sulla proposta il 25 novembre 2022³⁸³ e, successivamente, il Parlamento europeo ha adottato la propria posizione negoziale il 14 giugno 2023³⁸⁴ dando così avvio alle negoziazioni tra i delegati dei tre organismi europei (il cd. *Trilogo*).

La fase di negoziazione si è rivelata particolarmente tortuosa per una serie di ragioni che vanno oltre la naturale complessità della materia da regolare, che già di per sé richiede tempi e sforzi piuttosto elevati.

Innanzitutto la messa sul mercato di Chat GPT, nel novembre 2022, ha immediatamente messo in luce il rischio di un’obsolescenza improvvisa di una regolamentazione dell’intelligenza artificiale: secondo il testo inizialmente proposto dalla Commissione, infatti, strumenti di IA cd. *general purpose* (con

globale, cit., pp. 591 ss.; C. CASONATO - B. MARCHETTI, *Prime osservazioni sulla proposta di regolamento dell’Unione Europea in materia di intelligenza artificiale*, in *BioLaw Journal – Rivista di BioDiritto*, n. 3/2021, pp. 415-437; M. EBERS, *Standardizing AI - The Case of the European Commission’s Proposal for an Artificial Intelligence Act*, in L.A. DI MATTEO - M. CANNARSA - C. PONCIBÒ (a cura di), *The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics*, Cambridge, 2022; F. DONATI, *Diritti fondamentali e algoritmi nella proposta di regolamento sull’intelligenza artificiale*, in A. PAJNO - F. DONATI - A. PERRUCCI, *Intelligenza artificiale e diritto: una rivoluzione?*, volume I, 2022, pp. 111-126.

³⁸⁰ Parere del 22 settembre 2021, visionabile al seguente link: <https://www.eesc.europa.eu/en/our-work/opinions-information-reports/opinions/regulation-artificial-intelligence>.

³⁸¹ Parere del 29 dicembre 2021, visionabile al seguente link: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021AB0040>.

³⁸² Parere del 2 dicembre 2021, visionabile al seguente link: <https://cor.europa.eu/EN/our-work/Pages/OpinionTimeline.aspx?opId=CDR-2682-2021>.

³⁸³ Si cfr. il testo al seguente indirizzo: <https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/en/pdf>.

³⁸⁴ Si cfr. il testo al seguente indirizzo: https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf.

finalità generali) come Chat GPT sfuggivano alle previsioni del regolamento e alla definizione di IA *ivi* proposta³⁸⁵.

Ciò ha rallentato notevolmente i lavori, spingendo il Parlamento a proporre nei propri emendamenti – e poi far approvare – un’intera sezione dell’articolato del regolamento (Artt. 50-55) proprio a strumenti di IA “con finalità generali” come ChatGPT³⁸⁶.

Inoltre, in sede di negoziazione, attraverso una corposa attività di *lobbying* numerosi operatori economici, sia americani³⁸⁷ sia europei³⁸⁸, hanno esercitato notevoli pressioni sui legislatori e sui governi europei per allineare il regolamento *in fieri* ai propri interessi, soprattutto per quanto concerne i modelli *general purpose* come Chat GPT³⁸⁹.

Ciò detto, finalmente il 9 dicembre 2023 il Parlamento e il Consiglio hanno raggiunto l’accordo politico provvisorio sul testo emendato e approvato il 2 febbraio 2024. Infine, a seguito dell’adozione del testo da parte del Parlamento il 13 marzo 2024 e del Consiglio il 21 maggio 2024, il Regolamento è stato pubblicato nella gazzetta ufficiale dell’Unione europea il 12 luglio 2024.

3.2. Profili generali del Regolamento

L’analisi puntuale delle previsioni del regolamento – lungo e complesso, con 180 considerando, 113 articoli e 13 allegati³⁹⁰ – è già stata effettuata in altri scritti³⁹¹

³⁸⁵ N. HELBERGER, - N. DIAKOPOULOS, *ChatGPT and the AI Act*, in *Internet Policy Review*, 12 (1), 2023.

³⁸⁶ *Op. cit.*

³⁸⁷ Come rivela IL CORPORATE EUROPE OBSERVATORI, si v. *Big Tech lobbying is derailing the AI Act*, in *corporateeurope.org*, 24 novembre 2023.

³⁸⁸ È divenuto noto il caso di una società di intelligenza artificiale, Mistral AI, che durante le negoziazioni sul regolamento ha fatto grande attività di *lobbying* per allentare alcune norme del regolamento e subito dopo ha stretto una partnership con Microsoft, sollevando dubbi sul fatto che quell’attività fosse (anche) finalizzata a tranquillizzare la *big tech* americana. Si cfr. <https://guerredirete.substack.com/p/guerre-di-rete-gelo-mistral>.

³⁸⁹ E. FALLETTI, *L’Artificial Intelligence Act Proposal e la regolamentazione degli algoritmi predittivi: luci e ombre*, in *CERIDAP*, 4/2023, p. 174.

³⁹⁰ Per avere un’idea, basta compararlo con un altro Regolamento particolarmente elaborato, il GDPR, con “soli” 171 considerando, 99 articoli e nessun allegato.

³⁹¹ Si cfr. *ex multis*, G. LO SAPIO, *L’Artificial Intelligence Act e la prova di resistenza per la legalità algoritmica*, in *Federalismi*, n. 16/2024, pp. 265-290; E. FALLETTI, *L’Artificial Intelligence*

e, in questa sede, non riteniamo di doverla riproporre date le finalità del presente lavoro. Qui forniremo invece una concisa panoramica della struttura e degli obiettivi dell'*AI Act* che invece – questi sì – ci saranno utili per collocare nel più ampio contesto della regolazione dell'IA le previsioni attinenti alla giustizia algoritmica.

Ora, lo scopo del regolamento, indicato nel primo articolo, mostra le diverse esigenze dell'Unione di cui abbiamo già discusso: «migliorare il funzionamento del mercato interno e promuovere la diffusione di un'intelligenza artificiale», che deve essere, secondo l'impostazione *human centered* adottata anche nel GDPR³⁹², «antropocentrica³⁹³ e affidabile», garantendo da una parte la protezione dei diritti fondamentali e, dall'altra, promuovendo l'innovazione (art. 1 del Regolamento). A tal riguardo, nel *considerando n. 6* si specifica che «è essenziale che l'IA e il suo quadro normativo siano sviluppati conformemente ai valori dell'Unione sanciti dall'articolo 2 del trattato sull'Unione europea (TUE), ai diritti e alle libertà fondamentali sanciti dai trattati e, conformemente all'articolo 6 TUE³⁹⁴, alla Carta. Come prerequisito, l'IA dovrebbe essere una tecnologia antropocentrica. Dovrebbe fungere da strumento per le persone, con il fine ultimo di migliorare il benessere degli esseri umani».

Act Proposal e la regolamentazione degli algoritmi predittivi: luci e ombre, cit. pp. 173-217; M. HAATAJA - J. BRYSON, *The European Parliament's AI regulation: should we call it progress?*, in *Amicus Curiae*, Series 2, Vol 4, No 3, 2023, pp. 707-718.

³⁹² Si v. considerando n. 4 del GDPR in cui si dice che «il trattamento dei dati personali dovrebbe essere al servizio dell'uomo».

³⁹³ Secondo Luciano Floridi tale termine sarebbe in realtà obsoleto e ambiguo, evocando il ruolo invasivo, poiché centrale, dell'uomo nel rapporto con l'ambiente. L'Autore riconosce però che, nonostante il termine usato impropriamente, il Regolamento è attento al tema ambientale. Si cfr. L. FLORIDI, *The European Legislation on AI: a Brief Analysis of its Philosophical Approach*, in *Philosophy & Technology*, 34, 2021, pp. 215-222.

³⁹⁴ Art. 6 TUE: «1. L'Unione riconosce i diritti, le libertà e i principi sanciti nella Carta dei diritti fondamentali dell'Unione europea del 7 dicembre 2000, adattata il 12 dicembre 2007 a Strasburgo, che ha lo stesso valore giuridico dei trattati. Le disposizioni della Carta non estendono in alcun modo le competenze dell'Unione definite nei trattati. I diritti, le libertà e i principi della Carta sono interpretati in conformità delle disposizioni generali del titolo VII della Carta che disciplinano la sua interpretazione e applicazione e tenendo in debito conto le spiegazioni cui si fa riferimento nella Carta, che indicano le fonti di tali disposizioni.

2. L'Unione aderisce alla Convenzione europea per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali. Tale adesione non modifica le competenze dell'Unione definite nei trattati.

3. I diritti fondamentali, garantiti dalla Convenzione europea per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali e risultanti dalle tradizioni costituzionali comuni agli Stati membri, fanno parte del diritto dell'Unione in quanto principi generali».

Quanto all'ambito di applicazione, così come era stato fatto con il GDPR, il regolamento ha un'estensione applicativa *extra-territoriale*: indipendentemente dalla sede delle imprese produttrici, se all'interno o al di fuori dell'Unione europea, esse dovranno comunque rispettare le regole previste dall'*AI Act* quando «immettono sul mercato o mettono in servizio sistemi di IA o immettono sul mercato modelli di IA per finalità generali nell'Unione». Quanto agli utilizzatori (*deployers*) il Regolamento si applica sia se essi siano situati all'interno dell'Unione sia all'esterno, purché «l'output prodotto dal sistema di IA sia utilizzato nell'Unione» (Art. 2 del Regolamento)³⁹⁵. Si ritiene che tali previsioni, che manifestano la strategia dell'Unione di porsi come *leader* nel contesto normativo dell'IA³⁹⁶, potrebbero portare al cosiddetto “effetto Bruxelles”, già visto con l'emanazione del GDPR, in base al quale il mercato e il potere normativo dell'UE creano standard globali che vengono seguiti dalle aziende tecnologiche e, a loro volta, adottati da Paesi terzi³⁹⁷.

La definizione di IA – come già visto molto ampia (si v. *supra*, cap. introduttivo § 2) – rinuncia a calarsi nel dibattito filosofico più che pratico di cosa sia un'intelligenza artificiale, descrivendo piuttosto che cosa sia un “sistema di IA”. Per il Regolamento, secondo una definizione in realtà già adottata dall'OCSE³⁹⁸, un

³⁹⁵ Tra le più importanti esclusioni dall'ambito di applicazione del regolamento si trovano invece i sistemi di IA impiegati per scopi militari, di difesa o di sicurezza nazionale e i sistemi di IA sviluppati e messi in servizio al solo scopo di ricerca e sviluppo scientifici.

³⁹⁶ G. FINOCCHIARO, *La regolazione dell'intelligenza artificiale*, in *Riv. Trim. Dir. Pubbl.*, 4, 2022, p. 1089.

³⁹⁷ Sul tema, ancora, A. BRADFORD, *Brussels Effect. How the European Union rules the world*, cit. Già dopo l'approvazione della proposta alcuni Stati hanno avviato un processo di regolamentazione dell'IA. Il Canada ha introdotto il *Digital Charter Implementation Act* nel 2022. Nel dicembre 2022, il Brasile ha avviato l'elaborazione del disegno di legge n. 2338 al Senato, volto a regolamentare l'IA. Su queste e altre iniziative si v. L. COTINO HUESO, *The Council of Europe's Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*, in *CERIDAP*, 9 settembre 2024.

³⁹⁸ Nella *Recommendation Of The Council On Artificial Intelligence* del 22 maggio 2019, un *AI system* veniva definito come «*a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy*». Tale Raccomandazione è stata aggiornata di recente (3 maggio 2024), e ha adattato la nuova definizione a quella comparsa nell'*AI Act*, descrivendo un *AI system* come «*a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment*» ricalcando sostanzialmente la definizione del Regolamento europeo.

sistema di IA è «un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali»³⁹⁹.

La scelta di non definire l'IA quanto piuttosto un sistema di IA è coerente con l'approccio regolatorio seguito dal Regolamento, che ha deciso di non normare la tecnologia in quanto tale – per evitare rischi di obsolescenza a fronte dell'evoluzione quotidiana di essa – quanto piuttosto gli usi, più o meno rischiosi, dell'intelligenza artificiale. In altre parole, l'IA potrà sì cambiare forma ed evolversi, ma il suo uso in alcuni ambiti è considerato comunque un uso vietato o rischioso, secondo le previsioni del Regolamento che adesso brevemente si segnaleranno.

A tal proposito, come già accennato, il legislatore europeo ha seguito un approccio “basato sul rischio” (*risk-based approach*)⁴⁰⁰, descritto nello stesso regolamento in cui si specifica che «al fine di introdurre un insieme proporzionato ed efficace di regole vincolanti per i sistemi di IA è opportuno avvalersi di un approccio basato sul rischio definito in modo chiaro. Tale approccio dovrebbe adattare la tipologia e il contenuto di dette regole all'intensità e alla portata dei rischi che possono essere generati dai sistemi di IA. È pertanto necessario vietare determinate pratiche di IA inaccettabili, stabilire requisiti per i sistemi di IA ad alto rischio e obblighi per gli operatori pertinenti, nonché obblighi di trasparenza per determinati sistemi di IA» (*considerando n. 26*).

³⁹⁹ Per quanto ampia, nella Proposta della Commissione la definizione di sistema di IA era più circoscritta, essendo questo definito «un software sviluppato con una o più delle tecniche e degli approcci elencati nell'allegato I, che può, per una determinata serie di obiettivi definiti dall'uomo, generare output quali contenuti, previsioni, raccomandazioni o decisioni che influenzano gli ambienti con cui interagiscono». In virtù di questa definizione, come già abbiamo chiarito, probabilmente sistemi come i *general purposes AI systems* (come Chat GPT) avrebbero potuto sfuggire a tale categorizzazione, non avendo propriamente obiettivi e finalità esplicitamente definiti dal progettista.

⁴⁰⁰ J. BLACK, *Risk-based regulation: choices, practices and lessons being learnt. Risk and Regulatory Policy*, in *Improving the Governance of Risk*, OECD, Paris, 2010. Si v. anche A. SIMONCINI, *Verso la regolamentazione della Intelligenza Artificiale. Dimensioni e governo*, cit., p. 413.

Seguendo tale approccio, all'art. 5 del Regolamento vengono indicate le pratiche di intelligenza artificiale *vietate*, poiché inaccettabili e «contrarie ai valori dell'Unione relativi al rispetto della dignità umana, alla libertà, all'uguaglianza, alla democrazia e allo Stato di diritto e ai diritti fondamentali sanciti dalla Carta, compresi il diritto alla non discriminazione, alla protezione dei dati e alla vita privata e i diritti dei minori» (*considerando n. 28*).

Nello specifico, l'*AI Act* vieta: sistemi di sfruttamento della vulnerabilità delle persone o sistemiche utilizzano tecniche manipolative con finalità di distorcere il comportamento di una persona⁴⁰¹; sistemi di IA che assegnano un punteggio sociale sulla base dei comportamenti e delle caratteristiche di una persona⁴⁰²; sistemi per effettuare valutazioni del rischio relative a persone fisiche al fine di prevedere il rischio che una persona fisica commetta un reato, unicamente sulla base della profilazione di una persona fisica o della valutazione dei tratti e delle caratteristiche della personalità (su questo, attenendo al contesto di questo lavoro, ci soffermeremo maggiormente *infra*, § 4.3), sistemi di IA che ampliano banche dati di riconoscimento facciale tramite *scraping* non mirato di immagini facciali da internet o da telecamere a circuito chiuso⁴⁰³; sistemi di riconoscimento delle

⁴⁰¹ J. BERMUDEZ - R. NYRUP - S. DETERDING - C. MOUGENOT - L. MORADBAKHTI - F. YOU - R. CALVO, *What is a subliminal technique?*, in *Ieee Ethics-2023 Conference Proceedings*, 2023.

⁴⁰² Il tema del punteggio sociale (*social scoring*), impiegato in alcune parti della Cina, è molto sentito dal legislatore europeo in quanto «[i] sistemi di IA che permettono ad attori pubblici o privati di attribuire un punteggio sociale alle persone fisiche possono portare a risultati discriminatori e all'esclusione di determinati gruppi. Possono inoltre ledere il diritto alla dignità e alla non discriminazione e i valori di uguaglianza e giustizia. Tali sistemi di IA valutano o classificano le persone fisiche o i gruppi di persone fisiche sulla base di vari punti di dati riguardanti il loro comportamento sociale in molteplici contesti o di caratteristiche personali o della personalità note, inferite o previste nell'arco di determinati periodi di tempo. Il punteggio sociale ottenuto da tali sistemi di IA può determinare un trattamento pregiudizievole o sfavorevole di persone fisiche o di interi gruppi in contesti sociali che non sono collegati ai contesti in cui i dati sono stati originariamente generati o raccolti, o a un trattamento pregiudizievole che risulta ingiustificato o sproporzionato rispetto alla gravità del loro comportamento sociale. I sistemi di IA che comportano tali pratiche inaccettabili di punteggio aventi risultati pregiudizievoli o sfavorevoli dovrebbero pertanto essere vietati. Tale divieto non dovrebbe pregiudicare le pratiche lecite di valutazione delle persone fisiche effettuate per uno scopo specifico in conformità del diritto dell'Unione e nazionale». Per una panoramica dei meccanismi di *Social credit system* in Cina si cfr. T. JESSUP, *'I think therefore I AM'? - Artificial intelligence and the Chinese Social Credit System: A Human Rights Critical Analysis*, 17 maggio 2021.

⁴⁰³ Lo *Scraping* è la raccolta incontrollata dei dati che sono presenti sulla rete per la successiva raccolta e riutilizzo. Si cfr. M. A. KHDAR, *Web Scraping or Web Crawling: State of Art, Techniques, Approaches and Application*, in *International Journal of Advances in Soft Computing & Its Applications*, 13 (3), 2021.

emozioni utilizzati sul luogo di lavoro e negli istituti di formazioni, eccetto per motivi medici o di sicurezza; sistemi di categorizzazione biometrica che classificano le persone secondo i loro dati biometrici (salvo eccezioni come le attività di contrasto) e sistemi di identificazione biometrica in tempo reale in spazi accessibili al pubblico (salvo alcune eccezioni che comunque devono rispettare alcuni requisiti).

Il Capo III rappresenta la colonna portante del Regolamento e la parte più estesa dell'articolato legislativo (artt. 6-49), e disciplina i sistemi classificati «ad alto rischio», sistemi che «hanno un impatto nocivo significativo sulla salute, la sicurezza e i diritti fondamentali delle persone nell'Unione» (*considerando n. 46*), tra i quali il Regolamento menziona espressamente «la dignità umana, il rispetto della vita privata e della vita familiare, la protezione dei dati personali, la libertà di espressione e di informazione, la libertà di riunione e di associazione e il diritto alla non discriminazione, il diritto all'istruzione, la protezione dei consumatori, i diritti dei lavoratori, i diritti delle persone con disabilità, l'uguaglianza di genere, i diritti di proprietà intellettuale, il diritto a un ricorso effettivo e a un giudice imparziale, i diritti della difesa e la presunzione di innocenza e il diritto a una buona amministrazione» (*considerando n. 48*).

Per tali sistemi si prevedono una serie di requisiti e corrispondenti obblighi per i fornitori prima della messa sul mercato, nonché obblighi per gli utilizzatori (*deployers*), questi ultimi che verranno approfonditi per quanto di interesse nel corso del lavoro.

In generale, quanto ai requisiti obbligatori previsti dall'*AI Act* per i sistemi ad alto rischio (artt. 8-15) si prevede, tra gli altri, un sistema di gestione del rischio che richiede un riesame e un aggiornamento costanti e sistematici del sistema di IA (art. 9), cautele e limitazioni rispetto alla raccolta e alla *governance* dei dati per l'addestramento dei sistemi (art. 10), la presenza di una documentazione tecnica prima dell'immissione sul mercato che dimostri il rispetto dei requisiti richiesti dal Regolamento (art. 11), requisiti di tracciabilità (art. 12), trasparenza (art. 13), sorveglianza umana (art. 14), accuratezza e robustezza (art. 15).

Gli obblighi dei fornitori di sistemi ad alto rischio rappresentano il cuore della procedura di produzione e commercio dei sistemi di IA⁴⁰⁴: essi dovranno garantire il rispetto dei requisiti prescritti (art. 16), attraverso una valutazione di conformità (art. 43) oltre che obblighi di registrazione e di apposizione della marcatura CE (artt. 16, 48 e 49).

L'elencazione dei sistemi classificati ad alto rischio è inserita nell'Allegato III del Regolamento, cui fa rinvio l'art. 6, che comprende otto diverse categorie di sistemi di IA: i sistemi biometrici, quando non sono vietati a norma dell'art. 5 dell'*AI Act*; sistemi impiegati in infrastrutture critiche come traffico stradale o nella fornitura di acqua, gas, riscaldamento o elettricità; sistemi impiegati nel settore dell'istruzione e della formazione professionale per determinare l'accesso agli istituti, valutare i risultati di apprendimento o monitorare i comportamenti; sistemi impiegati per la gestione dei lavoratori (assunzione e successivi elementi del rapporto di lavoro); sistemi impiegati per determinare l'accesso a servizi e a prestazioni pubblici e privati essenziali (come, ad esempio, l'assistenza sanitaria o l'affidabilità creditizia); sistemi per l'attività di contrasto come la cd. polizia predittiva, il poligrafo o *risk assessment tool* quando non sono vietati ai sensi dell'art. 5, sistemi impiegati nella gestione della migrazione, dell'asilo e del controllo delle frontiere; infine, i sistemi di IA, nel settore dell'amministrazione della giustizia, e dei processi democratici, quali quelli usati per influenzare l'esito delle elezioni o il comportamento di voto delle persone.

L'articolo 6 prevede inoltre alcune deroghe rispetto alla classificazione indicata – che è alla base del perimetro d'indagine di questo lavoro – quando il sistema, pur rientrando astrattamente nella categoria *high risk*, «non presenta un rischio significativo di danno per la salute, la sicurezza o i diritti fondamentali delle persone fisiche anche nel senso di non influenzare materialmente il risultato del processo decisionale» (art. 6, § 3). La regolazione *hard* dei sistemi classificati ad alto rischio rischierebbe di risultare presto obsoleta a fronte della continua evoluzione della tecnologia: per questa ragione nel Regolamento sono previsti diversi elementi di

⁴⁰⁴ E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit., p. 311.

flessibilità regolatoria: l'art. 7 prevede per esempio la possibilità per la Commissione, a determinate condizioni, di modificare l'allegato III (sistemi *high risk*) aggiungendo o modificando i casi di uso dei sistemi di IA.

Il capo IV è dedicato agli obblighi di *trasparenza* che i fornitori devono garantire per sistemi di IA non classificati come ad alto rischio ma «destinati a interagire direttamente con le persone fisiche»⁴⁰⁵, i quali devono essere «progettati e sviluppati in modo tale che le persone fisiche interessate siano informate del fatto di stare interagendo con un sistema di IA, a meno che ciò non risulti evidente dal punto di vista di una persona fisica ragionevolmente informata, attenta e avveduta, tenendo conto delle circostanze e del contesto di utilizzo» (art. 50 § 1); o sistemi di IA «che generano contenuti audio, immagine, video o testuali sintetici»⁴⁰⁶, i quali devono garantire che gli *output* «siano marcati in un formato leggibile meccanicamente e rilevabili come generati o manipolati artificialmente».

Il Capo V (artt. 51-56) riguarda i modelli di IA con finalità generali (*general purpose AI*) e non era presente nell'originaria proposta della Commissione: essi sono utilizzati per scopi per i quali non sono stati intenzionalmente e specificamente progettati e in cui la tecnologia viene successivamente integrata in un altro sistema ad alto rischio. Come già segnalato, con l'entrata nel mercato di strumenti come Chat GPT, il Parlamento ha fortemente voluto aggiungere una regolamentazione specifica di tali strumenti, che presentano caratteristiche del tutto autonome. Se tali sistemi superano una certa potenza di calcolo, i fornitori dovranno adempiere ad alcune procedure, che diventano più stringenti se i modelli sono classificati, secondo degli standard previsti nel regolamento, «modelli di IA per finalità generali con rischio sistemico»⁴⁰⁷ (art. 55).

⁴⁰⁵ Sono per esempio i *chatbot*.

⁴⁰⁶ Si pensi alle funzionalità dell'ultima versione di ChatGPT o, quanto alla generazione di immagini, a MidJourney.

⁴⁰⁷ E cioè «un rischio specifico per le capacità di impatto elevato dei modelli di IA per finalità generali, avente un impatto significativo sul mercato dell'Unione a causa della sua portata o di effetti negativi effettivi o ragionevolmente prevedibili sulla salute pubblica, la sicurezza, i diritti fondamentali o la società nel suo complesso, che può propagarsi su larga scala lungo l'intera catena del valore» (considerando nn. 111 ss.).

Il Capo VI (artt. 57-63) si occupa delle misure a sostegno dell'innovazione, obbligando gli Stati membri alla creazione di spazi di sperimentazione normativa per l'IA (cd. *sandbox*) e di prova in condizioni reali, che garantiscono «un ambiente controllato che promuove l'innovazione e facilita lo sviluppo, l'addestramento, la sperimentazione e la convalida di sistemi di IA innovativi per un periodo di tempo limitato prima della loro immissione sul mercato o della loro messa in servizio conformemente a un piano specifico dello spazio di sperimentazione concordato tra i fornitori o i potenziali fornitori e l'autorità competente» (art. 57 § 5).

La architettura di *Governance* è delineata nel Capo VII dell'*AI Act*: si prevede l'istituzione di un *Ufficio per l'IA* (art. 64) con la «missione di sviluppare competenze e capacità dell'Unione nel settore dell'IA e di contribuire all'attuazione del diritto dell'Unione in materia di IA» (*considerando n. 148*); di un *Consiglio europeo per l'IA* composto dai rappresentanti degli Stati membri (art. 65) con il compito di fornire consulenza e assistenza alla Commissione e agli Stati membri al fine di agevolare l'applicazione coerente ed efficace del regolamento (art. 66); di un *Forum consultivo* nominato dalla Commissione con una composizione equilibrata tra interessi dell'industria (PMI e start-up) e interessi di altro tipo (società civile e accademia) e con il compito di produrre pareri, raccomandazioni e altri contributi (art. 67); di un *Gruppo di esperti scientifici indipendenti* con compiti di consulenza e sostegno all'Ufficio per l'IA, in particolare per quanto riguarda la segnalazione dei rischi sistemici e la classificazione e sperimentazione dei modelli di IA con finalità generali (artt. 68-69). Oltre a tali organi, il regolamento prevede per ogni Stato membro l'obbligo di designare almeno un'autorità di notifica e almeno un'autorità di vigilanza del mercato⁴⁰⁸ con il compito, in condizioni di indipendenza e imparzialità, di garantire l'applicazione e l'attuazione delle norme.

⁴⁰⁸ Nel disegno di legge in materia di IA approvato dal governo italiano il 23 aprile 2024, l'art. 18 identifica le *Autorità nazionali* nell'attuazione del Regolamento europeo. Tale articolo prevede che « 1. Al fine di garantire l'applicazione e l'attuazione della normativa nazionale e dell'Unione europea in materia di intelligenza artificiale, l'Agenzia per l'Italia digitale (AgID) e l'Agenzia per la cybersicurezza nazionale (ACN) sono Autorità nazionali per l'intelligenza artificiale.

Conseguentemente, nel rispetto dei principi di cui alla presente legge e ferme restando le funzioni già rispettivamente attribuite:

a) l'AgID è responsabile di promuovere l'innovazione e lo sviluppo dell'intelligenza artificiale, fatto salvo quanto previsto dalla lettera b). L'AgID provvede altresì a definire le procedure e ad

L'ultima parte del regolamento è dedicata: ai mezzi di ricorso (artt. 85-87), su cui ci soffermeremo diffusamente *infra* (cap. IX); all'elaborazione di codici di condotta per promuovere l'applicazione volontaria dei requisiti dei sistemi *high risk* anche ai sistemi non così classificati (art. 95); a misure regolatorie *soft* da affiancare alla normativa regolamentare, tra cui la possibilità della Commissione di emanare orientamenti per l'attuazione del regolamento (art. 96), nonché atti delegati (*secondary acts*) per aggiornare alcune parti del Regolamento all'evoluzione della tecnologia; le sanzioni per il mancato rispetto del Regolamento (artt. 99-101) e le disposizioni finali.

Tra queste ultime assumono importanza l'art. 112 sulle possibilità di riesame del Regolamento da parte della Commissione e l'art. 113 sulla sua entrata in vigore e sulla sua applicazione, che presentano alcune peculiarità.

Quanto all'art. 112, esso prevede che la Commissione possa valutare la necessità di modificare l'elenco stabilito nell'allegato III e l'elenco di pratiche di IA vietate di cui all'articolo 5, trasmettendo i relativi risultati della valutazione al Parlamento europeo e al Consiglio. Nei paragrafi successivi sono previsti una serie di obblighi periodici (ogni quattro anni) da parte della Commissione di riferire al Parlamento e al Consiglio rispetto ad alcuni aspetti di valutazione e implementazione del

esercitare le funzioni e i compiti in materia di notifica, valutazione, accreditamento e monitoraggio dei soggetti incaricati di verificare la conformità dei sistemi di intelligenza artificiale, secondo quanto previsto dalla normativa nazionale e dell'Unione europea;

b) l'ACN, anche ai fini di assicurare la tutela della cybersicurezza, come definita dall'articolo 1, comma 1, del decreto-legge 14 giugno 2021, n. 82, convertito, con modificazioni, dalla legge 4 agosto 2021, n. 109, è responsabile per la vigilanza, ivi incluse le attività ispettive e sanzionatorie, dei sistemi di intelligenza artificiale, secondo quanto previsto dalla normativa nazionale e dell'Unione europea. L'ACN è, altresì, responsabile per la promozione e lo sviluppo dell'intelligenza artificiale relativamente ai profili di cybersicurezza;

c) l'AgID e l'ACN, ciascuna per quanto di rispettiva competenza, assicurano l'istituzione e la gestione congiunta di spazi di sperimentazione finalizzati alla realizzazione di sistemi di intelligenza artificiale conformi alla normativa nazionale e dell'Unione europea, sentito il Ministero della difesa per gli aspetti relativi ai sistemi di intelligenza artificiale impiegabili in chiave duale.

2. Le Autorità nazionali per l'intelligenza artificiale assicurano il coordinamento e la collaborazione con le altre pubbliche amministrazioni e autorità indipendenti, nonché ogni opportuno raccordo tra loro per l'esercizio delle funzioni di cui al presente articolo. A quest'ultimo fine, presso la Presidenza del Consiglio dei ministri è istituito un Comitato di coordinamento, composto dai direttori generali delle due Agenzie e dal Capo del Dipartimento per la trasformazione digitale della stessa Presidenza.

3. Restano ferme le competenze, i compiti e i poteri del Garante per la protezione dei dati personali».

Regolamento. Se necessario, la Commissione presenta proposte di modifica tenendo conto, in particolare, degli sviluppi delle tecnologie e dell'effetto dei sistemi di IA sulla salute, sulla sicurezza e sui diritti fondamentali, nonché alla luce dei progressi della società dell'informazione (§ 10, sono questi i più importanti elementi *soft* nell'impianto regolatorio europeo).

Quanto infine all'art. 113, si prevede che il Regolamento entri in vigore venti giorni dopo la pubblicazione sulla Gazzetta ufficiale dell'Unione. Tuttavia, l'applicazione del regolamento avrà tempi "scaglionati" e diversi. In particolare, il regolamento si applica a decorrere dal 2 agosto 2026. Tuttavia: I capi I e II (pratiche vietate) si applicano a decorrere dal 2 febbraio 2025; Il capo III, sezione 4 (la designazione delle autorità dei notifica da parte degli Stati Membri, il capo V (i modelli di IA per finalità generali), il capo VII (l'istituzione della banca dati UE per i sistemi ad alto rischio), il capo XII (le sanzioni) e l'articolo 78 (riservatezza dei dati trattati) si applicano a decorrere dal 2 agosto 2025, ad eccezione dell'articolo 101 (sanzioni pecuniarie per i fornitori di modelli di IA per finalità generali); L'articolo 6, paragrafo 1 (classificazione dei sistemi ad alto rischio), e i corrispondenti obblighi di cui al Regolamento, si applicano a decorrere dal 2 agosto 2027.

4. La giustizia algoritmica nel prisma dell'*AI Act*

Visti i profili generali del Regolamento, nel presente paragrafo verranno prese in considerazione le previsioni specifiche che attengono al tema del presente lavoro, nel tentativo anche di comprendere quali sono stati gli *step* che dalla Proposta della Commissione hanno condotto alla versione finale del Regolamento approvato dal legislatore europeo.

Oltre alle previsioni esplicitamente dedicate all'amministrazione della giustizia, si considereranno anche le norme riguardanti gli strumenti di IA impiegati dalle autorità pubbliche per determinare il rischio di commissione di un reato o il rischio di recidiva degli individui. Tali sistemi costituiscono nell'*AI Act* una categoria autonoma in quanto non sono propriamente strumenti utilizzati nel settore

giurisdizionale quanto piuttosto nella attività di polizia e in generale di contrasto e prevenzione dei reati. Tuttavia, poiché una consistente parte di questo lavoro è dedicata proprio a sistemi di tal guisa impiegati negli Stati Uniti dalle autorità giurisdizionali, è utile comprendere come il nuovo Regolamento “vede” l’impiego di tali sistemi.

4.1. La giustizia algoritmica nella Proposta della Commissione e il parere del Comitato economico sociale europeo

Nella proposta iniziale della Commissione le previsioni sugli strumenti di intelligenza artificiale impiegati nel settore della giustizia erano del tutto residuali. Nel *considerando n. 40* si specificava che alcuni sistemi di IA destinati all’amministrazione della giustizia avrebbero dovuto essere classificati ad alto rischio, «in considerazione del loro impatto potenzialmente significativo sulla democrazia, sullo Stato di diritto, sulle libertà individuali e sul diritto a un ricorso effettivo e a un giudice imparziale».

In particolare, secondo la Commissione era opportuno, «al fine di far fronte ai rischi di potenziali distorsioni, errori e opacità, classificare come ad alto rischio i sistemi di IA destinati ad assistere le autorità giudiziarie nelle attività di ricerca e interpretazione dei fatti e del diritto e nell’applicazione della legge a una serie concreta di fatti». L’inserimento nella categoria *high risk* non doveva tuttavia comprendere i «sistemi di IA destinati ad attività amministrative puramente accessorie, che non incidono sull’effettiva amministrazione della giustizia nei singoli casi, quali l’anonimizzazione o la pseudonimizzazione di decisioni, documenti o dati giudiziari, la comunicazione tra il personale, i compiti amministrativi o l’assegnazione delle risorse».

Nell’articolato della Proposta, dunque, venivano classificati come ad altro rischio «i sistemi di IA destinati ad assistere un’autorità giudiziaria nella ricerca e

nell'interpretazione dei fatti e del diritto e nell'applicazione della legge a una serie concreta di fatti»⁴⁰⁹.

Tra i pareri istituzionali intervenuti a seguito della Proposta, particolarmente significativa è stata l'*opinion* del Comitato economico sociale europeo (CESE).

L'influenza che ha avuto il parere del Comitato sulle negoziazioni dell'*AI Act*, anche con riguardo alla giustizia algoritmica, emerge dall'importanza che il CESE conferisce all'intervento umano rispetto a decisioni, anche in campo giudiziario, automatizzate. Come vedremo, le indicazioni del CESE verranno prese in seria considerazione nel processo di approvazione del Regolamento, che ridefinirà alcune previsioni e prevederà in generale che «il processo decisionale finale deve rimanere un'attività a guida umana» (*considerando n. 61* del Regolamento). A questo riguardo, nel parere il CESE raccomanda «vivamente» che il Regolamento preveda «che determinate decisioni rimangano prerogativa degli esseri umani, in particolare negli ambiti in cui tali decisioni hanno una componente morale e implicazioni giuridiche o un impatto sociale, come ad esempio in campo giudiziario, nelle attività di contrasto, nei servizi sociali, nella sanità, nella politica degli alloggi, nei servizi finanziari, nei rapporti di lavoro e nell'istruzione»⁴¹⁰.

Inoltre, in riferimento proprio all'utilizzo della intelligenza artificiale nell'amministrazione della giustizia, il CESE afferma che esso rappresenta «un tema particolarmente delicato e dovrebbe essere affrontato con più sfumature e controlli di quanto non avvenga attualmente. Il riferimento ai sistemi di IA utilizzati per assistere l'autorità giudiziaria nella ricerca e nell'interpretazione dei fatti e del diritto e nell'applicazione delle norme a una serie concreta di fatti non considera che giudicare è attività assai più complessa della mera ricerca di schemi ricorrenti in dati storici (che in sostanza è ciò che gli attuali sistemi di intelligenza artificiale fanno)»⁴¹¹.

⁴⁰⁹ Art. 6 e Allegato III della Proposta.

⁴¹⁰ Parere del Comitato economico e sociale europeo sulla proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione, 22 dicembre 2021, §§ 1.9 e 3.7

⁴¹¹ *Ivi*, § 4.21.

Il CESE riconosce pertanto l'importanza di non sottovalutare il fenomeno e ammonisce rispetto alla complessità dell'attività giurisdizionale che non può essere relegata al mero appiattimento su dati statistici e predittivi.

4.2. L'idea di giustizia algoritmica nelle previsioni del Regolamento

A seguito del parere del CESE e del trilogio con il Consiglio e con il Parlamento, la giustizia algoritmica assume un peso molto maggiore nella visione del legislatore europeo. Non cambia la classificazione dei sistemi impiegati da un'autorità giudiziaria – inseriti sempre nella categoria *high risk* – ma il contorno di tale classificazione assume dei caratteri molto diversi rispetto a quelli visti nella proposta iniziale della Commissione. La formulazione testuale dell'allegato III, infatti, non muta nella sostanza⁴¹², in quanto sono considerati ad alto rischio «i sistemi di IA destinati a essere usati da un'autorità giudiziaria o per suo conto per assistere un'autorità giudiziaria nella ricerca e nell'interpretazione dei fatti e del diritto e nell'applicazione della legge a una serie concreta di fatti, o a essere utilizzati in modo analogo nella risoluzione alternativa delle controversie» (§ 8, lett. a)). E tuttavia, nei *considerando* e nelle ulteriori previsioni del Capo III relativi ai sistemi di alto rischio è possibile intravedere in quale contesto si è mossa la legislazione europea.

Innanzitutto, rispetto al *considerando n. 40* della proposta (si v. *supra* al § precedente), nel *considerando n. 61* del Regolamento il Parlamento europeo ha inteso precisare che «[l'] utilizzo di strumenti di IA può fornire sostegno al potere decisionale dei giudici o all'indipendenza del potere giudiziario, ma non dovrebbe sostituirlo: il processo decisionale finale deve rimanere un'attività a guida umana».

Si tratta più che altro di una dicitura simbolica – posto che già nella formulazione iniziale si parlava solo di «assistere» e non di sostituire il giudice – ma in ogni caso efficace anche per eliminare qualunque tipo di equivoco rispetto a una *sostituzione* del giudice con l'intelligenza artificiale. L'autorità giurisdizionale, insomma,

⁴¹² Viene tuttavia aggiunto l'impiego dell'IA nella risoluzione alternativa delle controversie, tema che non trattiamo, come indicato nel capitolo introduttivo.

rimane al suo posto, di tal che il ruolo del giurista non è quello di ipotizzare scenari senza più il giudice (ipoteticamente sostituito da sistemi artificiali) bensì di ridefinire i suoi compiti alla luce di un nuovo strumento, come già è stato fatto per le novità che nel corso dei decenni hanno interessato l'attività giurisdizionale.

Come anticipato nel paragrafo dedicato all'impostazione generale del Regolamento, occorre ora calare le previsioni dei sistemi di IA ad alto rischio (Capo III del regolamento) nel contesto della giustizia algoritmica.

In generale, come già visto, un sistema algoritmico impiegato nell'amministrazione della giustizia – per essere considerato *ad alto rischio* – deve incidere sul processo decisionale del giudice⁴¹³, tanto che nei *considerando* si esclude tale classificazione per strumenti di IA impiegati in attività accessorie.

E ciò accade quando, come visto, è soddisfatta *almeno* una delle quattro condizioni previste dal comma successivo: a) il sistema di IA è destinato a eseguire un *compito procedurale limitato*; b) il sistema di IA è destinato a migliorare il risultato di *un'attività umana precedentemente completata*; c) il sistema di IA è destinato a *rilevare schemi decisionali o deviazioni da schemi decisionali precedenti* e non è finalizzato a sostituire o influenzare la valutazione umana precedentemente completata senza un'adeguata revisione umana; d) il sistema di IA è destinato a eseguire *un compito preparatorio* per una valutazione pertinente ai fini dei casi d'uso elencati nell'allegato III (Art 6 § 3, secondo comma)⁴¹⁴.

Nulla quaestio se dunque un giudice assume la propria decisione senza l'ausilio di alcuno strumento di intelligenza artificiale, essendo quest'ultimo intervenuto solo in una fase accessoria o solo *successivamente* all'assunzione della decisione, magari rilevando errori lessicali nella decisione o delle deviazioni rispetto alla

⁴¹³ Che è la prospettiva di questo lavoro, come già indicato nel capitolo introduttivo e come si vedrà più avanti (si v. *infra*, parte III).

⁴¹⁴ Il Regolamento prevede inoltre che la Commissione possa modificare o integrare le condizioni «qualora vi siano prove concrete e affidabili dell'esistenza di sistemi di IA che rientrano nell'ambito di applicazione dell'allegato III ma non presentano un rischio significativo di danno per la salute, la sicurezza o i diritti fondamentali delle persone fisiche» (art. 6, § 6).

giurisprudenza maggioritaria⁴¹⁵. In questi casi siamo infatti fuori dall'ambito di applicazione del Regolamento per quanto concerne i sistemi ad alto rischio.

Fuori da questi casi, la giustizia algoritmica è un ambito considerato ad alto rischio dal legislatore europeo e deve, quindi, rispettare i requisiti e gli obblighi stringenti previsti nel Regolamento.

Essi verranno presi in considerazione dopo aver costituzionalizzato i problemi della giustizia algoritmica alla luce delle garanzie del giusto processo, al fine di verificare se le nuove regole sono in grado di porre una efficace barriera alle possibili violazioni che potrebbero presentarsi.

Quello che invece si può qua sottolineare è quale sia la giustizia algoritmica che il legislatore ha in mente e che intende regolare, considerandola un ambito ad alto rischio per il potenziale impatto sui diritti fondamentali.

La norma dice innanzitutto sistemi di IA impiegati da un'autorità giudiziaria *o per suo conto*: ipotizziamo quindi che anche se lo strumento venisse impiegato non direttamente dal giudice ma da altre articolazioni all'interno dell'apparato giurisdizionale (esempio gli ausiliari del giudice o il personale dell'Ufficio per il Processo), si ricadrebbe comunque nel perimetro della norma.

I sistemi ad alto rischio sono quelli impiegati *per assistere un'autorità giudiziaria*: anche qui, come già specificato nei *Considerando*, l'unica prospettiva possibile del legislatore europeo è di una giustizia algoritmica *ausiliaria* e non *sostitutiva*⁴¹⁶.

Infine i sistemi considerati sono quelli per *la ricerca e l'interpretazione dei fatti e del diritto e nell'applicazione della legge a una serie concreta di fatti*. Ora, il legislatore europeo non sembra voler identificare specificamente quali sistemi ha in mente, utilizzando una formulazione particolarmente ampia.

Quel che sembra potersi affermare, rispetto alle esperienze di giustizia predittiva che abbiamo delineato nel capitolo II, è che quelle ad oggi sperimentate in Europa

⁴¹⁵ Peraltro l'adozione di sistemi di questo tipo è stata già propugnata in dottrina. Si cfr. A. PUNZI, *Judge in the machine. E se fossero le macchine a restituirci l'umanità nel giudicare?* In A. CARLEO (a cura di) *Decisione robotica*, Il Mulino, Bologna, 2019, p. 330.

⁴¹⁶ Per questa classificazione in dottrina si v. M. BARBERIS, *Giustizia predittiva: ausiliare e sostitutiva. Un approccio evolutivo*, in *Milan Law Review*, vol. 3 n. 2, 2022, pp. 1-18.

rientrano a pieno nella definizione offerta nel Regolamento, in quanto servono specificamente, tramite elaborazioni statistiche, a ricavare un esito probabile della causa, e quindi ricomprendere sia l'interpretazione dei fatti (comuni al caso nuovo) sia l'applicazione della legge (mediata dall'applicazione fatta di quella legge nei precedenti elaborati dal sistema di IA).

Sembrano inoltre potersi ricomprendere nella formulazione legislativa anche quei casi in cui il giudice si serve di strumenti di IA quali ChatGPT per scrivere (parti del) la sentenza. Nelle esperienze che abbiamo indicato, infatti, il giudice si è servito di tali strumenti sia per interpretare i fatti, sia per chiarirsi le modalità di applicazione della legge a quei fatti, in conformità con le regole dell'ordinamento di riferimento.

Sembrerebbero poter rientrare anche i sistemi predittivi impiegati nelle giurisdizioni penali negli Stati Uniti qualora utilizzati, come si fa oltreoceano, direttamente dal giudice nella propria attività decisoria. Considerato il loro funzionamento e i dati processati al proprio interno, sarebbe difficile non considerarli dei sistemi che aiutano il giudice quantomeno nell'interpretazione dei fatti, considerato che il sistema ricostruisce la propria predizione di rischio di recidiva sulla base dei fatti (dati) riferibili all'imputato sottoposto a giudizio⁴¹⁷.

Tuttavia, il Regolamento non li ricomprende nei sistemi impiegati nell'amministrazione della giustizia, fornendo una regolamentazione autonoma, poiché in Europa, ad oggi, tali sistemi vengono impiegati ma al di fuori del processo, per esempio nelle attività di indagine o di polizia⁴¹⁸. Ciò, tuttavia, non ne cambia la visione del legislatore europeo che, come vedremo, li considera strumenti ad alto rischio e, in alcuni casi, addirittura vietati.

⁴¹⁷ In dottrina si cfr. G. VAN DIJCK, *Predicting Recidivism Risk Meets AI Act*, in *European Journal on Criminal Policy and Research*, 28, 2022, pp. 407-423, spec. p. 412, il quale ritiene che i *risk assessment tools* rientrino nella classificazione del Regolamento in quanto la valutazione del rischio può essere considerata un "fatto" a cui si applica la legge e che può essere interpretato e soppesato rispetto agli altri fatti del caso.

⁴¹⁸ Un esempio è il sistema *Hart* impiegato nel Regno Unito, si cfr. O. MARION *et al.*, *Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality*, in *Information & Communications Technology Law*, 2/2018, 223-250. Ne discute diffusamente anche M. GIALUZ, *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa*, in *Diritto penale contemporaneo*, 29 maggio 2019.

4.3. Gli strumenti di previsione del rischio di recidiva all'interno dell'AI Act

Gli strumenti di previsione del rischio di recidiva sono stati oggetto di dibattito nelle interlocuzioni successive alla presentazione da parte della Commissione della Proposta. In essa, infatti, tutti i sistemi di previsione del rischio di commissione di reati o di recidiva venivano ricompresi nella categoria *high risk* disciplinata dall'art. 6 della Proposta di regolamento e dall'Allegato III. Quest'ultimo prevedeva che fossero sistemi ad alto rischio impiegati nell'”Attività di contrasto” (*Law enforcement*)⁴¹⁹:

«a) i sistemi di IA destinati a essere utilizzati dalle autorità di contrasto per effettuare valutazioni individuali dei rischi delle persone fisiche al fine di determinare il rischio di reato o recidiva in relazione a una persona fisica (...);

(...)

e) i sistemi di IA destinati a essere utilizzati dalle autorità di contrasto per prevedere il verificarsi o il ripetersi di un reato effettivo o potenziale sulla base della profilazione delle persone fisiche di cui all'articolo 3, paragrafo 4, della direttiva (UE) 2016/680 o per valutare i tratti e le caratteristiche della personalità o il comportamento criminale pregresso di persone fisiche o gruppi;

f) i sistemi di IA destinati a essere utilizzati dalle autorità di contrasto per la profilazione delle persone fisiche di cui all'articolo 3, paragrafo 4, della direttiva (UE) 2016/680 nel corso dell'indagine, dell'accertamento e del perseguimento di reati (...)).».

Ora, rispetto alla Proposta, il Regolamento muta tale classificazione e inserisce nelle pratiche di IA *vietate* (art. 5, § 1 lett. d):

«l'immissione sul mercato, la messa in servizio per tale finalità specifica o l'uso di un sistema di IA per effettuare valutazioni del rischio relative a persone fisiche al fine di valutare o prevedere il rischio che una persona fisica commetta un reato, *unicamente sulla base della profilazione di una persona fisica o della valutazione*

⁴¹⁹ Si ribadisce che è questa la categoria in cui vengono classificati gli strumenti di previsione del rischio di recidiva e non la categoria dell'amministrazione della giustizia.

dei tratti e delle caratteristiche della personalità; tale divieto non si applica ai sistemi di IA utilizzati a sostegno della valutazione umana del coinvolgimento di una persona in un'attività criminosa, che si basa già su fatti oggettivi e verificabili direttamente connessi a un'attività criminosa» (corsivo nostro)⁴²⁰.

Sono invece classificati ad alto rischio (Allegato III, § 6):

«d) i sistemi di IA destinati a essere utilizzati dalle autorità di contrasto o per loro conto, oppure da istituzioni, organi e organismi dell'Unione a sostegno delle autorità di contrasto, per determinare il rischio di commissione del reato o di recidiva in relazione a una persona fisica *non solo sulla base della profilazione delle persone fisiche di cui all'articolo 3, paragrafo 4, della direttiva (UE) 2016/680* o per valutare i tratti e le caratteristiche della personalità o il comportamento criminale pregresso di persone fisiche o gruppi» (corsivo nostro).

e) i sistemi di IA destinati a essere utilizzati dalle autorità di contrasto o per loro conto, oppure da istituzioni, organi e organismi dell'Unione a sostegno delle autorità di contrasto, per effettuare la profilazione delle persone fisiche di cui all'articolo 3, paragrafo 4, della direttiva (UE) 2016/680 nel corso dell'indagine, dell'accertamento e del perseguimento di reati».

A corredo dell'articolato, il *considerando n. 59* fornisce una spiegazione della classificazione prevista, indicando una serie di elementi particolarmente rilevanti che ci serviranno nella delimitazione delle garanzie costituzionali rilevanti nel contesto della predizione algoritmica. Per quanto qui di interesse, tale considerando specifica che l'impiego di sistemi predittivi di IA potrebbero ostacolare «l'esercizio di importanti diritti procedurali fondamentali, quali il diritto a un ricorso effettivo

⁴²⁰ A spiegazione di questo articolo interviene il *Considerando n. 42* del Regolamento, il quale prevede che «[i]n linea con la presunzione di innocenza, le persone fisiche nell'Unione dovrebbero sempre essere giudicate in base al loro comportamento effettivo. Le persone fisiche non dovrebbero mai essere giudicate sulla base di un comportamento previsto dall'IA basato unicamente sulla profilazione, sui tratti della personalità o su caratteristiche quali la cittadinanza, il luogo di nascita, il luogo di residenza, il numero di figli, il livello di indebitamento o il tipo di automobile, senza che vi sia un ragionevole sospetto che la persona sia coinvolta in un'attività criminosa sulla base di fatti oggettivi verificabili e senza una valutazione umana al riguardo. Pertanto, dovrebbero essere vietate le valutazioni del rischio effettuate in relazione a persone fisiche intese a determinare la probabilità che queste ultime commettano un reato o volte a prevedere il verificarsi di un reato effettivo o potenziale unicamente sulla base della loro profilazione o della valutazione dei loro tratti della personalità e delle loro caratteristiche (...)).»

e a un giudice imparziale, nonché i diritti della difesa e la presunzione di innocenza, in particolare nel caso in cui tali sistemi di IA non siano sufficientemente trasparenti, spiegabili e documentati». Inoltre, aggiunge che «l'impatto dell'utilizzo degli strumenti di IA sul diritto alla difesa degli indagati non dovrebbe essere ignorato, in particolare la difficoltà di ottenere informazioni significative sul funzionamento di tali sistemi e la difficoltà che ne risulta nel confutarne i risultati in tribunale, in particolare per le persone fisiche sottoposte a indagini».

Ad ogni buon conto, rispetto alla Proposta, nella versione definitiva del Regolamento vi è quindi la specificazione che predizioni di rischio effettuate *solo sulla base della profilazione* costituiscono un rischio talmente elevato da essere vietate, salvo i casi in cui tali strumenti non vengano utilizzati nei confronti di soggetti già appartenenti sottoposti a indagine che si basi su fatti oggettivi e verificabili.

Il Regolamento vuole quindi evitare che sistemi di IA che profilano le persone vengano utilizzati, ad esempio, per porre sotto indagine un individuo solo per le sue caratteristiche personali e non per i fatti da esso compiuti.

Indicate le norme, occorre chiedersi in quale categoria del Regolamento rientrino i sistemi di *risk assessment* come COMPAS impiegati nella giurisdizione americana.

Non sembrano esserci dubbi sul fatto che tali sistemi si basino sulla profilazione in quanto, come visto quando abbiamo descritto il funzionamento di COMPAS (si v. *supra*, cap. I § 4) essi e utilizzano dati personali come il reddito, lo stato civile, il precedente abuso di alcool, l'uso di droghe ecc...⁴²¹.

Tanto precisato, non sembrerebbe che i sistemi impiegati oltreoceano siano vietati ai sensi del Regolamento sull'intelligenza artificiale. Ciò per due ordini di ragioni: in primo luogo strumenti come COMPAS non si basano *esclusivamente* sulla profilazione. In secondo luogo, anche se si basassero esclusivamente sulla

⁴²¹ La profilazione rilevante in questo contesto è quella prevista dall'art. 3, paragrafo 4, della direttiva (UE) 2016/680 secondo cui profilazione è «qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di tali dati personali per valutare determinati aspetti personali relativi a una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona fisica».

profilazione del soggetto⁴²², rientrerebbero nell'eccezione prevista nell'art. 5 secondo cui il divieto non si applica quando vengono utilizzati a sostegno della valutazione umana che si basa già su fatti oggettivi e verificabili direttamente connessi a un'attività criminosa. Come abbiamo visto, questi sistemi vengono infatti utilizzati nel procedimento americano nella fase di determinazione della pena, quando il soggetto – dopo la valutazione del fatto di reato e degli altri elementi rilevanti – è stato già dichiarato colpevole. La valutazione umana, in altre parole, rimane nell'identificazione dei fatti rilevanti, delle prove e delle leggi applicabili: l'algoritmo interviene in un secondo momento per adjuvare il giudice nell'applicazione della pena.

Riassumendo, sistemi di predizione di rischio di recidiva come quelli impiegati negli Stati Uniti rientrano nella categoria dei sistemi “ad alto rischio” dovendone rispettare i relativi requisiti. Come anticipato, se venissero impiegati anche in Unione europea non solo nell'attività di contrasto e prevenzione dei reati ma anche all'interno dei sistemi giudiziari sarebbero probabilmente classificabili come strumenti *high risk* anche ai sensi dell'art. 6 e dell'Allegato III: e ciò sia perché sono utilizzati per la ricerca e l'interpretazione dei fatti e sia perché, come si chiarirà a fondo, sono sicuramente influenti nel processo decisionale del giudice.

4.4. L'implementazione dell'AI Act nel disegno di legge di delega italiano sull'intelligenza artificiale

In concomitanza con le ultime fasi di approvazione dell'*AI Act*, il 23 aprile 2024 il Consiglio dei Ministri italiano ha deliberato un disegno di legge di delega in materia di intelligenza artificiale, mentre si scrive in discussione al Senato. Curiosamente, all'art. 1, dopo aver indicato le finalità della legge⁴²³, precisa che

⁴²² Come abbiamo specificato nel capitolo 1, non è possibile analizzare le caratteristiche di tutti i sistemi impiegati. Abbiamo esaminato specificamente COMPAS in quanto uno dei sistemi più evoluti tra quelli impiegati all'interno delle giurisdizioni americane.

⁴²³ «La presente legge reca principi in materia di ricerca, sperimentazione, sviluppo, adozione e applicazione di sistemi e modelli di intelligenza artificiale. Promuove un utilizzo corretto, trasparente e responsabile, in una dimensione antropocentrica, dell'intelligenza artificiale, volto a

«[I]e disposizioni della presente legge si interpretano e si applicano conformemente al diritto dell'Unione europea». Come noto, infatti, l'interpretazione della normativa nazionale conformemente al diritto dell'Unione europea è un principio pacifico, riconosciuto tanto dalla Corte di Giustizia⁴²⁴ quanto dalla Corte costituzionale⁴²⁵.

La definizione di sistema di intelligenza artificiale ricalca esattamente quella prevista nel Regolamento europeo (art. 2) e i principi generali (art. 3) ne fanno sicuramente eco: oltre al rischio ai diritti e alle libertà costituzionali si afferma che «[i] sistemi e i modelli di intelligenza artificiale devono essere sviluppati ed applicati nel rispetto della autonomia e del potere decisionale dell'uomo, della prevenzione del danno, della conoscibilità, della spiegabilità...» secondo i principi di autonomia e umana e trasparenza che già erano indicati nell'*AI Act*. Dopo aver indicato i principi in materia di informazione e riservatezza dei dati (art. 4), di utilizzo dell'IA per lo sviluppo economico (art. 5), di sicurezza nazionale (art. 6), il Capo II è dedicato alle disposizioni di settore.

Per quanto ci riguarda, la disposizione riguardante l'impiego dell'intelligenza artificiale nel settore giudiziario è l'art. 14 il quale prevede che «i sistemi di intelligenza artificiale sono utilizzati esclusivamente per l'organizzazione e la semplificazione del lavoro giudiziario nonché per la ricerca giurisprudenziale e dottrinale. Il Ministero della giustizia disciplina l'impiego dei sistemi di intelligenza artificiale da parte degli uffici giudiziari ordinari. Per le altre giurisdizioni l'impiego è disciplinato in conformità ai rispettivi ordinamenti» (comma 1); «[è] sempre riservata al magistrato la decisione sulla interpretazione della legge, sulla valutazione dei fatti e delle prove e sulla adozione di ogni provvedimento» (comma 2).

Quanto al primo comma, è interessante notare che la legge destina al solo Ministero della giustizia la disciplina dell'impiego dei sistemi di intelligenza

coglierne le opportunità. Garantisce la vigilanza sui rischi economici e sociali e sull'impatto sui diritti fondamentali dell'intelligenza artificiale».

⁴²⁴ *Ex multis*, Corte di Giustizia, causa C-282/10, *Maribel Dominguez c. Centre informatique du Centre Ouest Atlantique, Préfet de la région Centre*, 24 gennaio 2012.

⁴²⁵ *Ex multis*, Corte cost., sent. n. 28 del 2010.

artificiale, senza alcuna menzione del Consiglio Superiore della Magistratura. Una previsione di questo tipo rischia di generare alcune problematiche di natura costituzionale, legate all'autonomia e all'indipendenza della magistratura, soprattutto per la natura e il funzionamento dei sistemi di IA (si v. *infra*, cap. VIII). L'art. 14 appare comunque più restrittivo delle norme europee: se queste ultime classificano «ad alto rischio», e quindi non vietano, i sistemi di IA usati da un'autorità giudiziaria o per suo conto per assistere un'autorità giudiziaria nella ricerca, nell'interpretazione dei fatti e del diritto e nell'applicazione della legge a una serie concreta di fatti (come visto sembrando ricomprendere anche strumenti evoluti di giustizia predittiva), la disposizione nazionale sembra escludere qualunque utilizzo da parte del giudice di sistemi di IA che non sia per la *semplificazione* del lavoro giurisprudenziale nonché per la *ricerca* giurisprudenziale e dottrinale (rimangono quindi ammesse le banche dati anche qualora impiegassero evoluti algoritmi di *machine learning* per la ricerca)⁴²⁶.

Inoltre – ma qui sembra solo una specificazione del *considerando n. 61* del Regolamento, il quale afferma che il processo decisionale finale deve rimanere un'attività a guida umana – il comma 2 prevede l'esclusiva titolarità del magistrato per l'interpretazione della legge, la valutazione dei fatti e delle prove e l'adozione del provvedimento. Come vedremo, tali disposizioni, per quanto condivisibili, non sembrano fare i conti coi punti nodali e ancora irrisolti dell'impiego di algoritmi all'interno delle giurisdizioni (si v. *infra*, cap. VI).

5. La recente strategia della Corte di Giustizia Europa sulla giustizia algoritmica

Rimanendo in ambito eurounitario, merita un ultimo cenno la recente strategia della Corte di Giustizia dell'Unione europea sull'impiego dell'intelligenza artificiale. Essa dimostra che la Corte non è rimasta estranea rispetto al tema della giustizia algoritmica, e anzi l'approfondimento di come la Corte di Giustizia

⁴²⁶ Ci si chiede peraltro se sia stata una dimenticanza o una scelta del Governo l'omissione della ricerca normativa quale possibile utilizzo dell'IA.

intende muoversi è ancora più rilevante se si pensa che essa sarà l'organo giurisdizionale che sarà chiamato a interpretare il nuovo Regolamento sull'intelligenza artificiale.

La strategia è incentrata principalmente su tre diverse aree: «*improving the efficiency and effectiveness of our administrative and judicial processes, enhancing the quality and consistency of judicial decisions, and increasing access and transparency for EU citizens*»⁴²⁷.

Già dall'*executive summary* si evince che l'interesse della Corte è più che altro indagare l'impiego dell'IA in quelle attività che non incidono *direttamente* sull'*output* decisionale del giudice, quanto piuttosto l'utilizzo di algoritmi in quelle attività – che l'*AI Act* chiamerebbe “accessorie” – per le quali una più massiccia digitalizzazione dovrebbe garantire una maggior speditezza ed efficienza dell'apparato giurisdizionale. Ciò premesso, dal documento emergono comunque indicazioni interessanti rispetto agli usi dell'IA che questo lavoro intende indagare.

Sono tre gli obiettivi principali enunciati nella strategia: il primo, molto generico, è migliorare l'efficienza della Corte sia nelle procedure amministrative sia in quelle giurisdizionali. Ciò dovrebbe raggiungersi mediante quattro diverse direttrici: a) implementare l'IA nelle procedure della Corte aumenterebbe per aumentare la speditezza, diminuire il carico di lavoro, gestire meglio le risorse e, in definitiva, creare un sistema più efficiente per «*both the Court and the people it serves*»⁴²⁸; b) una *governance* dei dati⁴²⁹; c) la trasformazione dell'ecosistema

⁴²⁷ La strategia è stata pubblicata sul sito della Corte di Giustizia dell'Unione Europea, *Artificial Intelligence Strategy*, 2023. Sulla strategia si veda il commento di E. LONGO, *La “giustizia digitale” tra nuove forme di efficienza e garanzia dei diritti costituzionali*, cit.

⁴²⁸ Nel dettagliare l'obiettivo, la Corte esplicita le modalità con cui esso potrebbe raggiungersi: attraverso un miglioramento del sistema SIGA, il sistema digitale di *case management* presso la Corte; applicazioni di *speech-to-text machine* per la verbalizzazione automatica delle udienze; l'utilizzo di motori di ricerca più accurati grazie all'IA, che permette anche la ricerca semantica e non solo per parole chiave; efficientamento di applicazioni di NLP (*Natural Language Processing*) per la comprensione e il riassunto automatico di documenti processuali o materiali di ricerca; la diffusioni di assistenti virtuali per le attività amministrative, che assistano lo staff della Corte nello svolgimento di compiti di routine come la programmazione, la preparazione preliminare di documenti, lettere o presentazioni; la creazione di “*smart buildings*”, edifici e luoghi fisici che diventano sempre più connessi e intelligenti, con sensori e altre tecnologie integrate nell'edificio per monitorarne e ottimizzarne le prestazioni (pp. 10-11).

⁴²⁹ Secondo la Corte i dati devono costituire l'asset centrale della trasformazione digitale «*because without quality and representative data, A.I. algorithms cannot be correctly trained to*

lavorativo della Corte, che deve garantire lo sviluppo delle competenze adeguate a sostenere l'avvento dell'IA⁴³⁰; l'adozione di un IA «governabile»⁴³¹.

Il secondo obiettivo – ai nostri fini più interessante – mira a migliorare la qualità e la coerenza (*quality and consistency*) delle decisioni giurisdizionali da raggiungere attraverso tre diverse iniziative: a) sfruttare l'automazione⁴³²; b) migliorare la ricerca giuridica⁴³³; c) sfruttare la standardizzazione⁴³⁴.

Il terzo obiettivo è quello di garantire un migliore accesso alla giustizia e una maggiore trasparenza per i cittadini europei da raggiungere tramite: a) un aumento dell'accessibilità per i cittadini con disabilità⁴³⁵; b) un miglioramento dei sistemi

produce a quality output. Therefore, data governance is essential to achieve this objective». Molto interessante la strategia dei dati che la Corte intende prediligere: il processo giudiziario dovrà utilizzare algoritmi e dati creati e utilizzati all'interno della Corte a causa dell'elevata sensibilità dei dati relativi ai casi e dei requisiti di protezione dei dati (pp. 11-12).

⁴³⁰ Le iniziative chiave individuate dalla Corte sono: aggiornare lo staff tramite l'*Emerging Technologies Academy*, un'iniziativa lanciata dalla Corte nel 2022 con l'obiettivo di preparare il personale alla rivoluzione tecnologica; l'assegnazione tramite IA delle risorse umane ai settori che ne richiedono; adottare una strategia di *sourcing* che consenta di aumentare o diminuire in modo flessibile il personale necessario; attrarre talenti, reclutare in modo efficiente il personale con le nuove competenze necessarie (p. 12).

⁴³¹ Che non vuol dire solo l'adozione di strumenti di IA, ma soprattutto «*the continuous updating and supervision of such tools*» (pp. 12-13). Al riguardo la Corte propone la propria struttura di *governance* che comprende: un'*AI Management Board*, con la missione di garantire che l'acquisizione o la creazione di qualsiasi strumento di IA rispetti l'etica e i diritti fondamentali; l'*Informatics Steering Committee (CDI)*, che già esiste all'interno della Corte e si occuperà della gestione delle risorse; l'*AI+ Network*, rete creata nel 2020 composta da rappresentanti dei dipartimenti, delle cancellerie e della Corte, con il compito di individuare le aree in cui strumenti di IA potrebbero maggiormente recare beneficio ai lavori della Corte; *Architecture and data governance boards* che assicurino la robustezza tecnica dei sistemi di IA e la qualità dei dati utilizzati (pp. 20-21).

⁴³² Si prevede l'incorporazione nel SIGA di meccanismi di elaborazione automatica sia delle decisioni e delle conclusioni dell'Avvocato generale per estrarre riferimenti e “descrittori”, sia degli atti introduttivi per estrarne “indicatori” e orientamenti. Tutto questo garantendo l'*explainability*, vale a dire, secondo la Corte, «*capability of the machine to offer evidence on how it reached a particular recommendation or to highlight the parts that more likely have influenced the recommendation*» (p. 13).

⁴³³ Ciò tramite l'implementazione dell'IA nei motori di ricerca, che potrebbe assistere i giudici, i funzionari legali o i colleghi che si occupano di ricerca giuridica, analizzando rapidamente grandi quantità di dati, identificando i casi rilevanti e fornendo raccomandazioni. Inoltre, strumenti di giustizia predittiva, utilizzati e sperimentati da alcuni studi legali, dovrebbero essere studiati in modo approfondito perché potrebbero costituire «*high-risk algorithms*» (pp. 13-14).

⁴³⁴ E cioè creare un sistema di IA che dai documenti nazionali che arrivano in Corte crei dei modelli *standard* automatici riorganizzando il modo in cui le informazioni sono presentate. Questo, combinato con traduzioni automatiche, potrebbe aiutare a evitare le difficoltà di interpretazione di tali documenti che spesso incontra la Corte (p. 14).

⁴³⁵ Le tecnologie dell'IA, come i software *text-to-speech* e *speech-to-text* e gli assistenti virtuali potrebbero fornire supporto alle persone con disabilità visive, uditive o cognitive. Queste tecnologie

di accesso alla giustizia e aumentare la trasparenza⁴³⁶; c) una incorporazione del multilinguismo in ogni attività della Corte⁴³⁷; d) l'impegno per diventare un partner attivo nell'ecosistema dell'*e-justice*⁴³⁸.

Oltre agli obiettivi, la Corte delinea una serie di principi che dovrebbero guidare l'implementazione dell'IA in seno alla Corte di giustizia⁴³⁹, principi che, arrivati a questo punto della trattazione, appaiono ormai familiari: *fairness, impartiality and non-discrimination*⁴⁴⁰; *transparency*⁴⁴¹; *traceability*⁴⁴²; *privacy and data protection*⁴⁴³; *human oversight*⁴⁴⁴; *continuous improvement*⁴⁴⁵.

Più interessante risulta invece il capitolo del documento dedicato ai possibili rischi che la Corte individua, ai quali propone alcune strategie attraverso cui essi

potrebbero aiutare a rimuovere le barriere di accesso e consentire alle persone con disabilità di impegnarsi nell'apprendimento e nell'interazione con la CGUE (p. 14).

⁴³⁶ I *chatbot* e gli assistenti virtuali potrebbero fornire modalità più semplici per accedere alle informazioni fornite dalla Corte, nelle attività amministrative e giudiziarie. Inoltre, i video riassuntivi delle udienze che produce la Corte richiedono tanto tempo e non sono disponibili in tutte le lingue: utilizzando l'A.I., si potrebbero produrre i *briefing* utilizzando un *Avatar AI* che crea video basati su un testo in pochi minuti. Gli *Avatar AI* potrebbero essere utilizzati anche per la formazione interna e per produrre materiale didattico per studenti di scuole o università interessati all'attività della Corte o alla conoscenza delle istituzioni europee in generale. Così come strumenti di traduzione simultanea garantirebbero di accedere a informazioni in tempo reale che altrimenti richiederebbero l'interpretazione umana (pp. 14-15).

⁴³⁷ La traduzione basata sull'elaborazione del linguaggio naturale (NLP) offre già la possibilità di rompere le barriere linguistiche nelle comunicazioni. Si prevede che questa tecnologia si evolverà molto rapidamente e porterà i servizi di traduzione a un livello superiore all'interno della Corte stessa, nonché nella comunicazione con le parti, con i tribunali nazionali e con i cittadini dell'UE (pp. 15-16).

⁴³⁸ Secondo la Corte, nel giro di pochi anni si formerà uno spazio giuridico europeo (con l'aggregazione e l'interoperabilità dei dati) e la CGUE sarà un partner attivo all'interno di questo ecosistema. Per fare ciò, la Corte vuole mantenere alta la collaborazione con i tribunali nazionali e con le altre istituzioni europee. Inoltre, ritiene importante trovare i partner giusti nel mondo accademico e della ricerca, per avere accesso a competenze rilevanti in materia di IA.

⁴³⁹ Si v. p. 16.

⁴⁴⁰ Sia i dati che gli algoritmi creati o adottati devono evitare pregiudizi ed essere guidati dal principio di equità e imparzialità, affinché tutte le parti ricevano lo stesso trattamento durante il processo giudiziario o amministrativo.

⁴⁴¹ Il ragionamento alla base degli algoritmi di IA deve essere chiaro e comprensibile, sia per quelli creati *in house* che per quelli acquistati da altri.

⁴⁴² Le soluzioni di IA devono essere verificabili e spiegabili e devono esistere chiari meccanismi di responsabilità di chi utilizza l'IA.

⁴⁴³ L'IA deve rispettare i diritti di protezione dei dati delle persone e garantire che i dati personali siano gestiti in modo sicuro ed etico.

⁴⁴⁴ Qualsiasi strumento di IA deve essere sottoposto a una stretta supervisione umana e il pensiero critico umano deve essere continuamente sviluppato e incoraggiato.

⁴⁴⁵ I programmi di IA devono essere costantemente aggiornati per riflettere l'evoluzione degli standard legali ed etici.

potrebbero essere mitigati⁴⁴⁶. Quelli per noi più interessanti hanno un elemento in comune, l'importanza di un costante e attento controllo umano, che deve essere preparato adeguatamente al cambiamento e nelle mani del quale deve rimanere il potere giurisdizionale. Al riguardo, al rischio che l'uso dell'IA nel sistema giudiziario potrebbe creare preoccupazioni etiche (*ethical concerns*) rispetto al ruolo dell'essere umano nel processo decisionale, la Corte individua come possibile strategia mitigativa «*ensure the appropriate level of human oversight*», oltre a stabilire una «linea rossa» al di sopra della quale i rischi di utilizzo dell'IA superano i benefici. Per mitigare il rischio che la complessità dei sistemi di IA rendano arduo comprendere come sono prodotti gli *output (explainability)*, occorrerebbe, secondo la Corte, avere un personale con un adeguato livello di *training* durante tutto il ciclo di vita dell'IA, da chi progetta a chi, in questo caso il giudice, è chiamato ad utilizzarla.

Infine, due rischi collegati tra loro individuati dalla Corte sono da un lato il rischio di *over-reliance* rispetto ai risultati algoritmici (sull'argomento si v. *infra* cap. VI § 3), il quale deve venire mitigato da un continuo monitoraggio umano dei sistemi e l'incoraggiamento a un pensiero critico e proattivo rispetto all'IA, dall'altro il rischio di *hyper abuse*, secondo cui la rapida evoluzione della tecnologia dell'IA e la disponibilità di algoritmi potrebbero portare a un'adozione incontrollata di tali strumenti da parte degli utenti. Esso dovrebbe essere mitigato tramite il rispetto di un'adeguata *governance* e tramite la diffusione e comunicazione generalizzata delle *policies* e di una cultura dell'IA⁴⁴⁷.

Come si vede, grande importanza è assegnata dalla Corte alla componente umana del passaggio da una giustizia tradizionale a una giustizia che si serve dell'impiego di IA: sorveglianza continua, diffusione della cultura dell'IA,

⁴⁴⁶ Si v. pp. 17-18.

⁴⁴⁷ Tale ultimo rischio sembra preoccupare particolarmente la Corte di Giustizia, che dedica anche un capitolo alle soluzioni da adottare rispetto all'*hyper abuse*. Ruolo fondamentale dovrà averlo l'*Innovation Lab*, una struttura già presente dal 2019 con l'obiettivo di promuovere l'innovazione, fungere da piattaforma per lo scambio di idee. Nel contesto dell'IA esso dovrà continuare a essere il forum in cui le idee saranno discusse, potrà emanare linee guida ed essere profondamente coinvolto nella valutazione dell'adozione di strumenti di IA all'interno della Corte (p. 21).

formazione adeguata di tutti gli attori in gioco. Questo forse dimostra che la Corte non è rimasta estranea al lungo processo di approvazione del Regolamento sull'intelligenza artificiale⁴⁴⁸ che – lo ricordiamo – al *considerando n. 61* afferma un principio tanto semplice quanto decisivo: «il processo decisionale finale deve rimanere un'attività a guida umana». Per quanto l'IA si insinuerà – e ciò accadrà senz'altro – all'interno dei sistemi giurisdizionali, non sarà ammissibile delegare il potere agli algoritmi, che continueranno quindi ad essere dei meri strumenti, mai degli agenti.

Il documento si conclude indicando i passaggi necessari per la transizione dalla fase di sperimentazione alla fase di industrializzazione che comprendono l'adozione di una struttura di governance collegata alla governance già esistente, l'adozione di *policies* chiare per l'adozione dei sistemi di IA da parte della Corte, e l'aggiornamento e la formazione professionale per tutto il personale. L'ultima frase del documento manifesta l'approccio fiducioso della Corte: «*[b]ased on the experience gathered so far, the ECJ has firm foundations to step into the next phase of the adoption of artificial intelligence in a structured and responsible way*»⁴⁴⁹.

6. La giustizia algoritmica nei lavori del Consiglio d'Europa

Intorno agli stessi anni in cui l'Unione Europea mostrava interesse verso il fenomeno dell'intelligenza artificiale, anche il Consiglio d'Europa, attraverso diverse articolazioni interne, ha cominciato a produrre numerosi studi e documenti di *soft law* rispetto al tema, ovviamente secondo la missione di tale organizzazione internazionale, vale a dire la protezione dei diritti fondamentali.

L'approccio del Consiglio d'Europa, coerentemente con i propri meccanismi e le proprie finalità, è stato quello di studiare l'impatto dell'intelligenza artificiale in maniera settoriale, analizzandone i rischi per il rispetto dei diritti fondamentali in

⁴⁴⁸ Oltre alla considerazione dell'importanza del controllo umano, ci sono altre parti della strategia che sembrano riprodurre l'impostazione del regolamento, come quando la Corte parla di *risk-based approach* che l'*AI Management Board* dovrebbe adottare per definire quali applicazioni dovrebbero essere vietate, quali presentano un alto rischio o, invece, un rischio minimo (p. 4).

⁴⁴⁹ Si v. pp. 23-24.

una pluralità di contesti, dalla *data protection* al *gender equality*, dalla medicina ai diritti dei minori, dall'istruzione alla giustizia⁴⁵⁰.

Proprio rispetto al tema della giustizia algoritmica, a differenza del legislatore europeo, il Consiglio d'Europa, soprattutto tramite i lavori della CEPEJ (*European Commission for the efficiency of justice*), fin dagli inizi ha profuso grande impegno nel cercare di individuare le possibili frizioni che l'IA impiegata nei sistemi giudiziari avrebbe potuto provocare rispetto alle garanzie processuali, soprattutto in riferimento all'art. 6 CEDU. Questo interesse pioneristico ha portato all'adozione, nel 2018, nel primo documento europeo rivolto ad affrontare organicamente il tema dell'IA nel processo, la *Ethical Charter on the use of artificial intelligence in judicial systems and their environment*, già più volte incontrata nel corso del lavoro.

Proprio per l'approccio settoriale adottato dal Consiglio d'Europa e l'interesse di questo lavoro precipuamente verso la giustizia algoritmica, nei successivi paragrafi l'attenzione sarà rivolta specificamente al ruolo della regolazione dell'IA in campo giudiziario, senza dover necessariamente calare tale approfondimento nel contesto più ampio della regolazione dell'IA.

Ciò specificato, non si può qui tacere di un rilevante sviluppo che si è avuto recentemente all'interno del Consiglio d'Europa e di cui abbiamo già fatto cenno: la stipulazione del primo trattato giuridicamente vincolante sull'intelligenza artificiale.

6.1. Il trattato sull'intelligenza artificiale e le specifiche previsioni sulla giustizia algoritmica

La *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, adottata il 17 maggio 2024 e aperta alla firma, dal 5 settembre 2024⁴⁵¹, anche ai paesi non europei che hanno partecipato

⁴⁵⁰ Sul sito del Consiglio d'Europa sono presenti tutte le iniziative assunte sul tema dell'intelligenza artificiale: si cfr. il seguente link: <https://www.coe.int/en/web/artificial-intelligence/work-in-progress>.

⁴⁵¹ Per quanto qui interessa, sia gli Stati Uniti sia l'Unione europea hanno firmato il trattato.

all'elaborazione del trattato⁴⁵² e all'Unione europea, mira a fornire un quadro giuridico che copra l'intero ciclo di vita dei sistemi di intelligenza artificiale e i rischi che questi presentano con riguardo alla protezione dei diritti umani, alla democrazia e alla *rule of law*. A questo riguardo nel preambolo si legge:

[c]onvinced of the need to establish, as a matter of priority, a globally applicable legal framework setting out common general principles and rules governing the activities within the lifecycle of artificial intelligence systems that effectively preserves shared values and harnesses the benefits of artificial intelligence for the promotion of these values in a manner conducive to responsible innovation.

Il trattato, non dissimilmente dall'*AI Act*, adotta un approccio basato sul rischio, premettendo al riguardo che:

this Convention is intended to address specific challenges which arise throughout the lifecycle of artificial intelligence systems and encourage the consideration of the wider risks and impacts related to these technologies (...).

Quanto alle *General provisions* (art. 1-3), il trattato prevede che ogni Stato dovrebbe adottare le misure, graduate sulla base dei rischi in gioco, necessarie per garantire che durante tutto il ciclo di vita il sistema di IA⁴⁵³ sia conforme ai principi del rispetto dei diritti umani, della democrazia e della *rule of law*. Inoltre, l'ambito di applicazione è molto ampio (art. 3): la convenzione si applica all'utilizzo dei

⁴⁵² Si legge sul sito del Consiglio d'Europa che «[l]a convenzione è il risultato di due anni di lavoro svolto da un organismo intergovernativo, il Comitato sull'intelligenza artificiale (CAI), che ha riunito, allo scopo di redigere il trattato, i 46 Stati membri del Consiglio d'Europa, l'Unione europea e 11 Stati non membri (Argentina, Australia, Canada, Costa Rica, Giappone, Israele, Messico, Perù, Santa Sede, Stati Uniti d'America e Uruguay), nonché rappresentanti del settore privato, della società civile e del mondo accademico, che hanno partecipato in qualità di osservatori».

⁴⁵³ Definito all'art. 2 come a «*machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that may influence physical or virtual environments. Different artificial intelligence systems vary in their levels of autonomy and adaptiveness after deployment*», definizione che ricalca quella del Regolamento europeo che a sua volta aveva adottato la definizione accolta in seno all'OCSE.

sistemi di IA sia nel settore pubblico (compresi privati che operano nel pubblico), sia nel settore privato. Per quest'ultimo settore si prevedono inoltre per gli Stati due modalità per conformarsi ai principi e obblighi della Convenzione quadro: essi possono scegliere di essere sottoposti direttamente alle disposizioni applicabili della Convenzione o, in alternativa, prendere altre misure per conformarsi alle disposizioni della stessa rispettando appieno i loro obblighi internazionali in materia di diritti umani, democrazia e Stato di diritto.

Sono escluse dall'ambito di applicazione, salvo alcune eccezioni, la sicurezza e la difesa nazionali e l'attività di ricerca e sviluppo.

Indicate le caratteristiche generali della convenzione, di seguito si indicheranno le previsioni più rilevanti per quanto concerne il tema oggetto del presente lavoro.

L'art. 4 stabilisce, con una previsione molto generale, che i sistemi di IA devono rispettare i diritti umani, tra cui senz'altro rientrano il diritto alla tutela giurisdizionale e a un processo giusto.

L'art. 5 §1 mostra invece le preoccupazioni del Consiglio d'Europa anche con riguardo alle applicazioni di intelligenza artificiale impiegate nell'ambito della giustizia, stabilendo che i sistemi di IA non siano impiegati

to undermine the integrity, independence and effectiveness of democratic institutions and processes, including the principle of the separation of powers, respect for judicial independence and access to justice.

Nell'*Explanatory Report* che accompagna la Convenzione⁴⁵⁴ si precisa, nelle spiegazioni dell'art. 5, che i sistemi di IA possono «*influence court rulings, with potential implications for the integrity of the justice system*», come sarà evidente quando si analizzerà l'influenza sul processo decisionale del giudice di strumenti algoritmici (*infra* cap. VIII).

⁴⁵⁴ COMMITTEE ON ARTIFICIAL INTELLIGENCE (CAI), *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law: Explanatory Report*, 133rd Session of the Committee of Ministers, Strasburgo, 17 May 2024.

Tre sono i principi del giusto processo ritenuti fragili di fronte all'applicazione di strumenti di intelligenza artificiale: la separazione dei poteri e l'indipendenza della magistratura (strettamente connessi) e l'accesso alla giustizia.

Per quanto qui interessa, nelle previsioni successive il trattato richiede che le parti contraenti assicurino l'autonomia degli individui (art. 7), la trasparenza (art. 8) e la *accountability* e *responsability* per i danni prodotti dai sistemi di IA (art. 9). Tutte previsioni che hanno una rilevanza, come vedremo, se calate nel contesto della giustizia algoritmica⁴⁵⁵. Nell'*Explanatory Report*, già citato, rispetto alla trasparenza si sottolinea che essa include sia l'*explainability* sia l'*interpretability*.

Il termine *spiegabilità* si riferisce alla capacità di fornire, compatibilmente con la fattibilità tecnica e tenendo conto dello stato dell'arte generalmente riconosciuto, spiegazioni sufficientemente comprensibili sui motivi per cui un sistema di intelligenza artificiale fornisce informazioni, produce previsioni, contenuti, raccomandazioni o decisioni, il che è particolarmente cruciale in settori sensibili come l'assistenza sanitaria, la finanza, l'immigrazione, i servizi di frontiera e la *giustizia penale*, in cui è essenziale comprendere il ragionamento alla base delle decisioni prodotte o assistite da un sistema di intelligenza artificiale.

D'altro lato l'*interpretabilità* si riferisce alla capacità di capire come un sistema di intelligenza artificiale decide: in altre parole, alla misura in cui gli *output* dei sistemi di intelligenza artificiale possono essere resi accessibili e comprensibili sia agli esperti che ai non esperti. Si tratta di rendere il funzionamento interno, la logica e i processi decisionali dei sistemi di intelligenza artificiale comprensibili e accessibili agli utenti umani, compresi gli sviluppatori, le parti interessate, gli utenti finali e le persone interessate o lo impiegano. In questo senso nel *report* si riconosce un punto che sarà al centro della nostra riflessione: poiché la divulgazione delle informazioni ai fini della trasparenza può essere in contrasto con altri interessi (ad esempio il segreto industriale sull'algoritmo) nell'attuazione di questo principio le

⁴⁵⁵ Per quanto riguarda le altre previsioni del trattato, l'art. 10 prevede il divieto di discriminazione, l'art. 11 il rispetto della privacy, l'art. 12 il principio di *reliability* dei sistemi di IA e l'art. 13 l'istituzione, da parte degli Stati, di ambienti di test e sperimentazione dei sistemi di IA prima della messa sul mercato.

Parti contraenti sono tenute a trovare un giusto equilibrio tra i vari interessi in competizione⁴⁵⁶.

Il principio di trasparenza si lega indissolubilmente ai rimedi contro le decisioni assunte (anche) da sistemi di intelligenza artificiale, previsti dagli art. 14 e 15 del trattato. Essi contengono regole fondamentali nel contesto di decisioni prese mediante algoritmi, in particolare se si tratta di decisioni degli organi giudiziari, tenuto conto che «*due to certain unique characteristics of artificial intelligence systems, such as their technical complexity, their data-driven character and the relative opaqueness of the operations of some such systems, human interactions with artificial intelligence systems have been affected by the problem of opaqueness of artificial intelligence systems and information asymmetry*»⁴⁵⁷.

L'art. 14 prevede infatti che ciascuna Parte deve garantire la disponibilità di rimedi accessibili ed efficaci (*accessible and effective remedies*) per le violazioni dei diritti umani derivanti dalle attività dei sistemi di IA: a tal fine, le Parti garantiscono:

a. measures to ensure that relevant information regarding artificial intelligence systems which have the potential to significantly affect human rights and their relevant usage is documented, provided to bodies authorised to access that information and, where appropriate and applicable, made available or communicated to affected persons;

b. measures to ensure that the information referred to in subparagraph “a” is sufficient for the affected persons to contest the decision(s) made or substantially informed by the use of the system, and, where relevant and appropriate, the use of the system itself; and

c. an effective possibility for persons concerned to lodge a complaint to competent authorities.

⁴⁵⁶ *Explanatory Report*, cit., § 60-62.

⁴⁵⁷ *Explanatory Report*, cit., § 96

Il trattato, pertanto, non richiede sempre che alla persona soggetta a una decisione basata (anche) su un sistema di IA (ipotizziamo una decisione giudiziaria) siano comunicate le informazioni rilevanti, ma quanto meno che tale informazione sia resa disponibile a organismi autorizzati che possano adottare le relative misure, che comunque devono essere sufficienti affinché la persona soggetta alla decisione possa contestarla o contestare l'uso del sistema stesso anche con un ricorso rivolto all'autorità competente⁴⁵⁸.

Completa la parte sui rimedi l'art. 15 (*Procedural safeguards*) secondo cui qualora un sistema di intelligenza artificiale abbia un impatto significativo sul godimento dei diritti umani, le persone interessate possano godere di garanzie procedurali, salvaguardie e diritti effettivi (§ 1) e siano informate del fatto che stanno interagendo con sistemi artificiali che con un essere umano (§ 2).

L'art. 16 della Convenzione riguarda il sistema di controllo dei rischi dei sistemi di IA che obbliga le parti a identificare prevenire e mitigare i rischi di tali sistemi, eventualmente vietare certi utilizzi dell'IA «*where it considers such uses incompatible with the respect for human rights, the functioning of democracy or the rule of law*» (è evidente il parallelismo con l'approccio *risk based* seguito nel Regolamento europeo sull'intelligenza artificiale).

Infine, le ultime due previsioni che paiono importanti per il tema qui in oggetto sono inserite negli art. 19 e 20 e riguardano il tema della formazione e alfabetizzazione dell'IA (già più volte considerato nelle pagine del presente lavoro e, come vedremo, tema decisivo per l'implementazione della giustizia algoritmica). L'art. 19 prevede che le Parti si adoperino per garantire che le questioni importanti sollevate dai sistemi di intelligenza artificiale siano debitamente affrontate attraverso un *dibattito pubblico* e la *consultazione* di più *multistakeholder*, alla luce

⁴⁵⁸ Si legge nell'*Explanatory Report* che «[c]onsistent with the principles in Articles 8 (Principle of transparency and oversight) and 9 (Principle of accountability and responsibility), Article 14 of the Framework Convention requires Parties to adopt or maintain specific measures to document and make available certain information to the affected persons in order to support the aim of making available, accessible and effective remedies for violations of human rights in the context of activities in the lifecycle of an artificial intelligence system. The relevant content in the information-related measures should be context-appropriate, sufficiently clear and meaningful, and critically, provide a person concerned with an effective ability to use the information in question to exercise their rights in the proceedings in respect of the relevant decisions affecting their human rights» (§ 99).

delle implicazioni (sociali, economiche, giuridiche, etiche, ambientali ecc...) che comportano. L'art. 20, ancora più specificamente, prevede che le Parti incoraggino e promuovano un'adeguata *alfabetizzazione digitale* e per tutta la popolazione, comprese competenze specifiche per coloro i quali sono responsabili dell'identificazione, della valutazione, della prevenzione e della mitigazione dei rischi posti dai sistemi di intelligenza artificiale⁴⁵⁹. Non è un caso che nell'*Explanatory report* si dica esplicitamente che in considerazione di quanto sia essenziale la formazione per i responsabili dell'identificazione, della valutazione, della prevenzione e della mitigazione dei rischi posti dall'intelligenza artificiale, l'art. 20 fa riferimento soprattutto a uno specifico gruppo di destinatari che include, tra gli altri, le *autorità giudiziarie*⁴⁶⁰.

Ancora non possiamo anticipare la rilevanza che avrà il trattato; e tuttavia, sebbene gli obblighi e i principi enucleati nel trattato appaiano ancora molto generici e di difficile implementazione in mancanza di una legislazione di attuazione da parti dei singoli Stati aderenti alla Convenzione, la portata giuridica di siffatta Convenzione non deve essere sottovalutata. Come noto infatti, qualora la Convenzione verrà ratificata dall'Italia, l'art. 117 primo comma della Costituzione italiana impone al legislatore statale e regionale di conformarsi non solo alla Costituzione, ma anche «ai vincoli derivanti dall'ordinamento comunitario e dagli obblighi internazionali». Come è stato efficacemente scritto «*the Convention not only has a symbolic and meta-legal value, but is also a normative instrument, capable of quasi-constitutional integration into the legal systems of the Member States, and has great interpretative potential. This is why the IA Convention*

⁴⁵⁹ I successivi articoli della convenzione riguardano i meccanismi di cooperazione per l'effettiva implementazione e le eventuali modifiche della Convenzione. In particolare, si prevede l'istituzione di una Conferenza delle Parti con il compito di consultarsi periodicamente sullo stato di applicazione della Convenzione (artt. 23-24); si incoraggiano le Parti firmatarie ad assistere Stati terzi ad agire e adottare misure nel rispetto della convenzione (art. 25); si prevede che ogni Parte istituisca o designi una o più autorità per vigilare, in modo indipendente e imparziale, sull'osservanza degli obblighi previsti dalla convenzione (art. 26).

⁴⁶⁰ *Explanatory report*, cit. § 124: «*in view of how essential training is to those responsible for the identification, assessment, prevention and mitigation of risks posed by artificial intelligence, the provision refers additionally to this specific group of addressees (such actors include, for instance, judiciary, national supervisory authorities, data protection authorities, equality and human rights bodies, ombuds, consumer protection authorities, artificial intelligence providers and artificial intelligence users)*».

supersedes dozens of declaratory and soft-law instruments that were already superfluous, unwieldy and even tedious»⁴⁶¹.

6.2. I lavori della CEPEJ e i principi della Carta Etica sull'impiego dell'IA nei sistemi giudiziari (e il suo seguito)

Per concludere il discorso sugli approcci regolativi dell'intelligenza artificiale occorre infine dare conto dei lavori svolti dalla CEPEJ in tal senso e, soprattutto, dei principi sanciti nel primo documento europeo che ha accostato intelligenza artificiale e giustizia, *l'European Ethical Charter on the use of artificial intelligence in judicial systems and their environment* (Carta etica).

La CEPEJ (Commissione europea per l'efficienza della giustizia) è stata istituita nel 2002 con l'obiettivo di migliorare il funzionamento dei sistemi giudiziari degli Stati membri⁴⁶² attraverso strumenti tradizionalmente adottati dal Consiglio d'Europa tra cui: analizzare i risultati e i dati dei sistemi giudiziari; identificare le difficoltà che incontrano; definire modalità concrete per migliorare, da un lato, la valutazione dei risultati e, dall'altro, il funzionamento dei sistemi stessi; fornire assistenza agli Stati membri, su loro richiesta; proporre alle istituzioni competenti del Consiglio d'Europa i settori in cui sarebbe auspicabile elaborare nuovi strumenti normativi.

Tra i compiti della CEPEJ c'è altresì quello di esplorare il potenziale offerto dalle nuove tecnologie per migliorare l'efficienza della giustizia, compito che almeno dal 2015 la Commissione svolge con costanza e con la produzione di numerosi studi e documenti, il cui principale *focus* risiede nell'accompagnare gli Stati membri verso una sempre maggiore digitalizzazione della giustizia in conformità, soprattutto, alle regole del giusto processo *ex art. 6 CEDU*⁴⁶³.

⁴⁶¹ L. COTINO HUESO, *The Council of Europe's Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*, in CERIDAP, 9 settembre 2024.

⁴⁶² COUNCIL OF EUROPE - COMMITTEE OF MINISTERS, *Resolution Res(2002)12 establishing the European Commission for the efficiency of justice (CEPEJ)*, 18 September 2002.

⁴⁶³ Così si esprime la CEPEJ nell'*Action Plan 2022-2025, "Digitalisation for a better justice"*, dicembre 2021. Per una panoramica degli altri studi della CEPEJ sulla digitalizzazione e sulla

Tra questi assume particolare rilevanza ai nostri fini, come detto, la Carta etica, adottata dalle CEPEJ nella sessione plenaria del 3-4 dicembre 2018⁴⁶⁴, che fissa cinque principi sull'uso dell'intelligenza artificiale nei sistemi giudiziari che dovrebbero «essere sottoposti a regolare applicazione, monitoraggio e valutazione da parte di attori pubblici e privati, al fine del continuo miglioramento delle prassi»⁴⁶⁵.

Significativamente, il primo principio enunciato dalla Carta etica è il *rispetto dei diritti fondamentali*. La Carta precisa a tal riguardo che gli strumenti di IA utilizzati per *dirimere una controversia* non devono pregiudicare le garanzie di accesso al giudice e, in generale, il diritto a un *fair trial* (soprattutto quanto a *parità delle armi*, diritto *al contraddittorio e indipendenza dei giudici* nel loro processo decisionale)⁴⁶⁶.

Strettamente connesso al principio del rispetto dei diritti umani è il principio di *non discriminazione*: tenuto conto che i sistemi di intelligenza artificiale operano grazie all'elaborazione e al raggruppamento di grandi quantità di dati occorre evitare di immettere all'interno della macchina informazioni codificate che possano riprodurre e aggravare le discriminazioni già esistenti. In questo senso – prosegue la Carta – sarebbe necessaria una pregnante vigilanza e un continuo aggiornamento di tali sistemi, in particolar modo quando il sistema artificiale elabora dati sensibili della persona.

Il terzo principio enunciato dalla Carta Etica è il principio di *qualità e sicurezza*, che richiede l'instaurazione di un ambiente tecnologico sicuro. Tale deve essere perseguito attraverso un approccio quanto più possibile interdisciplinare: a tal fine,

cyberjustice si v. E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit. pp. 94-106.

⁴⁶⁴ Vista l'importanza della Carta, quasi ogni lavoro che si occupa di intelligenza artificiale e giustizia si premura di citarla e commentarla. Tra i numerosi contributi, si cfr., per la lucidità dei commenti, C. BARBARO, *Uso dell'intelligenza artificiale nei sistemi giudiziari: verso la definizione di principi etici condivisi a livello europeo?*, cit. pp. 189 ss.; S. QUATTROCOLO, *Quesiti nuovi e soluzioni antiche? Consolidati paradigmi normativi vs rischi e paure della giustizia digitale "predittiva"*, in *Cassazione penale*, fasc. 4, 2019, pp. 1748 ss.; S. QUATTROCOLO, *Intelligenza artificiale e giustizia: nella cornice della Carta etica europea gli spunti per un'urgente discussione tra scienze penali e informatiche*, in www.lalegislazionepenale.it, 18 dicembre 2018.

⁴⁶⁵ Carta Etica, p. 5 (traduzione nostra).

⁴⁶⁶ Su queste problematiche si cfr. *infra*, cap. VIII.

i progettisti di algoritmi di *machine learning* dovrebbero fare affidamento sulle competenze di professionisti, sia in ambito giuridico (giudici, avvocati, pubblici ministeri...), sia nel campo delle scienze sociali quali economisti, sociologi e filosofi. Inoltre – e tale aspetto assume notevole importanza per quanto riguarda la giustizia predittiva (*supra*, cap. II) – i dati derivanti da precedenti decisioni giudiziarie immessi in un *software* di IA dovrebbero provenire da fonti certificate e non dovrebbero poter essere modificati fino a che il processo di apprendimento della macchina non si concluda. Ciò, ovviamente, per garantire la tracciabilità del procedimento e per scongiurare il rischio di alterare il contenuto dei dati trattati.

Il principio di *trasparenza, imparzialità ed equità* si sostanzia invece nella ricerca di un equilibrio tra i diritti di proprietà intellettuale sugli algoritmi e l'ineludibile necessità che il processo decisionale sia trasparente, imparziale ed equo. Le soluzioni sono diverse: si potrebbe pensare a una totale trasparenza tecnica (e dunque l'accesso ai dati della macchina) ovvero, qualora ciò non fosse possibile, sarebbe necessario perlomeno garantire almeno la spiegazione, in un linguaggio comprensibile, del funzionamento dell'algoritmo e del procedimento seguito (vedremo nei successivi capitoli quali sono le soluzioni ad oggi adottate, sia negli Stati Uniti sia in Europa e in Italia).

In ultimo la Carta etica enuncia il principio del *controllo da parte dell'utilizzatore (under user control)*, secondo cui l'autonomia degli utilizzatori (e quindi i giudici) dei sistemi di IA non deve essere in alcun modo eliminata o limitata ma, semmai, accresciuta. A questo proposito, il giudice dovrebbe essere in grado in ogni momento di rivedere la decisione assunta dalla macchina e di non essere vincolato a essa qualora il caso concreto richieda una decisione diversa. Inoltre, il soggetto sottoposto a una determinazione algoritmica dovrebbe essere sempre informato del carattere vincolante o meno dell'*output* del sistema, mantenendo il diritto di opporvisi per far giudicare il caso da un tribunale che rispetti i requisiti previsti dall'art. 6 CEDU. Inoltre – e questo aspetto risulta, come già visto, di primaria importanza nel presente lavoro – dovrebbero essere previsti programmi di *alfabetizzazione* informatica, destinati agli utilizzatori, e dibattiti che coinvolgano i professionisti della giustizia.

Pur sancendo principi privi di un contenuto giuridico obbligatorio, il valore e l'impatto della Carta etica è stato ed è significativo. Essa ha avuto il merito di indirizzare un dibattito scientifico che fino ad allora aveva difficoltà a identificare le reali problematiche costituzionali derivante dall'implementazione di algoritmi all'interno dei sistemi di giustizia: la messa a punto dei predetti principi – senz'altro incompleti e da saggiare man mano che la giustizia algoritmica si affaccerà realmente nel panorama europeo – ha permesso alla dottrina di interrogarsi sulle opportunità e i rischi di tale mutamento di paradigma, interpretando tali principi e animando così un dibattito necessario.

Inoltre, la Carta lega strettamente lo sviluppo della giustizia algoritmica ai diritti fondamentali sanciti dalla CEDU, così facendo eco a norme che sono invece vincolanti e cogenti che dovrebbero indirizzarne il futuro sviluppo⁴⁶⁷.

Non ultimo, è evidente l'impatto della Carta nello sviluppo della normativa dell'Unione europea che, come visto e come ancora si vedrà, contiene al suo interno delle norme che risentono dell'approccio seguito dalla CEPEJ.

Del resto, l'interlocuzione tra la CEPEJ e la Commissione europea sui temi dell'intelligenza artificiale è dimostrata dallo sviluppo successivo della Carta etica in cui la CEPEJ, all'interno di alcuni documenti ha messo a punto i principi espressi nella stessa. Al riguardo, in un documento di dicembre 2021⁴⁶⁸ la CEPEJ, anche «*in light of discussions with other partners working on AI, in particular the European Commission*», che aveva qualche mese prima adottato la Proposta di regolamento sull'IA (si v. *supra*, § 3), raccomanda uno specifico piano di azione per migliorare l'implementazione dei principi della Carta Etica, con cinque *step* da seguire: la predisposizione di uno strumento di valutazione pratico (*assessment tool*) per l'operatività della Carta⁴⁶⁹; la messa a punto di un progetto pilota sulla valutazione

⁴⁶⁷ In questo senso cfr. E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit. p. 288.

⁴⁶⁸ CEPEJ, *Revised roadmap for ensuring an appropriate follow-up of the CEPEJ Ethical Charter on the use of artificial intelligence in judicial systems and their environment*, 2023.

⁴⁶⁹ Esso è stato di recente adottato. Si cfr. CEPEJ, *Assessment Tool for the Operationalisation of the European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment*, dicembre 2023. Si tratta di uno strumento pratico (un questionario di 29 domande) volto a sensibilizzare e identificare le potenziali violazioni di un sistema di IA ai principi della Carta CEPEJò. Esso identifica dei rischi di utilizzo dell'IA nei sistemi giudiziari ed è destinato a chi

di conformità di un sistema di IA ai principi della Carta; l'istituzione di un Comitato consultivo per l'intelligenza artificiale (*Artificial Intelligence Advisory Board, AIAB*) della CEPEJ per monitorare l'effettivo emergere delle applicazioni dell'IA nel settore giudiziario, discutere le problematiche attuali e proporre nuove strategie sull'uso dell'IA nel sistema giudiziario nel rispetto dei diritti fondamentali; l'istituzione di un Centro risorse (*Resource Centre on (Cyberjustice and) Artificial Intelligence*) accessibile al pubblico, regolarmente aggiornato dal comitato consultivo, in cui registrare le applicazioni esistenti nel campo della giustizia che utilizzano l'intelligenza artificiale; attività di formazione nei confronti di tutti coloro che sono nel circuito dell'IA, dagli sviluppatori di strumenti informatici ai potenziali utenti delle applicazioni (giudici, pubblici ministeri, avvocati...).

7. Breve nota conclusiva: l'azione regolatoria europea come sentiero sicuro della giustizia algoritmica

Dopo numerosi tentativi di regolamentazione *soft*, l'Europa si è dotata di due strumenti normativi vincolanti (con destinatari evidentemente diversi), il Regolamento sull'Intelligenza artificiale adottato dall'Unione europea e la Convenzione quadro sull'intelligenza artificiale adottato da Consiglio d'Europa⁴⁷⁰. L'obiettivo è quello di consegnare agli ordinamenti europei, ma con ambizioni globali, degli *standard* normativi che non abbiamo al centro esclusivamente le comprensibili ambizioni di mercato e innovazione, ma lo sforzo di conformare lo sviluppo al rispetto dei diritti umani, della democrazia e dello stato di diritto. Come efficacemente affermato, «*Europe wants to position itself as a shining beacon, a cornerstone in EU AI Act not only for the whole continent but also globally*»⁴⁷¹.

all'interno dell'amministrazione giudiziaria si occupa di implementare sistemi artificiali all'interno della giurisdizione.

⁴⁷⁰ Sul confronto tra i due atti si cfr. J. ZILLER, *The Council of Europe Framework Convention on Artificial Intelligence vs. the EU Regulation: two quite different legal instruments*, in CERIDAP, 2/2024.

⁴⁷¹ L. COTINO HUESO, *The Council of Europe's Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*, in CERIDAP, 9 settembre 2024.

Come detto sia l'*AI Act* sia la Convenzione hanno ambizioni che oltrepassano i confini europei, avendo l'aspirazione di influenzare la regolazione di paesi terzi. Soprattutto la Convenzione propone un testo particolarmente flessibile e "di principio" che mira a incentivarne l'adesione: in quest'ottica, per l'elaborazione della Convenzione quadro il Comitato dei Ministri del Consiglio d'Europa ha aperto la negoziazione a paesi terzi che hanno partecipato all'adozione del testo, e non è un caso che, per esempio, sono previste più modalità flessibili per garantire il rispetto del trattato anche per gli attori privati dell'IA: la ricerca di un testo condiviso anche con paesi estranei al Consiglio d'Europa, e soprattutto gli Stati Uniti, ha infatti avuto l'effetto di mitigarne la rigidità, in particolare nel settore privato⁴⁷².

Proprio gli Stati Uniti, come visto in questo lavoro, vivono nell'ambiguità di avere l'industria di intelligenza più sviluppata al mondo non accompagnata da una normativa in grado di mitigarne i rischi; la firma al trattato del Consiglio d'Europa ha sicuramente una portata simbolica ma la speranza è anche quella di spingere il legislatore statunitense ad accompagnare lo sviluppo economico con delle regole giuridiche più stringenti.

Detto questo, non è possibile prevedere quanto e quando la legislazione europea sarà in grado di dispiegare il suo potenziale ma non è peregrino ritenere che l'influenza sia sulla legislazione globale sia sull'interpretazione da parte delle Corti possa profondamente cambiare.

Lo ripetiamo, siamo realmente in una fase ancora iniziale di comprensione del fenomeno dell'intelligenza artificiale e di quali siano le migliori strategie normative per regolarlo. L'Europa sembra essere il continente in cui vi è oggi la volontà politica più decisa per dotare tutto il contesto mondiale di regole incentrate sul rispetto di diritti fondamentali. Ci sembra di poter dire, tornando al più circoscritto campo di indagine del presente lavoro, che l'approccio regolativo europeo sembra il sentiero più sicuro per lo sviluppo di una giustizia algoritmica che sia anch'essa rispettosa dello stato di diritto e dei diritti fondamentali. In uno spazio oggi così

⁴⁷² Si cfr. ancora J. ZILLER, *The Council of Europe Framework Convention on Artificial Intelligence vs. the EU Regulation: two quite different legal instruments*, cit., p. 214.

decisivo per la tutela dei diritti, e cioè il processo giurisdizionale, lasciare la regolamentazione in mano a singoli attori, privati o pubblici che siano, senza una visione organica del fenomeno, lascerebbe da una parte gli organi giurisdizionali privi di coordinate chiare, dall'altra le parti del processo senza fondamentali strumenti di tutela (come si vedrà analizzando la giurisprudenza americana nella prossima parte del lavoro).

PARTE III

GIUSTIZIA ALGORITMICA E GIUSTO PROCESSO

Nota introduttiva: la giustizia algoritmica nel prisma dell'AI Act

Nel capitolo introduttivo (§3) si è segnalata la prospettiva costituzionalistica del lavoro e si poneva in luce come le tensioni maggiori tra l'impiego di algoritmi nel settore della giustizia e le garanzie costituzionali sorgessero nei casi in cui gli algoritmi incidono sul processo decisionale del giudice. Come in quella sede precisato, è lo stesso Regolamento sull'IA che impone una normazione più rigida per tali algoritmi. Conviene qui quindi estendere tali premesse per chiarire quale prospettiva (normativa) assume il Regolamento e inserire al suo interno la prospettiva (costituzionalistica) di questo lavoro.

Come visto, gli strumenti di IA destinati a venire utilizzati da un'autorità giudiziaria sono considerati "ad alto rischio" (Art. 6, § 2 e Allegato III del Regolamento), e tale classificazione impone il rispetto di una serie di requisiti (si v. *supra* cap. V § 3).

Ciò nondimeno, l'art. 6 § 3 del Regolamento, pone una deroga a tale classificazione quando il sistema non comporta «un rischio significativo di danno per la salute, la sicurezza o i diritti fondamentali delle persone fisiche» e ciò «*anche nel senso di non influenzare materialmente il risultato del processo decisionale*» (corsivo nostro).

È lecito chiedersi che cosa intenda il legislatore con "materialmente" (in modo analogo, in inglese viene impiegata la parola *materially*), data la vaghezza del suo significato: è lo stesso legislatore europeo a tentare di concretizzare il contenuto della norma, specificando che, perché il sistema non sia considerato ad alto rischio, deve verificarsi *almeno una* delle seguenti condizioni (art. 6 § 3 comma 2):

- a) il sistema di IA svolge un compito procedurale limitato;

b) sistema di IA è destinato a migliorare il risultato di un'attività umana precedentemente completata;

c) il sistema di IA è destinato a rilevare schemi decisionali o deviazioni da schemi decisionali precedenti e non è finalizzato a sostituire o influenzare la valutazione umana precedentemente completata senza un'adeguata revisione umana;

d) il sistema di IA è destinato a eseguire un compito preparatorio per una valutazione pertinente ai fini dei casi d'uso elencati nell'allegato III.

In tutti questi casi il Regolamento presume quindi che un sistema ad alto rischio non abbia un'influenza sul processo decisionale dell'utilizzatore e quindi non debba essere classificato come *high risk*.

A cornice dell'articolato del Regolamento, seguendo la medesima impostazione, il *considerando n. 61* specifica che, per quanto riguarda gli strumenti di IA nel campo della giustizia, la loro classificazione ad alto rischio non dovrebbe comunque comprendere strumenti destinati ad attività amministrative puramente accessorie «che non incidono sull'effettiva amministrazione della giustizia nei singoli casi, quali l'anonimizzazione o la pseudonimizzazione di decisioni, documenti o dati giudiziari, la comunicazione tra il personale, i compiti amministrativi».

Come si vede, si tratta di attività che sì, in senso lato, potrebbero avere un'influenza sulla decisione del giudice (una migliore organizzazione dei tribunali e delle singole attività amministrative al suo interno, per esempio, potrebbero certamente comportare una migliore organizzazione del singolo giudice in termini di tempo e concentrazione da dedicare al singolo caso da decidere), ma non nel senso che sembra ritenere il legislatore europeo. E forse è qui che si comprende il termine *materialmente*: devono essere sistemi che si inseriscono in modo *diretto* nel processo decisionale, sottoposti in modo *diretto* all'attenzione e alla considerazione del singolo giudice o collegio giudicante.

Quando ciò accade – e vi è quindi una decisione influenzata materialmente da un algoritmo – le ricadute sul rispetto delle garanzie del giusto processo perdono quell'aspetto di fumosità e vaghezza che sembrano alle volte assumere nelle

riflessioni su intelligenza artificiale e giustizia e, come si dirà, emergono limpide in tutta la loro concretezza. Senza qui ripetere quanto già specificato nel capitolo introduttivo, alcune delle predette garanzie sono correttamente identificate dal legislatore europeo (e confrontarsi con esse sarà dunque necessario): imparzialità del giudice, diritto a un ricorso effettivo, diritto di difesa (*Considerando nn. 48, 59 e 61*).

Su queste – e altre – implicazioni costituzionali si costruirà il ragionamento dei prossimi capitoli, muovendo dalla copiosa giurisprudenza statunitense che da anni, su queste tematiche, ha tentato di fornire delle soluzioni – non di rado insoddisfacenti – per decidere sulle violazioni del *due process of law* determinate dall'impiego di algoritmi predittivi (cap. VIII). Ciò, tuttavia, non prima di aver correttamente individuato quali siano oggi i principali problemi di natura teorica e pratica degli algoritmi impiegati nei sistemi di giustizia (cap. VI) e aver fornito un quadro delle garanzie processuali sia del *due process* statunitense, sia del giusto processo nella prospettiva costituzionale di tutela multilivello dei diritti fondamentali (cap. VII).

I NODI PROBLEMATICI TEORICI E PRATICI DEGLI ALGORITMI IMPIEGATI NELLA GIUSTIZIA

SOMMARIO: Premessa - 1. Algoritmi e dati - 1.1. La disponibilità e la selezione dei dati - 2. L'opacità algoritmica - 2.1. L'opacità intrinseca dei sistemi di IA - 2.2. Il ruolo dei privati e il *trade secret* - 3. L'autorità dell'algoritmo nei processi decisionali e l'*automation bias* - Breve conclusione: l'intreccio dei nodi

Premessa

Di “società algoritmica” e delle sue implicazioni sul diritto costituzionale si parla da tempo e la letteratura è in costante aumento⁴⁷³. Se ne parla ponendo in luce gli aspetti che sono risultati nell'esperienza più sintomatici di una tensione crescente con il rispetto dei diritti fondamentali: esemplificando, la crescita esponenziale e accumulatrice dei dati e il loro trattamento per fini di profilazione o financo di sorveglianza⁴⁷⁴; i problemi relativi all'opacità dei sistemi di IA; il ruolo dei privati nella raccolta e selezione dei dati e la progettazione di essi; le potenzialità discriminatorie degli algoritmi. Tutte queste caratteristiche vengono spesso calate nel contesto del diritto (e dei diritti) mettendone in luce le principali e crescenti preoccupazioni.

Ora, l'orizzonte più limitato di questo lavoro mira a evidenziare quali paiono oggi essere i noti problematici più impellenti degli strumenti algoritmici che ad oggi sono impiegati nei sistemi di giustizia, individuabili sia attraverso le elaborazioni teoriche degli specialisti, sia attraverso le esperienze pratiche che già si sono analizzate. Peraltro la loro individuazione, arrivati a questo punto del lavoro, è

⁴⁷³ Si cfr. i contributi citati *supra* alla nota 27.

⁴⁷⁴ S. ZUBOFF, *The Age of Surveillance Capitalism, The Fight for a Human Future at the New Frontier of Power* London, 2019. Nello spiegare il meccanismo del capitalismo della sorveglianza, l'Autrice afferma che «*although some of these data are applied to product or service improvement, the rest are declared as a proprietary behavioral surplus, fed into advanced manufacturing processes known as “machine intelligence,” and fabricated into prediction products that anticipate what you will do now, soon, and later. Finally, these prediction products are traded in a new kind of marketplace for behavioral predictions that I call behavioral futures markets*».

altresì intuibile da come oggi le istituzioni hanno deciso di regolare la giustizia algoritmica, soprattutto nel continente europeo. I principi di qualità e sicurezza dei dati, le varie declinazioni del principio di trasparenza nei diversi strumenti normativi, i principi relativi al *training* e alla formazione di coloro che sono chiamati a utilizzare sistemi algoritmici sono la conseguenza delle caratteristiche proprie degli algoritmi e del rapporto tra contributo umano e contributo algoritmico.

Non solo, anche quando alcune forme di giustizia algoritmica sono state sindacate dai giudici – come si vedrà nell’esperienza americana – nelle sentenze sono stati individuati pressoché i medesimi problemi.

Prima di interrogarsi sulla compatibilità della giustizia algoritmica con le garanzie costituzionali, quindi, occorre in questo capitolo chiarire quali siano i motivi di fondo per cui tali interrogativi meritano approfondimento nella prospettiva di questo lavoro e nello specifico di questa parte.

Facendo un’opera di astrazione e categorizzazione ci sembra di poter individuare tre distinte questioni che meritano di essere evidenziate e che sono in strettissima correlazione tra loro: il problema dei dati e le sue numerose implicazioni, il problema della trasparenza (*recte* opacità) algoritmica; e il problema dell’autorità che l’algoritmo esercita sul percorso decisionale.

Si è ritenuto di concentrare l’approfondimento su tali tematiche non solo perché sono al centro del dibattito teorico tra gli studiosi, ma soprattutto perché hanno dimostrato di avere una notevole importanza pratica: come si vedrà nell’analisi della giurisprudenza americana sono tutte questioni che assumono grande importanza nel sindacato dei giudici americani sulla compatibilità degli algoritmi con le garanzie del giusto processo. Inoltre, come detto, tutti gli strumenti normativi tentano di trovare soluzioni per fronteggiare la *governance* dei dati, l’opacità algoritmica e l’idoneità dell’algoritmo a deresponsabilizzare l’essere umano.

Rispetto all’approccio induttivo ed empirico segnalato nel capitolo introduttivo, tali questioni accomunano tutti i sistemi algoritmici oggi impiegati, sicché è possibile discorrerne in termini generali: non si rinuncerà tuttavia, quando necessario, a sottolineare i diversi contorni che queste possono assumere rispetto ai diversi impieghi degli algoritmi nel settore della giustizia.

1. Algoritmi e dati

Nel 2006 un matematico britannico, Clive Humby, coniava la frase, diventata celebre, «*Data is the new oil*» («I dati sono il nuovo petrolio»). Sebbene spesso la frase venga intesa monodimensionalmente nel senso di paragonare il valore dei dati al valore del petrolio, in realtà il paragone mirava a porre in luce un'altra similitudine: come il petrolio, che per creare valore deve essere trasformato in gas, plastica, prodotti chimici ecc., anche i dati sono preziosi ma devono essere “raffinati”, cioè elaborati e ordinati perché abbiano valore e possano utilmente venire impiegati.

Impiegati da chi?

Come già visto gli algoritmi – e in particolare gli algoritmi di intelligenza artificiale che impiegano tecniche di apprendimento automatico – eseguono i loro compiti attraverso l'elaborazione dei dati, per funzionare hanno bisogno di un'ampia disponibilità di dati, che costituiscono la condizione *sine qua non* per lo sviluppo dell'IA. Non a caso tale tecnologia ha vissuto il suo rilancio definitivo, e ormai sembra irrefrenabile, prima con l'avvento di internet e poi con lo sviluppo continuo di tecnologie ICT (*information and communication technologies*). Ciò ha consentito una produzione e diffusione di dati sulla rete senza precedenti e in continuo aumento (si parla, come già visto di epoca dei *Big data*).

In ambito giuridico, come visto, il processo è stato più lento ma da anni, in Europa ma anche in Italia, si è avuto un processo di digitalizzazione e datificazione necessariamente prodromico all'idea di inserire procedure algoritmiche all'interno dei sistemi giudiziari.

Si aprono a questo punto due distinti problemi, uno legato alla *disponibilità dei dati* e uno alla *selezione dei dati*.

1.1. La disponibilità e la selezione dei dati

Senza i dati gli algoritmi non hanno modo di funzionare.

La disponibilità dei dati è quindi il primo aspetto che occorre considerare per valutare la realizzabilità delle applicazioni di IA nei settori della giustizia. Nella Carta Etica, più volte menzionata, si afferma infatti che «*the availability of data is an essential condition for the development of AI, enabling it to perform certain tasks previously carried out by humans in a non-automated manner. The more data available, the more AI is able to refine models improving their predictive ability. An open data approach to judicial decisions is therefore a prerequisite for the work of legal tech companies specialising in search engines or trend analysis (“predictive justice”)*»⁴⁷⁵.

La disponibilità dei dati giudiziari è quindi un primo aspetto che merita considerazione: mentre in Cina si rendono disponibili al pubblico milioni di dati giudiziari e negli Stati Uniti ci sono da anni i primi progetti (come il *Caselaw Access project* della *Harvard Law School*) che mirano digitalizzare e rendere liberamente disponibile il panorama giurisprudenziale sia federale sia statale, in Italia solo di recente, sulla spinta del PNRR e degli obiettivi di digitalizzazione della PA e della giustizia, si sono cominciate a costruire banche dati liberamente accessibili e il più complete possibili (e quindi utili per l’addestramento di algoritmi).

Come è stato rilevato, solo la pubblicazione e l’accesso pubblico dell’intero patrimonio delle decisioni giurisprudenziali italiane può far sì che i progetti di applicazioni di *machine learning* relative alla giustizia predittiva evolvano dallo stadio sperimentale attuale e diventino e uno strumento potenzialmente utile per i giudici⁴⁷⁶.

La disponibilità di dati elaborabili non è tuttavia sufficiente per progettare un sistema di IA, occorre ovviamente selezionarli, classificarli e implementarli per l’addestramento di un sistema, nel caso di *machine learning*, o più semplicemente fornirli (in formato leggibile) al sistema algoritmico per fare in modo che quest’ultimo esegua i compiti previsti. Come si intuisce questa attività richiede tutt’oggi un intervento umano decisivo, composto da scelte che non possono in

⁴⁷⁵ Carta Etica, p. 18

⁴⁷⁶ A. SANTOSUOSSO - G. SARTOR, *La giustizia predittiva: una visione realistica*, cit.

alcun caso essere considerate neutrali⁴⁷⁷. Gli esseri umani esercitano ancora, in altre parole, un controllo sostanziale su una miriade di aspetti del funzionamento di un algoritmo. Sono ancora tenuti a fornire all'algoritmo i dati o quantomeno indicare lo spazio in cui il sistema li acquisisca, a costruire un particolare tipo di algoritmo e a mettere a punto i dettagli del processo di ottimizzazione dello stesso.

Gli esseri umani svolgono un ruolo importante anche nel *training* dell'algoritmo e nella valutazione delle prestazioni dello stesso. Dopo l'addestramento, è sempre un essere umano che valuta l'accuratezza dell'algoritmo apportando le modifiche necessarie per aumentare l'accuratezza⁴⁷⁸.

Il contributo decisivo dell'essere umano quando si parla di dati aiuta a comprendere che spesso, quando si discute giustizia algoritmica, si concentra tutta l'attenzione sull'algoritmo, mentre la scelta e la disponibilità dei dati è lasciata in secondo piano. Invece, se è vero che già l'algoritmo potrebbe essere discriminatorio (pensiamo a un sistema di reclutamento del personale che scarta automaticamente tutte le persone candidate di sesso femminile), esso potrebbe invece funzionare correttamente ma essere istruito e allenato con dati non rappresentativi o essi stessi discriminatori o inadeguati. Insomma, dall'intersezione tra progettazione degli algoritmi, da una parte, e selezione dei dati, dall'altra, le moderne tecnologie di IA non sono affatto oggettive o neutrali.

Come affermava Hildebrandt, professoressa esperta di *law and technology*, la tecnologia «*in itself is neither good nor bad, it is never neutral*», ma che anzi essa «*invites certain behaviors and inhibits others or even enforces certain behaviors while prohibiting others*»⁴⁷⁹.

Solo per spiegare il concetto con un esempio famoso, già negli anni 80' del secolo scorso Langdon Winner, celebre politologo che indagava le intersezioni tra potere e tecnologia, ha notoriamente sostenuto la non neutralità della tecnologia utilizzando l'esempio di un cavalcavia autostradale a bassa quota di Long Island. Il

⁴⁷⁷ Sulla percezione dell'IA come neutrale ed esatta si v. *ex multis*, T. ARAUJO *et al.*, *In AI we trust? Perceptions about automated decision-making by artificial intelligence*, in *AI and Society*, 35, 2020, 611-623.

⁴⁷⁸ C. COGLIANESE - D. LEHR, *Transparency and Algorithmic Governance*, cit., p. 15.

⁴⁷⁹ M. HILDEBRANDT, *Technology and the end of law*, in B. KEIRSBILCK - W. DEVROE - E. CLAES (a cura di), *Facing the limits of the law*, Springer, 2009, pp. 443-464, spec. p. 451.

cavalcavia era situato abbastanza in basso da impedire agli autobus di passare sotto il ponte autostradale e la strada che passava sotto il cavalcavia era l'unico metodo di viaggio verso una specifica zona della città, tagliando di fatto fuori da quella parte della città coloro che non possedevano un'auto (spesso a quel tempo persone non bianche e meno abbienti): in questo modo, spiegava l'Autore, il cavalcavia diventava una barriera tecnologica che consentiva la discriminazione di alcuni gruppi di persone⁴⁸⁰.

La tesi della non neutralità si applica a tutte le tecnologie, il che significa che gli algoritmi che popolano il nostro mondo digitale, e che verosimilmente saranno presenti in maniera consistente all'interno delle giurisdizioni, sono strumenti carichi di scelte umane, tutt'altro che neutre. Non è un caso che, come sottolineato anche da alcune normative che abbiamo analizzato *supra* nel capitolo V, spesso i legislatori hanno ragionato sulla necessità che la progettazione degli algoritmi e la *governance* dei dati non siano affidate solo a esperti di ingegneria o informatica, ma che siano invece frutto della collaborazione di attori provenienti da più campi scientifici, dall'etica, alla filosofia, al diritto.

In definitiva, quindi, gli algoritmi riflettono un giudizio umano e una scelta umana «*because human beings program predictive algorithms, their biases and values are embedded into the software's instructions, known as the source code and predictive algorithms*»⁴⁸¹.

Tenuto conto di quanto detto, il primo aspetto fondamentale da considerare è la qualità del dato, se esso cioè sia idoneo a rendere il sistema algoritmico capace di prendere decisioni o fornire previsioni adeguate rispetto all'obiettivo assegnato al sistema: è fondamentale quindi garantire che il *dataset* utilizzato nel processo di apprendimento sia accurato e adeguato allo scopo prefissato. Se il *dataset* è difettoso, il programma non riuscirà a prendere decisioni adeguate (*garbage in, garbage out*).

⁴⁸⁰ L. WINNER, *Do artifacts have politics?* Daedalus, 109 (1), 1980, pp. 121-136.

⁴⁸¹ D. K. CITRON - F. PASQUALE, *The scored society: Due process for automated predictions*, in *Washington Law Review*, 89, 1, 2014, pp. 1-33, spec. p. 4.

Pensiamo al funzionamento di alcuni sistemi di predizione del rischio di recidiva: abbiamo assunto COMPAS come strumento paradigmatico (si v. *supra*, cap. I, § 4). In quella occasione abbiamo spiegato che, sebbene non fosse accessibile né il *dataset* né il codice algoritmico, nella Guida si chiariva che il *dataset* era costituito dai dati di arresti e esecuzioni penali di un circoscritto periodo di tempo (2004-2005).

Ora, sono subito evidenti alcune importanti conseguenze rispetto alla selezione dei dati raccolti: innanzitutto non sembra ci sia alcun aggiornamento dei dati rispetto a utilizzi che si collocano molto più avanti nel tempo rispetto alla situazione carceraria nel periodo di raccolta dati. In secondo luogo sono dati *non rappresentativi*, e ciò perché addestrano un sistema che viene utilizzato in diverse giurisdizioni con lo stesso *dataset*: come sollevato in giudizio dalle difese degli imputati (si v. *infra*), si lamenta il fatto che il sistema predittivo impiegato nella giurisdizione del Wisconsin dovrebbe essere addestrato sui dati del Wisconsin e non su dati di altre giurisdizioni. In terzo luogo, una raccolta dei dati della situazione carceraria negli Stati Uniti non può che riflettere le scelte – squisitamente umane – che hanno preceduto l'impiego dell'algoritmo.

Ciò significa che se fino al momento della raccolta dati il sistema penale americano aveva delle inclinazioni discriminatorie, anche il sistema algoritmico rifletterà tali inclinazioni. Non è un caso che proprio COMPAS è stato sottoposto ad alcune analisi da parte di organizzazioni indipendenti che hanno rilevato la tendenza discriminatoria dell'algoritmo⁴⁸².

Nel senso delle potenzialità discriminatorie dei dati inseriti nei *risk assessment tool*, nel 2014 l'*Attorney General* degli Stati Uniti si esprime così: «*Although these measures were crafted with the best of intentions, I am concerned that they may inadvertently undermine our efforts to ensure individualized and equal justice. By basing sentencing decisions on static factors and immutable characteristics – like the defendant's education level, socioeconomic background, or neighborhood –*

⁴⁸² J. ANGWIN - J. LARSON - S. MATTU - L. KIRCHNER, *Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks*, cit.

they may exacerbate unwarranted and unjust disparities that are already far too common in our criminal justice system and in our society»⁴⁸³.

La selezione dei dati assume rilevanza anche nelle ipotesi di quella che abbiamo definito qui “giustizia predittiva”, protagonista oggi delle sperimentazioni nel continente europeo (si v. *supra*, cap. II).

Dimenticando per ora i problemi rispetto all'accettabilità costituzionale di siffatti strumenti, per renderli quanto meno utili per un futuro impiego nei sistemi giudiziari, occorre capire quali dati (*rectius* quali precedenti giurisprudenziali) inserire all'interno del sistema. Se infatti quei dati sono inadeguati (perché – per esempio – riflettono un orientamento che in realtà è minoritario⁴⁸⁴) è evidente che il giudice, ad oggi incapace di comprendere il funzionamento di tali sistemi, si potrebbe affidare a un risultato che non riflette minimamente la giurisprudenza maggioritaria su una determinata controversia.

Oltre alla qualità del dato, di vitale importanza nell'ipotesi di giustizia algoritmica è *chi* si occupa della scelta e dell'implementazione del *dataset*. E ciò per almeno due distinte ma connesse ragioni: se, come oggi frequentemente accade, la progettazione degli algoritmi e il loro addestramento è lasciato nelle mani di soggetti *privati*, si pongono notevoli problemi di trasparenza del sistema e del rischio di *esternalizzare* la responsabilità del giudice verso i progettisti dell'algoritmo (si v. *infra*, cap. VIII § 5); ma anche se la progettazione e l'addestramento di sistemi di giustizia predittiva viene affidata a soggetti pubblici – abbiamo visto che in alcune giurisdizioni italiane la maggior parte dei sistemi algoritmici vengono progettati e allenati dal Ministero della Giustizia, spesso in collaborazione con le Università – occorre chiedersi se un sistema di giustizia predittiva fornito al giudice dal Ministero possa in qualche modo minare alle garanzie di indipendenza e di separazione dei poteri che sono alla base del nostro sistema costituzionale (si v. *infra*, cap. VIII § 6).

⁴⁸³ ERIC HOLDER, U.S. Attorney Gen., *Speech at the National Association of Criminal Defense Lawyers 57th Annual Meeting and 13th State Criminal Justice Network Conference*, 1° agosto 2014.

⁴⁸⁴ A. SANTOSUOSSO - G. SARTOR, *La giustizia predittiva: una visione realistica*, cit., p. 1775.

2. L'opacità algoritmica

Di opacità algoritmica la letteratura, sia tecnica⁴⁸⁵ sia giuridica⁴⁸⁶, ha da tempo concentrato una parte consistente delle sue riflessioni.

L'importanza che assume la trasparenza nella "società algoritmica" nel suo insieme si intuisce anche dagli strumenti normativi che abbiamo analizzato nella seconda parte di questo lavoro: sia nei primi tentativi di regolazione da parte degli Stati Uniti, sia nella regolazione all'interno dell'Unione Europea e del Consiglio d'Europa, ci sono disposizioni che impongono la trasparenza algoritmica, anche se la soglia di tale trasparenza spesso non viene specificata.

Per descrivere l'opacità algoritmica vengono spesso usati, nel dibattito scientifico ma anche pubblico, termini interscambiabili: in particolare si fa riferimento al fatto che gli algoritmi sono "segreti", "opachi", "black box" "inintelligibili" "incomprensibili"⁴⁸⁷, e tutti rendono bene l'idea di imperscrutabilità di ciò che accade all'interno del procedimento algoritmico prima della produzione dell'*output*. In realtà, per chiarire la portata dell'opacità algoritmica occorre distinguere, ai nostri fini, almeno due diversi tipi di opacità: una è la conseguenza del funzionamento dei più moderni sistemi algoritmici di intelligenza artificiale, l'altra è una opacità *giuridica* legata ad alcuni diritti che si oppongono alla necessità della spiegabilità⁴⁸⁸.

⁴⁸⁵ A. SELBST - S. VENKATASUBRAMANIAN - I. E. KUMAR, *Deconstructing design decisions: why courts must interrogate machine learning and other technologies*, in *85 Ohio State Law Journal*, 2024, pp. 415 ss.

⁴⁸⁶ Si cfr. *ex multis*, C. COGLIANESE - D. LEHR, *Transparency and Algorithmic Governance*, cit.; A. DEEKS, *The judicial demand for explainable artificial intelligence*, in *Columbia Law Review*, Vol. 119, No. 7, 2019, pp. 1829- 1850; L. EDWARDS - M. VEALE, *Enslaving the Algorithm: From a "Right to an Explanation" to a "Right to Better Decisions"?*, in *IEEE Security & Privacy*, 16 (3), 2019, pp. 46-54.

⁴⁸⁷ A. SELBST - S. BOROCAS, *The Intuitive Appeal of Explainable Machines*, in *Fordham Law Review*, 87 (3), 2018, pp. 1089 ss.

⁴⁸⁸ Alcuni distinguono invece tre diversi tipi di opacità: «Opacity might arise because certain technologies are proprietary. Secondly, a source of opacity might be complex systems that require specialist knowledge and skill to understand them.7 3 Thirdly, there are systems that are simply naturally opaque. In machine learning, black box models are created by an algorithm directly from data. This means that humans - even those who design these systems - cannot understand how variables are being combined to make predictions», si v. W. GRAVETT, *Jailed by a "black box": the impact of opaque algorithms on the right to a fair trial in the Unites State of America*, in *84 THRHR*, 2021, pp. 299 ss, spec. p. 308.

2.1. L'opacità intrinseca dei sistemi di IA

Il primo aspetto della trasparenza che occorre analizzare riguarda i più moderni sistemi algoritmici, quelli cioè che impiegano tecniche di apprendimento automatico per produrre i propri *output*.

Esso è stato definito come il problema della *technical black box*, da affiancare alla *legal black box* che, come si vedrà, si riferisce a ostacoli legali – quali il *trade secret* – che si frappongono alla conoscenza dell'algoritmo⁴⁸⁹.

La *black box*⁴⁹⁰ tecnica si riferisce al fatto che gli algoritmi implementati nei *tools*, soprattutto quelli che fanno utilizzo di tecniche di *machine learning*, il più delle volte non sono trasparenti: infatti per gli algoritmi più moderni non è sempre facile verificare il processo matematico che dagli *input* conduce a un *output* e il peso che l'algoritmo assegna ai singoli dati elaborati⁴⁹¹. E il problema si aggrava considerando che il procedimento algoritmo non rimane sconosciuto solo all'utente finale (che può essere, per quanto qui di interesse, la parte o il giudice) ma anche allo stesso programmatore. Ciò significa che spesso nemmeno attraverso un accesso completo ai codici dell'algoritmo è possibile davvero verificare come esso abbia prodotto il proprio risultato.

La difficoltà di generare spiegazioni di algoritmi di apprendimento automatico è ulteriormente aggravata dal tipo di “*big data*” su cui essi spesso operano. Sebbene in linea di principio sia gli algoritmi di apprendimento automatico sia le tecniche più tradizionali possano operare sugli stessi insiemi di dati, la combinazione tra le tecniche induttive del *machine learning* e l'insieme di grandissimi *dataset* che

⁴⁸⁹ Tale distinzione, tra gli altri, appare in H. LIU - C. LIN - Y. CHEN, *Beyond State v Loomis: artificial intelligence, government algorithmization and accountability*, cit., pp. 135 ss.; si v. altresì W. GRAVETT, *Jailed by a “black box”: the impact of opaque algorithms on the right to a fair trial in the Unites State of America*, cit.

⁴⁹⁰ F. PASQUALE, *The Black Box Society*, Harvard University Press, 2015, dove l'Autore compara il ruolo degli algoritmi con l'allegoria della caverna di Platone, dove le persone intrappolate in essa sono solo in grado di vedere delle ombre proiettate da un fuoco alle loro spalle. I prigionieri non possono conoscere chi proietta le immagini e come vengono proiettate, ma per essi tutta la realtà è rappresentata dalle ombre.

⁴⁹¹ L. FRIDMAN - L. DING - B. JENIK - B. REIMER, *Arguing Machines: Human Supervision of Black Box AI Systems That Make Life-Critical Decisions*, 2019.

contengono interrelazioni tra i dati rende ancora più difficile spiegare come un algoritmo di apprendimento automatico produce i propri *output*⁴⁹².

In sintesi, diversamente dalla opacità prodotta da ostacoli giuridici di cui si dirà subito, l'ostacolo tecnico⁴⁹³ derivante dalle caratteristiche dei sistemi più evoluti sembrerebbe di più difficile soluzione: anche se ad oggi gli strumenti come *i risk assessment tools* o gli strumenti sperimentati nelle aule giudiziarie in Europa non sembrano essere *software* così complessi da poter risultare inaccessibili allo scrutinio umano, non vi è dubbio che la ricerca anche in campo giuridico stia andando in questa direzione, e ciò per la ragione che più si intende migliorare la *performance* del sistema algoritmico, più esso necessita di meccanismi di elaborazione complessi e di difficile comprensione.

Occorre qui comunque segnalare che la ricerca scientifica più evoluta oggi si sta concentrando sul concetto di *explainable artificial intelligence* (XAI), cercando di “spiegare” l'IA e di superare il problema della scatola nera⁴⁹⁴. I ricercatori di XAI, in particolare, stanno sviluppando modelli di IA che possano essere compresi e interpretati dagli esseri umani che possano essi stessi generare spiegazioni dei risultati dell'apprendimento automatico. In particolare, due sono le strade per garantire una maggiore trasparenza dei sistemi algoritmici moderni: l'*explainability* dei sistemi di IA che permette di comprendere le variabili e i procedimenti interni degli algoritmi, e l'*interpretability*⁴⁹⁵ – meno ambiziosa – che consiste di valutare il sistema sulla base dell'output finale, senza necessariamente comprendere ogni singolo passaggio del procedimento algoritmico⁴⁹⁶.

⁴⁹² C. COGLIANESE - D. LEHR, *Transparency and Algorithmic Governance*, cit., p. 17.

⁴⁹³ Sul tema si v. C. CASONATO, *Intelligenza artificiale e giustizia: potenzialità e rischi*, in *DPCE online*, 2020/3, p. 3369-3388.

⁴⁹⁴ A. ADADI - M. BERRADA, *Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI)*, in *6 IEEE ACCESS*, 2018. V. anche Cfr. R. GUIDOTTI ET AL. *A survey of methods for explaining black box models*, in *ACM Comput. Surv.* 51, 5, 2018; R. CALEGARI - G. CIATTO, A. OMICINI, *On the integration of symbolic and sub-symbolic techniques for XAI: A survey*, 7-32, in *Intelligenza Artificiale* 14, 1/2020.

⁴⁹⁵ A. SELBST - S. BOROCAS, *The Intuitive Appeal of Explainable Machines*, cit.

⁴⁹⁶ Una strada che di recente sta venendo intrapresa è l'uso di scenari controfattuali: il modello seleziona campioni alternativi con caratteristiche diverse, confronta i diversi risultati e identifica come e perché differiscono. Per esempio, il modello sarà in grado di rilevare che il risultato sarebbe stato diverso se la caratteristica X fosse mancata o se quella Y fosse stata aggiunta, si v., *ex plurimis*, M. HARRADON - J. DRUCE - B. RUTTEMBERG, *Casual Learning and Explanation of Deep Neural Networks via Autoencoded Activations*, 2018. V. anche A. DEEKS, *The judicial demand for*

2.2. Il ruolo dei privati e il trade secret

Un distinto aspetto dell'opacità algoritmica, come detto, ha a che vedere con il mercato dell'intelligenza artificiale e il valore commerciale degli algoritmi. Alle richieste di accesso al codice dell'algoritmo e al suo funzionamento vengono opposti i diritti di privativa e di segreto sull'algoritmo stesso, alla cui segretezza le imprese produttrici non vogliono rinunciare per ragioni economiche.

Nei casi statunitensi che analizzeremo, la tensione tra i diritti di proprietà industriale delle imprese e i diritti delle parti in giudizio sono più volte venuti in rilievo, con risultati come si vedrà non del tutto soddisfacenti. Per fornire uno sguardo di insieme, i *risk assessment tools* impiegati nelle giurisdizioni statunitensi sono normalmente progettati da società private con scopo di lucro e l'algoritmo al loro interno è contrattualmente coperto da segreto industriale (*trade secret*)⁴⁹⁷. Tali società, quali Equivant (la produttrice di COMPAS), stipulano contratti milionari di licenza annuali con gli Stati interessati per l'utilizzo dei propri algoritmi⁴⁹⁸ nei quali si riconosce espressamente il segreto industriale dei codici e l'impossibilità per lo Stato contraente di modificare, alterare, divulgare, trasferire o vendere l'algoritmo a terze parti⁴⁹⁹. Pertanto, il segreto industriale può essere invocato dalla società produttrice per proteggere i sistemi algoritmici da qualsiasi scrutinio in tutte le fasi del processo. In particolare, a rimanere segreti sono tanto l'algoritmo e il codice sorgente, quanto i dati di *training* con cui il sistema è stato addestrato. Non

explainable artificial intelligence, cit., pp. 1829 ss. Ciò detto, non è sempre possibile spiegare *ex post* i procedimenti algoritmici, sia per il fatto che non sempre è facile identificare, in un modello *black box*, i fattori decisivi che hanno portato a un determinato risultato, sia perché anche qualora sia identificato il fattore causativo, non sempre può servire quale vera e propria spiegazione del procedimento. Si cfr., sul punto, M. SCHERER, *Artificial Intelligence and legal Decision-Making: the wide open?*, in *Journal of International Arbitration*, 36, 5, 2019, pp. 563 ss.

⁴⁹⁷ Il fenomeno è mondiale, si cfr. T. SOURDIN, *Judges, Technology and Artificial Intelligence. The Artificial Judge*, cit., p. 273 secondo cui «[i]nternationally, much of the app and bot development in the justice sector is currently taking place in the private sector and is unregulated».

⁴⁹⁸ K. FREEMAN, *Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis*, cit., p. 92.

⁴⁹⁹ È disponibile *online* un contratto di licenza del 2015 sottoscritto da Northpointe (ora Equivant) e dallo Stato del Wisconsin per l'utilizzo di COMPAS. Si v. il seguente indirizzo: <https://archive.epic.org/algorithmic-transparency/crim-justice/EPIC-16-06-23-WI-FOIA-201600805-2010InitialContract.pdf>.

vi è dunque possibilità di conoscere il procedimento attraverso cui l'algoritmo ha raggiunto la propria conclusione, e ciò vale per le parti, per i propri difensori e persino per il giudice. In questo senso si parla di *legal black box*⁵⁰⁰, con ciò intendendosi gli ostacoli di natura giuridica che si frappongono alla conoscibilità dell'algoritmo.

Non solo i *risk assessment tools*, anche in alcuni esperimenti di giustizia predittiva (per i quali si rinvia al capitolo II), vi è non di rado l'intervento dei privati nella progettazione. E sembra che oggi, per le ragioni che presto diremo, tale realtà non possa venire facilmente modificata. È vero, d'altra parte, che diversi progetti in Europa già oggi prevedono la partecipazione esclusiva di soggetti pubblici (il Ministero della Giustizia o le Università) ma – tralasciando per adesso i problemi che comunque ciò solleva rispetto alle garanzie di indipendenza della magistratura – sono progetti, come visto, ancora lontani da ciò che oggi intendiamo con intelligenza artificiale e giustizia algoritmica, progetti che invece richiederanno investimenti economici molto maggiori, ad oggi possibili, sembra, quasi esclusivamente per soggetti privati.

Questioni di tal guisa mettono bene in luce la tensione che si avverte oggi tra interessi contrapposti e difficilmente conciliabili che attengono alla costante tensione tra pubblico e privato nel contesto dell'intelligenza artificiale⁵⁰¹: da una parte gli interessi di società che per ragioni di guadagni e di mercato mirano a mantenere il segreto commerciale sui propri algoritmi, dall'altra i diritti processuali di una parte che intende conoscere le informazioni sui cui la decisione si fonda⁵⁰². La tensione, in mancanza di una seria regolazione pubblica, è destinata peraltro ad acuirsi, poiché la tecnologia progredirà ulteriormente e serviranno sempre maggiori

⁵⁰⁰ Si v. ancora H. LIU - C. LIN - Y. CHEN, *Beyond State v Loomis: artificial intelligence, government algorithmization and accountability*, in *International Journal of Law and Information Technology*, 2019, vol. 27, pp. 122-141, spec. pp. 135 ss.; si v. altresì W. GRAVETT, *Jailed by a "black box": the impact of opaque algorithms on the right to a fair trial in the Unites State of America*, cit.

⁵⁰¹ Sui rapporti tra pubblico e privato nell'era digitale si cfr. M. BETZU, *I poteri privati nella società digitale: oligopoli e antitrust*, cit., pp. 739-760, nonché ID, *poteri pubblici e poteri privati nel mondo digitale*, in *Rivista Gruppo di Pisa*, cit., pp. 166-191.

⁵⁰² E. KIESOW - N. MASLEY, *Adjudication of artificial intelligence and automated decision-making cases in Europe and the USA*, in *European Journal of Risk Regulation*, 2023, pp. 1-19, spec. p. 14.

investimenti per progettargliela e commercializzargliela. Investimenti che già oggi sono appannaggio quasi esclusivo di soggetti privati (per esempio le grandi *Big Tech*), gli unici in grado di sostenere spese di tale portata. Oggi gli usi dell'apprendimento automatico sono sempre più diffusi nel settore privato, prendendo parte a sistemi di raccomandazione per i consumatori, a campagne di *marketing*, dell'ottimizzazione delle catene di approvvigionamento, alle auto a guida autonoma e di molto altro ancora⁵⁰³. Il valore dell'apprendimento automatico, come visto, deriva dalla sua capacità di apprendere da solo come individuare schemi utili in enormi insiemi di dati e mettere insieme le informazioni in modo da ottenere previsioni o stime straordinariamente accurate. Dati i vantaggi di questi algoritmi, anche le autorità governative hanno molte opportunità di sfruttarli, pur non avendo le risorse per farlo⁵⁰⁴.

A questo proposito è noto che il costo elevato dei modelli di AI sta allontanando i ricercatori pubblici dal settore, compromettendo lo studio indipendente di questa tecnologia. Mentre aziende come Meta, Google e Microsoft investono miliardi di dollari nell'AI, si sta creando un enorme divario di risorse anche con le università più ricche del Paese. Negli ultimi *report* annuali dell'Università di Stanford si scrive che mentre fino al 2014 la maggior parte dei modelli di *machine learning* sono stati rilasciati dal mondo accademico, da quel momento l'industria tecnologica privata ha preso il sopravvento: nel 2022 ci sono stati creati trentadue significativi modelli di apprendimento automatico da imprese private rispetto ai soli tre prodotti dall'accademia⁵⁰⁵. Nel 2023 invece i modelli progettati dall'industria privata sono stati cinquantuno, mentre quindici quelli creati dal mondo accademico⁵⁰⁶. Ciò perché, come detto, la costruzione di sistemi di IA all'avanguardia richiede sempre più spesso grandi quantità di dati, potenza di calcolo e denaro – risorse che gli operatori del settore dell'industria privata possiedono in quantità maggiore rispetto a organizzazioni pubbliche e alle università.

⁵⁰³ D. M. WEST - J. R. ALLEN, *How Artificial Intelligence Is Transforming the World*, in *Brookings.edu*, 24 aprile 2018.

⁵⁰⁴ C. COGLIANESE - D. LEHR, *Transparency and Algorithmic Governance*, cit., p. 6.

⁵⁰⁵ Il report del 2023 – riferito all'anno 2022 – è visionabile al seguente indirizzo: https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf.

⁵⁰⁶ Si veda ancora il report del 2024 già citato.

Tutto ciò ha dei riflessi evidenti sulla giustizia algoritmica, anch'essa immersa come già visto nelle dinamiche che avvolgono l'intera società. L'esplosione del potere privato sui mezzi tecnologici unito alla sempre crescente importanza di tali mezzi rischia di condurre, in mancanza di cautele giuridiche, a un impiego generalizzato nell'ambito dell'amministrazione della giustizia, la quale dovrebbe ovviamente rimanere estranea alle logiche di mercato che oggi governano l'industria dell'intelligenza artificiale.

3. L'autorità dell'algoritmo nei processi decisionali e l'*automation bias*

Un ultimo nodo che pare essere fondamentale nell'analisi condotta in questa parte del lavoro sull'influenza di algoritmi decisionali sul processo decisionale del giudice affonda le sue radici nelle scienze psicologiche⁵⁰⁷ ma ha, come vedremo, implicazioni giuridiche relevantissime. Nella ricerca svolta sulla giurisprudenza statunitense sembra che il fenomeno, come si vedrà, abbia un radicamento evidente nelle stesse decisioni dei giudici. Esso è noto con il fenomeno del "*technology effect*"⁵⁰⁸ o, nel campo di decisioni algoritmiche, della "*algorithmic authority*"⁵⁰⁹ che produce il cd. *automation bias*.

Il "pregiudizio da automazione" (*automation bias*) si riferisce alla tendenza dell'essere umano a fidarsi quasi ciecamente della tecnologia e delle decisioni automatizzate e rischia di trasformare «*a computer program's suggested answer into a trusted final decision*»⁵¹⁰. Anziché guardare all'algoritmo come un elemento complementare che dovrebbe rimanere oggetto di controllo critico, l'essere umano

⁵⁰⁷ A. TVERSKY - D. KAHNEMAN, *Judgment under Uncertainty: Heuristics and Biases*, in *Science, New Series*, Vol. 185, No. 4157. 1974, pp. 1124-1131.

⁵⁰⁸ Tra i tanti studi si veda, B.B. CLARK - C. ROBERT - S. A. HAMPTON, *The Technology Effect: How Perceptions of Technology Drive Excessive Optimism*, in 31 *Journal of Business and Psychology*, 2016, pp. 87-102.

⁵⁰⁹ C. LUSTIG *et. al.*, *Algorithmic Authority: the Ethics, Politics, and Economics of Algorithms that Interpret, Decide, and Manage*, in *Chi Conference*, 2016, pp. 1057- 1062. Con il medesimo significato si parla anche di "*Philosophy anchoring*", si cfr. M. E. DONOHUE, *A replacement for justitia's scales?: machine learning's role in sentencing*, in *Harvard Journal of Law & Technology*, 32(2), 2019, pp. 657-678.

⁵¹⁰ D.K. CITRON, *Technological Due Process*, cit., 1272.

tende arbitrariamente ad accettare l'*output* algoritmico come infallibile⁵¹¹ o, comunque, come esatto.

Nel contesto della giustizia algoritmica, quindi, se il giudice viene assistito da strumenti automatizzati che producono degli *output* – come la previsione di un indice di recidiva o la previsione dell'esito di una controversia – tenderà a seguire il risultato algoritmico. Pur in ambiti diversi, nelle Corti statunitensi si utilizza da tempo l'espressione “*unringing the bell*” quando si invita la giuria a non considerare un'opinione personale espressa dal procuratore e che non dovrebbe venire considerata in sede di decisione⁵¹². Una volta che la suddetta opinione è stata recepita, diventa difficile “spegnere la campanella” ed eliminare dalla mente del decidente quell'elemento. Come è stato efficacemente sintetizzato da una Corte federale «*if you throw a skunk into the jury box, you can't instruct the jury not to smell it*» (se tiri una puzzola nella stanza della giuria, non riesci a istruire la giuria a non sentirne l'odore)⁵¹³.

Ancora, quanto ad alcune prove particolarmente impattanti ma di difficile comprensione⁵¹⁴, un professore di scienze cognitive parlava di “*Christmas tree phenomenon*” – nel senso che il giudice avrebbe fatto troppo affidamento sulle “*pretty lights*” di tali prove e non avrebbe avuto la forza né gli strumenti per sindacarne il risultato⁵¹⁵.

Ciò accade *a fortiori* nel contesto algoritmico, in cui il risultato del procedimento automatizzato, percepito come oggettivo ed esatto, assume, nella decisione del giudice, un peso *preponderante* e *quasi esclusivo*, e ciò anche per ragioni di

⁵¹¹ A. R. COWGER, JR., *The Threats of Algorithms and AI to Civil Rights, Legal Remedies, and American Jurisprudence: One Nation Under Algorithms*, London, 2020, p. 14.

⁵¹² K. FREEMAN, *Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis*, cit., p. 99.

⁵¹³ *United States v. Rodriguez*, 585 F.2d 1234 (5th Cir. 1978).

⁵¹⁴ Si trattava di immagini cerebrali, *functional magnetic resonance imaging (fMRI)*.

⁵¹⁵ D. MOOBS, *The Implications of Brain Imaging Studies for the Law*, presentato alla conferenza *Law, Mind and Brain interdisciplinary colloquium*, University College London, febbraio 2006. Su questo tema si v. anche N. FEIGENSON, *Brain Imaging and Courtroom Evidence: On the Admissibility and Persuasiveness of fMRI*, in *International Journal of Law in Context*, n. 3, 2006, p. 246.

convenienza pratica, poiché lo sforzo motivazionale sarebbe maggiore nel caso in cui si decidesse di sovvertire il risultato prodotto dall’algoritmo⁵¹⁶.

Una volta, infatti, che si ingaggia l’algoritmo come “prova principe” – caratteristica che possiede per sua stessa natura, come abbiamo spiegato – risulta molto agevole per il giudice assumere la propria decisione motivandola alla luce di quello che l’algoritmo gli ha indicato. L’“*anchoring effect*”⁵¹⁷, insomma, impedisce al giudice di poter vagliare attentamente tutti gli elementi presenti e finisce per attrarre la decisione finale⁵¹⁸. Nel caso specifico dei *risk assessment tools* si è scritto che un giudice che ha davanti agli occhi uno strumento algoritmico di previsione del rischio di recidiva «potrebbe essere portato a irrogare una pena maggiore senza aver neanche la minima consapevolezza del ruolo avuto dall’“*anchoring*” nella decisione medesima»⁵¹⁹.

L’effetto ancoraggio (o euristica dell’ancoraggio), per vero, è stata a lungo studiata prima dell’avvento di applicazioni algoritmiche e ha messo in luce come i giudici, mossi dal desiderio di semplificare la realtà, tendono a farsi un’idea del caso sottoposto al proprio esame fin dall’inizio e diventa poi assai difficile superare tale pregiudizio (il cd. *bias di conferma*)⁵²⁰. Se tale effetto è stato riconosciuto come fondamentale da diversi studi nelle decisioni dei giudici⁵²¹, è evidente che

⁵¹⁶ In Italia Simoncini e Suweis parlano di “forza di argomentazione pratica dell’algoritmo”. Si cfr. A. SIMONCINI - S. SUWEIS, *Il cambio di paradigma nell’intelligenza artificiale e il suo impatto sul diritto costituzionale*, cit., p. 100. Già prima Simoncini parlava di «travolgente forza pratica dell’algoritmo», si cfr., A. SIMONCINI, *L’algoritmo incostituzionale; intelligenza artificiale e il futuro delle libertà*, in *Rivista di BioDiritto*, cit., p. 81.

⁵¹⁷ Questa è l’espressione utilizzata in H. LIU - C. LIN - Y. CHEN, *Beyond State v Loomis: artificial intelligence, government algorithmization and accountability*, cit., p. 130, in cui gli autori affermano: «[t]hat is, the Court failed to consider the psychological ‘anchoring effect’ for courts using scientific and technological tools, as numerous studies have demonstrated how judges (and human individuals) are submissive to computer-generated numbers and results that may further frame and condition the view of judges».

⁵¹⁸ In una indagine condotta in Virginia sull’utilizzo di uno strumento di valutazione del rischio è emerso che la grande maggioranza dei giudici era propenso ad attenersi al risultato prodotto dall’algoritmo, si cfr. B. L. GARRETT - J. MONAHAN, *Judging Risk*, cit., p. 467.

⁵¹⁹ L. MALDONATO, *Algoritmi predittivi e discrezionalità del giudice: una nuova sfida per la giustizia penale* in *Diritto penale contemporaneo*, 2/2019, p. 410.

⁵²⁰ J. NIEVA-FENOLL, *Intelligenza artificiale e processo*, trad.it di P. COMOGLIO, Giappichelli, Torino, 2019, pp. 38 ss.

⁵²¹ Si cfr. *ex multis*, M. J. SAKS - R. F. KIDD, *Human Information Processing and Adjudication: Trial by Heuristics*, in *Law & Society Review*, Vol. 15, No. 1, 1980, pp. 123-160; E. GARRIDO MARTIN - C. HERRERO ALONSO, *Influence of the prosecutor’s Plea on the Judge’s Sentencing in Sexual Crimes: Hypothesis of the Theory of Anchoring by Tversky and Kahneman*, 1997.

l'algoritmo, per le sue caratteristiche di apparenza di oggettività, oltre che di incomprendibilità, una volta portato all'attenzione del giudice senza le garanzie di cui discuteremo nei prossimi capitoli, tenderà a divenire la base decisiva per la decisione finale.

A questo riguardo, non è un caso che in tutti i tentativi di regolazione dell'impiego dell'IA, in generale e nello specifico nella giustizia algoritmica, sono presenti previsioni riguardanti la "supervisione" umana, la "sorveglianza" umana, il cd. "*human in the loop*" (si v. *supra*, Parte II). Da ultimo – e il tema verrà approfondito più avanti nel corso del lavoro quando si tratterà delle soluzioni per garantire il giusto processo nel contesto della giustizia algoritmica (si v. *infra*, cap. IX) – anche il legislatore europeo ha riconosciuto i rischi dell'*automation bias*, che nella versione italiana del Regolamento viene definito «distorsione da automazione». Al § 4 lett. b) dell'art. 14 del nuovo Regolamento, rubricato "Sorveglianza umana", si prevede che colui che sarà deputato a tale sorveglianza debba «restare consapevole della possibile tendenza a fare automaticamente affidamento o a fare eccessivo affidamento sull'output prodotto da un sistema di IA ad alto rischio («distorsione dell'automazione»), in particolare in relazione ai sistemi di IA ad alto rischio utilizzati per fornire informazioni o raccomandazioni per le decisioni che devono essere prese da persone fisiche».

In definitiva, come si vedrà nei successivi capitoli, l'influenza dell'*automation bias* oggi deteriora l'esercizio della discrezionalità giudiziaria in quanto i giudici attribuiscono alla tecnologia e ai risultati algoritmici un peso maggiore rispetto ad altri fattori, indipendentemente dal fatto che lo facciano consapevolmente o meno⁵²².

Breve conclusione: l'intreccio dei nodi

Questo breve capitolo è servito per indicare le maggiori problematiche che oggi sembrano dover essere considerate quando si discute di giustizia algoritmica. Preme

⁵²² A. GREEN, *Humans May Put Too Much Trust in Robots, Study Finds*, in *Mental Floss*, marzo 2016.

peraltro sottolineare, in questo paragrafo conclusivo, che tali problematiche non rappresentano blocchi isolati, ma anzi si intrecciano tra loro e si autoalimentano, venendo così a provocare un'unica, forte, influenza sul processo decisionale del giudice, come vedremo nei successivi capitoli.

L'intreccio è facilmente spiegabile.

Più vengono utilizzate tecniche sofisticate di *machine learning*, fondate su quantità sempre più grandi di dati, più i sistemi sconteranno defezioni per quanto concerne la loro trasparenza.

La complessità dei sistemi e l'opacità dell'algoritmo impiegato nella giustizia comporta, di conseguenza, l'impossibilità del giudice, che ad oggi difetta degli strumenti formativi necessari, di comprendere ed eventualmente sindacare il risultato prodotto dallo stesso (vedremo che tale atteggiamento è frequente nei casi giurisprudenziali statunitensi).

La mancanza di comprensione unita alla apparenza di neutralità ed "esattezza" dell'algoritmo comporta l'adesione automatica del giudice al risultato dello stesso, producendo quello che abbiamo descritto come *automation bias*.

Lo strettissimo intreccio tra i diversi nodi problematici che abbiamo tentato di delineare in questo capitolo generano dubbi non irrilevanti rispetto alla compatibilità della giustizia algoritmica con le garanzie costituzionali del giusto processo. Dopo un rapido inquadramento di tali garanzie nel *due process* statunitense e nel giusto processo costituzionale europeo, tali dubbi verranno costituzionalizzati partendo dall'analisi della giurisprudenza americana, che ha già dovuto fare i conti con una forma di giustizia algoritmica.

IL *DUE PROCESS OF LAW* STATUNITENSE E IL *FAIR TRIAL* NELLA PROSPETTIVA EUROPEA
MULTILIVELLO

SOMMARIO: Premessa - 1. Cenni generali sul *due process of law* nella Costituzione degli Stati Uniti - 1.1. Il dibattito giurisprudenziale tra *abuse of discretion* e violazione del *due process of law* - 1.2. *L'abuse of discretion standard* - 1.3. Le principali garanzie del *due process of law* - 2. Il *fair trial* nella prospettiva multilivello: le disposizioni rilevanti

Premessa

Se si ragiona di processo *giusto* non risulta facile affibbiargli una definizione astratta e univoca: ciò forse per il motivo che il giusto processo è un principio giuridico talmente radicato nella nostra tradizione giuridica che estrarne un solo significato, quello *originale*, sembra particolarmente complesso (oltre che forse non utile). Già la dichiarazione universale dei diritti umani, adottata nel 1948, infatti, prevedeva che «*Everyone is entitled in full equality to a fair and public hearing by an independent and impartial tribunal, in the determination of his rights and obligations and of any criminal charge against him*» (art. 10).

Oltreoceano tale idea è stata espressa in una *concurring opinion* del giudice Frankfurter nel caso *Joint Anti-Fascist Refugee Committee v. McGrath*, deciso dalla Corte Suprema degli Stati Uniti nel 1951: secondo il Giudice: «*'due process', unlike some legal rules, is not a technical conception with a fixed content unrelated to time, place and circumstances. Expressing as it does in its ultimate analysis respect enforced by law for that feeling of just treatment which has been evolved through centuries of Anglo-American constitutional history and civilization, 'due process' cannot be imprisoned within the treacherous limits of any formula. Representing a profound attitude of fairness between man and man, and more particularly between the individual and government, 'due process' is compounded of history, reason, the*

*past course of decisions, and stout confidence in the strength of the democratic faith which we profess»*⁵²³.

Anche nella tradizione europea si riscontra la difficoltà di definire il giusto processo in astratto: la Corte europea dei diritti, coerentemente con il suo approccio casistico, non definisce che cosa sia il *fair trial*, ma piuttosto costruisce dal caso sottoposto al suo esame le garanzie che devono sussistere perché il diritto sia rispettato.

La Corte costituzionale, dal canto suo, ha tentato di fornire una definizione di giusto processo, ancorché non in astratto ma anch'essa tentando di ricostruirne la fisionomia attraverso le garanzie costituzionali. In una sentenza la Corte ha affermato che il giusto processo è una «formula in cui si compendiano i principi che la Costituzione detta in ordine tanto ai caratteri della giurisdizione, sotto il profilo soggettivo e oggettivo, quanto ai diritti di azione e difesa in giudizio»⁵²⁴. Coerentemente, in altra sentenza, la Corte ha inteso il “giusto processo” come «espressione necessaria del diritto ad una tutela giurisdizionale mediante azione (art. 24 della Costituzione) avanti ad un giudice con le garanzie proprie della giurisdizione, cioè con la connaturale imparzialità, senza la quale non avrebbe significato né la soggezione dei giudici solo alla legge (art. 101 della Costituzione), né la stessa autonomia ed indipendenza della magistratura (art. 104, primo comma, della Costituzione)»⁵²⁵.

Dunque nel giusto processo si condensano le garanzie sia legate alla funzione giurisdizionale sia quelle legate alle parti nel processo: secondo questa impostazione, dopo aver delineato nel presente capitolo i caratteri generali del *due process of law* statunitense e del *fair trial* europeo, non cercheremo di trovare una risposta alla domanda se la giustizia algoritmica, per come oggi la si conosce, sia compatibile con il processo “giusto” *tout court*, quanto piuttosto verificare la compatibilità della giustizia algoritmica rispetto alla singole garanzie che di volta in volta vengono in rilievo.

⁵²³ Corte Suprema degli Stati Uniti, *Joint Anti-Fascist Refugee Committee v. McGrath*, 341 U.S. 123 (1951).

⁵²⁴ Corte cost., sent. n. 131 del 1996.

⁵²⁵ Corte cost., sent. n. 287 del 1999.

Un' ultima precisazione si rende necessaria: mentre negli Stati Uniti già sono sorti dei casi in cui la giurisprudenza è stata chiamata a sindacare la compatibilità della giustizia algoritmica con il *due process*, in Europa – in cui come visto non esiste un'esperienza strutturata di impiego di algoritmi da parte dei giudici nel processo – non è ancora presente casistica rilevante sul tema. Ciò comporta che, nell'ottica comparativa che abbiamo segnalato all'inizio del lavoro, si prenderanno dapprima in considerazione le garanzie del *due process of law* e come esse sono state interpretate dai giudici nordamericani di fronte all'impiego di algoritmi; successivamente, forti dell'insegnamento e delle coordinate di tale giurisprudenza, proveremo a indentificare quali sono le garanzie del *fair trial* che appaiono più rilevanti nel contesto algoritmico.

Come facilmente si intuirà, infatti, i numerosi nodi giuridici che sono emersi nel dibattito statunitense con riguardo all'IA nel giudiziario hanno una portata che difficilmente può essere confinata al contesto locale, seppure molto esteso come quello statunitense. Anzi, le interpretazioni dei giudici americani e la letteratura giuridica che le ha accompagnate, essendo oggi il primo e più importante contenitore da cui attingere per trarre dei punti fermi su cui impostare il discorso, devono essere un insegnamento per il futuro inserimento diffuso dell'IA nel panorama giudiziario europeo.

Peraltro, benché non esistano ancora decisioni delle Corti europee specificamente riguardanti la giustizia algoritmica, qualche indicazione emerge già da alcune affermazioni svolte in sentenze che riguardavano in generale le nuove tecnologie. Per esempio la Corte europea dei diritti in una sentenza ha affermato che «*any State claiming a pioneer role in the development of new technologies bears special responsibility for striking the right balance in this regard*⁵²⁶». Quando cioè la giustizia algoritmica diventerà una realtà nel continente europeo, gli Stati avranno una «speciale responsabilità» nel garantire il giusto equilibrio tra i diritti in gioco. La Corte di giustizia dell'Unione europea, dal canto suo, riguardo all'impiego dell'intelligenza artificiale e la sua compatibilità con il giusto processo,

⁵²⁶ Corte Edu, *S. and Marper v. the United Kingdom* [GC], 2008, § 112).

ha specificato che «tenuto conto dell'opacità che caratterizza il funzionamento delle tecnologie di intelligenza artificiale, può risultare impossibile comprendere la ragione per la quale un dato programma sia arrivato ad un riscontro positivo. In tali circostanze, l'uso di siffatte tecnologie potrebbe privare gli interessati anche del loro diritto a un ricorso giurisdizionale effettivo sancito dall'articolo 47 della Carta»⁵²⁷.

Come vedremo, i profili già oggi individuati dalle Corti quanto all'impiego di nuove tecnologie e dell'intelligenza artificiale sono decisivi anche nel più ristretto contesto della giustizia algoritmica.

1. Cenni generali sul *due process of law* nella Costituzione degli Stati Uniti

La giurisprudenza statunitense sull'impiego degli strumenti algoritmici di previsione del rischio di recidiva ha messo in luce le garanzie rilevanti del *due process of law* che vengono in rilievo e ha anche evidenziato il mutamento nell'esercizio della discrezionalità giudiziaria di fronte ai risultati algoritmici (*abuse of discretion*, strettamente legato al giusto processo).

Mentre le singole garanzie verranno di volta in volta prese in considerazione quando si vaglieranno specificamente le problematiche del giusto processo alla luce del nuovo contesto di giustizia algoritmica, nei successivi paragrafi si darà rapidamente conto dei caratteri generali sia del *due process* sia dell'*abuse of discretion*, entrambi rilevanti nell'impianto del presente lavoro.

1.1. Il dibattito giurisprudenziale tra *abuse of discretion* e violazione del *due process of law*

Se il celebre caso *Loomis* (si v. *infra*, § 3) ha avuto il merito di portare il dibattito su un piano diverso – le possibili violazioni del *due process*, come si vedrà – già la giurisprudenza precedente ha vagliato la legittimità dei *tools* predittivi non già alla

⁵²⁷ CGUE, *Ligue des droits humains v. Conseil des ministres*, C-817/19, 21 giugno 2022.

luce del principio costituzionale del giusto processo quanto dell'abuso di discrezionalità (*abuse of discretion*) da parte del giudice che determina la pena (anche) sulla base del risultato prodotto dall'algoritmo predittivo.

Sebbene nel diritto statunitense *l'abuse of discretion* non rientri formalmente in uno dei corollari del *due process* – essendo uno *standard of review* utilizzato dalle Corti superiori per sindacare decisioni considerate arbitrarie delle Corti di prime cure – le due prospettive verranno analizzate parallelamente, sia perché gli argomenti che fondano la lamentata illegittimità dell'impiego degli algoritmi predittivi sono spesso i medesimi, sia perché, come si vedrà in alcune sentenze più recenti, *l'abuse of discretion* viene invocato insieme alle violazioni del *due process*, costituendone causa e conseguenza. Così, per esempio, un eventuale esercizio abusivo della discrezionalità potrebbe comportare una violazione del giusto processo qualora abbia impedito alla parte di esercitare il proprio diritto di difesa; così anche un'eventuale violazione del *due process* – quale per esempio la arbitraria valutazione di una prova – può condurre a un abuso di discrezionalità.

La stretta correlazione tra *abuse of discretion* e violazioni del giusto processo emerge limpidamente da alcune sentenze che verranno prese in considerazione del corso dell'analisi giurisprudenziale: in un argomento sollevato in *Loomis*, la mancata *disclosure* dell'algoritmo costituisce tanto un abuso di discrezionalità del giudice quanto una violazione del giusto processo a causa della segretezza delle informazioni sulla cui base la pena viene determinata⁵²⁸. In altre sentenze i giudici, spiegando che l'imputato lamentava la violazione del giusto processo e, in alternativa, l'esercizio abusivo della discrezionalità per aver la corte considerato l'algoritmo senza essere consapevole delle sue limitazioni, precisano che «*the alternative argument* (quello rispetto all'*abuse of discretion*) *is in essence a due process argument*»⁵²⁹.

⁵²⁸ Si v. *infra* nel caso *Loomis* al § 49.

⁵²⁹ *State v. Guise* (v. *infra* § 3).

1.2. L'abuse of discretion standard

Ai nostri fini non interessa indagare a fondo la natura giuridica dell'*abuse of discretion* come *standard of review* delle decisioni delle corti circondariali. È sufficiente qui dare conto di che cosa si intenda con tale termine e come questo parametro è costantemente impiegato per sostenere l'illegittimità dell'impiego dei *risk assessment tools* nel *sentencing*.

Nel diritto nordamericano – e non ci sono differenze rilevanti tra il diritto federale e il diritto delle singole giurisdizioni statali – le *trial courts* godono di ampia discrezionalità (*great discretion*).

Tale ampia discrezionalità incontra però il limite dell'abuso di essa, che si concreta in forme diverse, dall'abuso di discrezionalità per l'applicazione di una legge erronea, l'abuso di discrezionalità per essersi il giudice basato su una ricostruzione dei fatti palesemente errata, fino all'abuso di discrezionalità per avere il giudice determinato la condanna a fronte di un'evidente assenza di prove.

Per usare le parole delle Corti federali, l'abuso di discrezionalità si verifica quando la sentenza è «*unreasonable given the totality of the circumstances*»⁵³⁰. E ancora «*A district court abuses its sentencing discretion only if the sentence exceeded the bounds of permissible choice*»⁵³¹ o quando essa è «*arbitrary, capricious, whimsical, or manifestly unreasonable*»⁵³².

Per quanto qui interessa, nel processo penale l'esercizio abusivo della discrezionalità viene invocato dinanzi alle Corti d'Appello o alle Corti Supreme, come vedremo, allorché il giudice di prime cure abbia fondato la decisione sulla determinazione della pena sulla base di un algoritmo predittivo o, al contrario, lo abbia trascurato. In alcuni casi, si lamenta l'utilizzo del risultato algoritmico quale circostanza aggravante, in altri l'esercizio abusivo della discrezionalità viene fatto discendere dall'appiattimento della decisione del giudice al risultato della valutazione del rischio che farebbe venire meno la responsabilità del giudice ad

⁵³⁰ *United States v. Caiba- Antele*, 705 F.3d 1162, 1165 (10th Cir. 2012).

⁵³¹ *United States v. McComb*, 519 F.3d 1049, 1053 (10th Cir. 2007).

⁵³² *United States v. DeRusse*, 859 F.3d 1232, 1236 (10th Cir. 2017).

applicare una sentenza individualizzata; in altri ancora l'*abuse of discretion* dipenderebbe dalla mera considerazione del risultato algoritmico.

1.3. Le principali garanzie del due process of law

Come noto, la *due process clause* è contenuta, con la stessa formulazione, nel V e nel XIV emendamento della Costituzione americana. Mentre il V emendamento, riferito al Governo federale, afferma che «*no person shall...be deprived of life, liberty, or property, without due process of law*», il XIV, ratificato nel 1868, si rivolge ai singoli Stati, prevedendo anch'esso che nessuno Stato possa «*deprive any person of life, liberty, or property, without due process of law*».

Vale la pena di sottolineare che l'interpretazione sia di cosa si intenda per «vita» «libertà» e «proprietà» sia di quando il processo può considerarsi «giusto» è cambiata nel corso dei quasi due secoli e mezzo dall'approvazione della Costituzione americana. Inoltre, la loro interpretazione è influenzata dalla legislazione dei singoli stati e dall'evoluzione della giurisprudenza sia federale sia nazionale.

Allo scopo di fornire solo alcune coordinate di natura generale, è utile chiarire che se un interesse non rientra nel significato della espressione «vita, libertà o proprietà», la *due process clause* non impone alcuna procedura particolare per la sua privazione; inoltre, la clausola si applica soltanto alle azioni che hanno natura pubblica, essendo esclusi i rapporti tra privati.

Fatte queste premesse, se non è possibile fornire una lista esaustiva di corollari del *due process*, in un articolo del 1975 che ha avuto una grossa influenza anche sulla giurisprudenza, il giudice Henry J. Friendly ha tentato di fornire un elenco di procedure che dovrebbero essere garantite affinché il processo possa ritenersi *due*⁵³³.

Secondo il giudice Friendly, gli elementi necessari per una *fair hearing*, e quindi del giusto processo, sono: 1) un tribunale imparziale; 2) la notifica dell'azione

⁵³³ H.J. FRIENDLY, *Some Kind of Hearing*, in *123 University of Pennsylvania Law Review*, 1975, pp. 1267 ss.

proposta e delle motivazioni addotte a sostegno di essa; 3) l'opportunità di presentare ragioni per le quali l'azione proposta non dovrebbe essere intrapresa; 4) il diritto di chiamare testimoni; 5) il diritto di conoscere le prove a proprio carico; 6) il diritto di avere decisioni su prove conosciute dalla parte, 7) il diritto all'assistenza tecnica; 8) la verbalizzazione del processo; 9) la motivazione; 10) la presenza del pubblico; 11) la *Judicial Review*⁵³⁴. In ogni caso, nel procedimento penale – e ciò ci servirà per vagliare la legittimità della giurisprudenza della Corti statunitensi – alla base del *due process of law* giace il principio per cui le condanne devono essere fondate su elementi di prova che conducano a decidere per la colpevolezza *oltre ogni ragionevole dubbio*⁵³⁵.

Come si vedrà nel corso dell'analisi della giurisprudenza, molti dei corollari qui indicati verranno in rilievo nel contesto di decisioni giudiziarie assunte tramite impiego di strumenti algoritmici di valutazione del rischio di recidiva.

2. Il *fair trial* nella prospettiva multilivello: le disposizioni rilevanti

Come segnalato nel capitolo introduttivo di questo lavoro, la prospettiva delle garanzie costituzionali del giusto processo di fronte alla giustizia algoritmica seguirà l'impostazione multilivello di tutela dei diritti.

Ora, mentre le singole garanzie e le relative interpretazioni giurisprudenziali, come abbiamo già specificato per il *due process*, verranno prese singolarmente in considerazione nel prossimo capitolo, in questo breve paragrafo segnaleremo solo le norme e la prospettiva da cui muoveremo.

Il giusto processo è riconosciuto sia dalla Carte e Convenzioni sovranazionali sia dalla Costituzione italiana. La Convenzione europea dei diritti dell'uomo, all'art. 6 stabilisce che «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

⁵³⁴ *Ibidem.* pp. 1279 ss.

⁵³⁵ C. BASSU - M. BETZU - F. CLEMENTI - G. COINU, *Diritto costituzionale degli Stati Uniti d'America. Una introduzione*, cit., pp. 112-113.

sulle controversie sui suoi diritti e doveri di carattere civile o sulla fondatezza di ogni accusa penale formulata nei suoi confronti».

Nel diritto dell'Unione europea la disposizione rilevante per il nostro discorso è l'art. 47 della Carta dei diritti fondamentali, il quale prevede ai primi due paragrafi, che «Ogni individuo 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 (§ 1) e «Ogni individuo 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 individuo ha la facoltà di farsi consigliare, difendere e rappresentare» (§ 2)⁵³⁶.

Chiarisce la relazione tra le norme della Convenzione e le norme della Carta l'art. 52 § 3 di quest'ultima secondo cui «[l]addove la presente Carta contenga diritti corrispondenti a quelli garantiti dalla convenzione europea per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali, il significato e la portata degli stessi sono uguali a quelli conferiti dalla suddetta convenzione. La presente disposizione non preclude che il diritto dell'Unione conceda una protezione più estesa». Quindi mentre l'art. 47 § 1 corrisponde all'art. 13 CEDU⁵³⁷, il § 2 corrisponde ai diritti garantiti dall'art. 6 § 1 CEDU, senza le limitazioni di quest'ultimo legate alle controversie di carattere civile⁵³⁸.

Nella Costituzione italiana le norme per noi rilevanti sono l'art. 24 e l'art. 111 Cost., oltre che le disposizioni, di cui si tratterà specificamente *infra*, sulla

⁵³⁶ Sul versante dell'obbligo degli Stati membri di garantire il giusto processo rileva altresì l'art. 19 § 2 del Trattato sul funzionamento dell'Unione Europea secondo cui «Gli Stati membri stabiliscono i rimedi giurisdizionali necessari per assicurare una tutela giurisdizionale effettiva nei settori disciplinati dal diritto dell'Unione».

⁵³⁷ Nella prospettiva del giusto processo rientrerebbe anche l'art. 13 CEDU secondo cui «Ogni persona i cui diritti e le cui libertà riconosciuti nella presente Convenzione siano stati violati, ha diritto a un ricorso effettivo davanti a un'istanza nazionale, anche quando la violazione sia stata commessa da persone che agiscono nell'esercizio delle loro funzioni ufficiali». Tuttavia, ai nostri fini è sufficiente limitare l'indagine al solo art. 6 CEDU per due ragioni, sia perché l'art 6 § 1 è norma speciale rispetto all'art. 13 e quindi quest'ultimo viene assorbito dai requisiti del giusto processo (*ex multis*, Corte Edu, *Grzęda v. Poland* [GC], 2022, § 352; sia perché, come regola generale, l'art. 13 CEDU non si applica quando la violazione dei diritti avviene durante il processo, che è la prospettiva da cui muove questo lavoro (*ex multis*, *Menesheva v. Russia*, 2006, § 105).

⁵³⁸ Si cfr. B. RANDAZZO, *Access to Justice in a Multilevel Constitutional System, Protecting Human Rights*, Giappichelli, 2024, p. 28 e CJEU, *Trade Agency Ltd v. Seramico Investments Ltd*, C-6919/10, § 52.

indipendenza della magistratura, in particolare gli artt. 101, secondo comma, e 104, primo comma, Cost.

I primi due commi dell'art. 24 prevedono che «Tutti possono agire in giudizio per la tutela dei propri diritti e interessi legittimi (primo comma)» e «La difesa è diritto inviolabile in ogni stato e grado del procedimento» (secondo comma).

Come noto, il legislatore costituzionale italiano, sulla scorta non solo del testo della Convenzione ma altresì della giurisprudenza della Corte europea dei diritti⁵³⁹, ha successivamente introdotto ai primi commi dell'art. 111 Cost il principio secondo cui «La giurisdizione si attua mediante il giusto processo regolato dalla legge» (comma 1) e «Ogni processo si svolge nel contraddittorio tra le parti, in condizioni di parità, davanti a giudice terzo e imparziale. La legge ne assicura la ragionevole durata».

Sebbene la giurisprudenza costituzionale abbia più volte specificato che i primi due commi dell'art. 111 non hanno una vera portata innovativa, ma piuttosto ricognitiva, nel senso della mera riproduzione o esplicitazione di contenuti normativi già presenti a livello costituzionale (in particolare nei diritti alla tutela giurisdizionale e alla difesa sanciti dall'art. 24 Cost)⁵⁴⁰, la cristallizzazione di tali principi nella Costituzione, secondo la dottrina, non è comunque priva di conseguenze, soprattutto in riferimento alla collocazione dei diritti di difesa e di tutela in giudizio nell'alveo dei principi supremi dell'ordinamento⁵⁴¹.

⁵³⁹ Lo svolgimento del processo «nel contraddittorio delle parti» e «in condizioni di parità» ai sensi dell'art. 111 Cost. non è previsto invece dal testo dell'art. 6 CEDU ma è stato riconosciuto, come visto, dalla giurisprudenza della Corte EDU.

⁵⁴⁰ Così Corte cost., sent. n. 59 del 2009 «la citata legge costituzionale n. 2 del 1999 - nell'introdurre nel corpo dell'art. 111 Cost. un nuovo secondo comma, per il quale «Ogni processo si svolge nel contraddittorio tra le parti, in condizioni di parità [...]» - si è limitata, sul punto, ad esplicitare nel testo della Costituzione principi che la giurisprudenza di questa Corte aveva già tratto dagli artt. 3 e 24 Cost. e che aveva posto a fondamento del «giusto processo» [...]. Pertanto, con riferimento ai suddetti principi, l'art. 111, secondo comma, Cost. non ha carattere innovativo» (punto 6.4.1 del *Considerato in diritto*. Si cfr. anche, *ex multis*, sentenze n. 241 del 1999; n. 290 del 1998; n. 432 del 1995).

⁵⁴¹ Per una ricognizione della dottrina sul tema all'indomani della riforma costituzionale si cfr., M. CECCHETTI, *La riforma dell'art. 111 Cost.: tra fonti preesistenti, modifica della Costituzione, diritto intertemporale ed esigenze di adeguamento della legislazione ordinaria*, in U. DE SIERVO (a cura di), *Osservatorio sulle fonti*, Torino, Giappichelli, 2001, pp. 35-66. Secondo l'A. «la stretta connessione con i principi supremi che caratterizza la riforma attuata con la l. cost. n. 2/1999 potrebbe anche rendere plausibile ritenere che le nuove norme ricavabili dall'art. 111 siano, da questo punto di vista, dotate di una forza di resistenza passiva di carattere «supercostituzionale», nel

Ciò precisato, come autorevolmente sostenuto⁵⁴² è la giurisprudenza della Corte europea dei diritti la «*most influencing*» nel processo di integrazione multilivello e, di conseguenza, nella definizione dei requisiti del giusto processo. Ciò in quanto la Corte europea «*deals with applications alleging violations of rights of as many as 47 (now 46) different legal systems, adopting a very broad interpretation of the Convention in order to update its guarantees to present-day conditions*». Inoltre, sul piano nazionale, la giurisprudenza della Corte europea dei diritti conforma anche l'interpretazione delle disposizioni nazionali *ex art. 117*, primo comma Cost, come interpretato dalla Corte costituzionale nelle note “sentenze gemelle”⁵⁴³.

Per questa ragione, per indagare la compatibilità della giustizia algoritmica con le garanzie del giusto processo prenderemo in considerazione principalmente la giurisprudenza della Corte EDU, segnalando quando necessario per una maggiore definizione del ragionamento anche la giurisprudenza della Corte di Giustizia dell'Unione europea e della Corte costituzionale.

Un'ultima notazione è necessaria: come noto la giurisprudenza della Corte europea ha una natura casistica, e ogni principio andrebbe indagato rispetto al caso sottoposto al suo esame. Ciò precisato, ai nostri fini – considerati alcuni principi ormai “consolidati” all'interno dell'interpretazione della Corte – riteniamo possibile astrarre dal caso la giurisprudenza rilevante e adattarla al contesto qui in esame.

senso che la loro eventuale rimozione o l'incisione del loro nucleo essenziale di valore, anche ad opera di una successiva legge approvata con il procedimento di revisione, comporterebbe inevitabilmente la violazione dei principi fondamentali dell'ordinamento costituzionale».

⁵⁴² Si cfr., B. RANDAZZO, *Access to Justice in a Multilevel Constitutional System, Protecting Human Rights*, cit., p. 82.

⁵⁴³ Corte cost., sentt. n. 348 e 349 del 2007 in cui la Corte ha risolto in termini innovativi il rapporto tra legge ordinaria, Costituzione e i vincoli derivanti dai trattati internazionali, in particolare la Convenzione europea dei diritti. Sul tema si v. B. RANDAZZO, *Interpretazione delle sentenze della Corte europea dei diritti ai fini dell'esecuzione (giudiziaria) e interpretazione della sua giurisprudenza ai fini dell'applicazione della CEDU*, in *Rivista AIC*, 2/2015. Si v. anche, ID, *Il 'riaccostamento' del giudizio costituzionale nella prospettiva di un sistema integrato di giustizia costituzionale*, in *Federalismi*, 3, 2021, pp. 144-60.

CAPITOLO VIII

L'INCIDENZA DELLA GIUSTIZIA ALGORITMICA SULLA DECISIONE GIUDIZIARIA E LE RICADUTE SULLE GARANZIE DEL GIUSTO PROCESSO

SOMMARIO: 1. Premessa: l'impatto degli algoritmi sulla discrezionalità come prodromo alle violazioni del giusto processo - SEZIONE I. LA GIURISPRUDENZA AMERICANA SULL'IMPIEGO DEI *RISK ASSESSMENT TOOLS* NEL PROCEDIMENTO PENALE - 2. L'insegnamento giurisprudenza statunitense: l'impatto dei *risk assessment tools* sulla discrezionalità giudiziaria - 2.1. I primi giudizi sull'impiego dei *tools* algoritmici nel *sentencing*: la deresponsabilizzazione del giudice - 2.2. (segue) I primi argomenti a favore dell'impiego dei *tools* nel *sentencing* - 3. (segue) L'insegnamento della giurisprudenza statunitense: i *risk assessment tools* alla prova del *due process of law*. 3.1. (segue) Il caso *Loomis* - 3.2. (segue) La giurisprudenza successiva a *Loomis* - SEZIONE II. I DIVERSI LIVELLI DI INFLUENZA DELL'ALGORITMO SULLA DECISIONE DEL GIUDICE - 4. L'autorità algoritmica nel procedimento decisionale del giudice - 4.1. Autorità algoritmica nel giudizio penale nordamericano: la pretesa non decisività dell'algoritmo come garanzia di una sentenza individualizzata - 4.2. L'influenza dell'autorità algoritmica nelle altre ipotesi di giustizia algoritmica - 5. Opacità dell'algoritmo ed esternalizzazione della funzione giudicante - SEZIONE III. LE RICADUTE SULLE GARANZIE DEL GIUSTO PROCESSO - 6. La giustizia algoritmica alla prova dell'indipendenza e sull'imparzialità del giudice - 7. La *fairness* del processo giurisdizionale - 7.1. La giustizia algoritmica alla prova del contraddittorio e della parità delle armi - 7.2. La giustizia algoritmica alla prova della motivazione della decisione - 8. Nota conclusiva: la giustizia algoritmica alla prova del diritto di difesa e a una tutela giurisdizionale effettiva

1. Premessa: l'impatto degli algoritmi sulla discrezionalità come prodromo alle violazioni del giusto processo

Quando si discute di giustizia algoritmica, si citano spesso, negli Stati Uniti ma anche in Europa, le parole pronunciate nel 2017 dal *Chief Justice* della Corte suprema degli Stati Uniti, John G. Roberts Jr., il quale, alla domanda se riuscisse a prevedere un giorno in cui macchine intelligenti guidate dall'intelligenza artificiale (IA) sarebbero state in grado di assistere i tribunali nell'accertamento dei fatti o,

persino, nel «*judicial decision making*», la sua risposta apparve particolarmente netta: «*[i]t's a day that's here, and it's putting a significant strain on how the judiciary goes about doing things*»⁵⁴⁴.

Nonostante siano parole, in effetti, particolarmente efficaci, esse hanno teso la mano a ricerche e osservazioni che hanno limitato il proprio campo di indagine alle possibili violazioni delle garanzie del giusto processo dalla prospettiva delle parti in causa (in particolare, a titolo di esempio, il diritto di difesa, il diritto di accesso al giudice, il diritto a una tutela giurisdizionale effettiva)⁵⁴⁵.

I lodevoli sforzi della dottrina europea hanno lasciato in ombra, tuttavia, ciò su cui invero le parole del Giudice Roberts parrebbero voler mettere in guardia: la giustizia algoritmica sta mettendo a dura prova (*significant strain*) il modo stesso di esercitare la funzione giurisdizionale, le modalità attraverso cui il giudice opera, gli strumenti giuridici che accompagnano la decisione giudiziaria.

La «*tremendous revolution*» – come definita dal Giudice – non conosce quindi solo un mutamento di paradigma dal punto di vista di coloro che vengono raggiunti da decisioni in tutto o in parte assunte da sistemi algoritmici ma, al contrario, determina un significativo cambio di prospettiva dal lato di chi tali sistemi è chiamato a utilizzare, nella specie il giudice e il suo potere discrezionale.

E anzi, come si vedrà, il confine tra violazione e non violazione del giusto processo nel contesto della giustizia algoritmica si muove in dipendenza di come muta l'esercizio del potere discrezionale di fronte agli algoritmi impiegati nel processo. In altre parole, una volta che il processo diventa la sede dell'utilizzo di strumenti algoritmici, se mancano le cautele, tecniche e giuridiche, necessarie (di cui discuteremo nell'ultima parte del lavoro) si produce un impatto sulla decisione discrezionale e interpretativa del giudice che ha delle ricadute notevoli sul rispetto delle garanzie costituzionali.

Questo capitolo intende quindi indagare come un sistema automatizzato (come lo sono i *tools* impiegati nel processo americano, gli strumenti di giustizia

⁵⁴⁴ L'intervista integrale è disponibile su *YouTube* al seguente indirizzo: <https://www.youtube.com/watch?v=TuZEKIRgDEg>.

⁵⁴⁵ Su cui ci si soffermerà nel capitolo successivo.

predittiva, o ancora strumenti di IA generativa come Chat GPT), possa avere un'influenza su tale scelta discrezionale e come tale influenza incida sul giusto processo.

Per discutere giuridicamente dell'impatto degli algoritmi sulla discrezionalità giudiziaria si porrà inizialmente l'attenzione sulla giurisdizione statale statunitense che, come detto, da anni sperimenta l'impiego di sistemi algoritmici in sede processuale. In effetti, mancando esperienze significative di giustizia algoritmica negli ordinamenti europei, ci è parso fondamentale vedere nell'esperienza nordamericana le modalità con cui i giudici affrontano tale impatto. Solo dopo potremo estendere, con il metodo della comparazione, i risultati raggiunti dalla giurisprudenza americana anche alle ipotesi oggi più ricorrenti in Europa e ai possibili sviluppi futuri della giustizia algoritmica.

Precisiamo in ultimo che esula dagli scopi del presente lavoro ricostruire l'ampio dibattito costituzionale sui confini della discrezionalità interpretativa del giudice, costruito principalmente sul disposto dell'art. 101, secondo comma, della costituzione secondo cui i "i giudici sono soggetti soltanto alla legge".

Preme soltanto qui ribadire, come già indicato nel capitolo introduttivo, il profondo mutamento del ruolo del giudice, e dello stesso principio della soggezione esclusiva alla legge, avvenuto a seguito almeno di due momenti della modificazione del sistema delle fonti, l'entrata in vigore della Costituzione e la progressiva realizzazione dell'ordinamento sovranazionale (qui intendendosi sia l'integrazione europea sia l'incidenza della giurisprudenza della Corte europea dei diritti dell'Uomo)⁵⁴⁶.

Gli strumenti interpretativi oggi in mano al giudice comune sono sempre più complessi da maneggiare e la concezione del giudice quale "bocca della legge" è ormai superata⁵⁴⁷.

⁵⁴⁶ Tra i molti vedi. R. ROMBOLI, *Il ruolo del giudice in rapporto all'evoluzione del sistema delle fonti ed alla disciplina dell'ordinamento giudiziario*, in *Associazione per gli studi e le ricerche parlamentari*, Quaderno n. 16, Torino, 2006, 63-87.

⁵⁴⁷ Mauro Cappelletti affermava che il suo scopo non era quello di «dimostrare la verità banale, anche se infinite volte in ogni epoca e con tanta inesauribile perseveranza negata o nascosta, della creatività della giurisprudenza», si cfr. M. CAPPELLETTI, *Giudici legislatori?*, Milano, 1984, I.

Il mutamento e la maggiore complessità del sistema delle fonti ha reso il giudice protagonista della tutela dei diritti, tanto che si è parlato di “giudice dei diritti” come evoluzione del “giudice della legalità”⁵⁴⁸.

Lo spazio attraverso cui il giudice esercita tale rinnovato ruolo è ovviamente, prima di tutti, il processo.

A differenza del legislatore, il giudice “produce” o, se si vuole, “inventa”⁵⁴⁹ il diritto solo attraverso la processualità del suo agire, vale a dire a fronte della richiesta delle parti e rispettando il rito processuale e le sue garanzie.

Ed è per questo che le analisi dei successivi capitoli avranno ad oggetto dapprima l’influenza degli algoritmi sulla discrezionalità del giudice e le sue ricadute sul suo ruolo di decisore terzo, imparziale e indipendente; in seguito, la giustizia algoritmica verrà indagata dal punto di vista dei soggetti destinatari di decisioni assunte anche mediante algoritmi, per verificare la tenuta delle garanzie costituzionali.

Per istruire il discorso, come già anticipato, si muoverà dalla folta giurisprudenza americana che ha sindacato la legittimità dei *risk assessment* alla luce dell’*abuse of discretion* e del *due process of law*.

SEZIONE I. LA GIURISPRUDENZA AMERICANA SULL’IMPIEGO DEI *RISK ASSESSMENT TOOLS* NEL PROCEDIMENTO PENALE

2. L’insegnamento giurisprudenza statunitense: l’impatto dei *risk assessment tools* sulla discrezionalità giudiziaria

⁵⁴⁸ In questo senso, si. v. F. G. PIZZETTI, *Il giudice nell’ordinamento complesso*, Milano, 2003, pp.243 ss. nonché R. ROMBOLI, *Il ruolo del giudice in rapporto all’evoluzione del sistema delle fonti ed alla disciplina dell’ordinamento giudiziario*, cit.

⁵⁴⁹ P. GROSSI, *Ritorno al diritto*, cit. Già in tempi remoti Calamandrei scriveva efficacemente «non sappiamo più che farci dei giudici di Montesquieu, “êtres inanimés” fatti di pura logica. Vogliamo dei giudici con l’anima, giudici engagés, che sappiano portare con vigile impegno umano il grande peso di questa immane responsabilità che è il rendere giustizia», si v. P. CALAMANDREI, *Opere giuridiche. Volume I. Problemi generali del diritto e del processo*, RomaTre-Press, 2019, p. 650.

Come già anticipato, il dibattito europeo e italiano sulla giustizia algoritmica ha spesso trascurato l'ampia casistica giurisprudenziale che da anni si sta sviluppando oltreoceano rispetto all'impiego di siffatti sistemi nel processo penale, finendo non di rado per adottare una prospettiva molto lontana da ciò oggi sembra essere il piano di realtà su cui il giurista per sua natura è chiamato a muoversi.

Poco indagata è stata, infatti, la cospicua casistica giurisprudenziale che, almeno dal 2008 sino ad oggi, è stata chiamata a sindacare la legittimità dell'utilizzo dei *risk assessment tools* nella fase di *sentencing* del processo penale.

L'oggetto di siffatto sindacato riguarda sia il rispetto della *due process clause* – dal caso *Loomis* in poi – sia, nella maggior parte delle decisioni, l'abuso di discrezionalità del giudice (*abuse of discretion*, si v. *supra* cap. VII) nel momento in cui questi fonda la propria decisione (anche) sulle risultanze di un algoritmo predittivo.

L'analisi della giurisprudenza americana suggerirà dei percorsi di riflessione sull'impiego dell'algoritmo e sulla sua influenza sulla decisione, sulle garanzie del giusto processo, sul ruolo del giudice e sull'esercizio del suo potere anche dal punto di vista del diritto costituzionale europeo, in particolar modo su come cambia il suo potere decisionale, tenendo a mente in ogni caso che, come si è autorevolmente affermato in tempi risalenti, «potere e discrezionalità vengono [...] a identificarsi»⁵⁵⁰.

Mentre a livello federale ancora non sono stati discussi casi significativi rispetto all'impiego dei *tools* nel *sentencing*⁵⁵¹ in alcune giurisdizioni statali – in particolare l'Indiana, il Wisconsin, il Michigan e lo Iowa – si è sviluppato un numero rilevante di pronunce da cui è possibile trarre alcune tendenze che appaiono in continua evoluzione. Ciò peraltro non sorprende, se si considera che la stragrande maggioranza dei procedimenti penali sono celebrati all'interno delle giurisdizioni statali più che nella giurisdizione federale⁵⁵².

⁵⁵⁰ C. MORTATI, *Note sul potere discrezionale*, 1936, ora in *Problemi di diritto pubblico nell'attuale esperienza costituzionale repubblicana*, Raccolta di scritti, vol. III, 1972, p. 1009.

⁵⁵¹ Ma si cfr. *infra* (cap. IX § 3) dei casi affrontati dalle Corti Federali in altri ambiti che offrono degli spunti interessanti per la riflessione qui condotta.

⁵⁵² Negli ultimi dati disponibili si vede come mentre i casi federali rimangono nell'ordine delle decine di migliaia (intorno ai sessantacinque mila), nelle giurisdizioni statali i numeri si aggirano

2.1. I primi giudizi sull'impiego dei tools algoritmici nel sentencing: la deresponsabilizzazione del giudice

Solo con *Loomis* l'attenzione si sposta sul rispetto delle garanzie del *due process of law*, mentre nella giurisprudenza precedente il *focus* rimane sull'esercizio abusivo della discrezionalità da parte del giudice al cospetto di risultati algoritmici.

Nelle prime pronunce rinvenute rispetto all'utilizzo di algoritmi predittivi nel *sentencing* emerge una forte cautela da parte dei giudici. In particolare, due sentenze del 2008 adottate in Indiana risultano particolarmente significative in quanto per la prima volta toccano, seppur marginalmente, la problematica dell'impatto degli algoritmi sul processo decisionale del giudice.

Nel caso *Cardwell v. State*⁵⁵³, invero, non vi era stata alcuna applicazione di sistemi algoritmici per la determinazione della pena: e tuttavia, la Corte suprema dell'Indiana, all'esito di un procedimento instaurato per *abuso di discrezionalità* del giudice circondariale, coglie l'occasione per svolgere alcune considerazioni di portata generale sull'impiego di "automatismi" nell'ambito del *sentencing*, considerazioni che verranno riprese anche nella giurisprudenza successiva.

Il caso riguardava la condanna di *Cardwell* a trentaquattro anni di carcere per lesioni dolose a danni di una minore. Il ricorrente aveva impugnato la sentenza *per abuse of discretion* poiché la Corte territoriale non aveva valutato attentamente tutte le circostanze attenuanti. Nell'accogliere il ricorso, e rinviare alla Corte distrettuale con l'indicazione di applicare al ricorrente una pena di diciassette anni, la Corte suprema, riferendosi con toni duri al sistema federale di linee guida obbligatorie per il *sentencing*⁵⁵⁴, afferma perentoriamente che il sistema giurisdizionale dell'indiana «*has never adopted a mechanical approach to sentencing, and we have not identified any inflexible system that did not raise more problems than it solved*».

intorno a dieci milioni di casi annuali. Si cfr., per la consultazioni dei dati i seguenti indirizzi, consultati il 15 agosto 2024: <https://www.ussc.gov/about/annual-report-2023> (federale) <https://www.courtstatistics.org/court-statistics/interactive-caseload-data-displays/csp-stat-nav-cards-first-row/csp-stat-criminal> (statale).

⁵⁵³ *Cardwell v. State*, 895 N.E.2d 1219 (2008).

⁵⁵⁴ Si v. *supra* capitolo I § 2.

Con parole ancora più decise la Corte ritiene che

[a]ny effort to force a sentence to result from some algorithm based on the number and definition of crimes and various consequences removes the ability of the trial judge to ameliorate the inevitable unfairness a mindless formula sometimes produces.

Le affermazioni della Corte dell'Indiana sono interessanti sia di per sé sia se rapportate alla giurisprudenza successiva.

Quanto al primo aspetto, esse toccano un punto che è centrale nel dibattito sull'impiego di algoritmi predittivi nel sistema giurisdizionale, vale a dire se sia desiderabile che il risultato di un algoritmo possa in qualche modo vincolare la decisione del giudice: secondo la Corte non sarebbe auspicabile in quanto impedirebbe al giudice di mitigare l'inevitabile ingiustizia («*inevitable unfairness*») che un automatismo – definito molto efficacemente una «*mindless formula*» – talvolta produce, e applicare così una sentenza individualizzata. Inoltre, esse assumono importanza anche perché rappresentano, come subito si vedrà, un primo approdo cui la giurisprudenza si confronterà per sviluppare le proprie considerazioni rispetto all'incidenza di algoritmi predittivi sul processo decisionale del giudice.

In tal senso, nello stesso anno la sentenza *Rhodes v. State*⁵⁵⁵ della Corte di Appello dell'Indiana riprende le parole espresse dalla *Supreme Court* in *Cardwell*, adattandole a un caso in cui vi era stata in effetti la considerazione da parte del giudice del risultato prodotto da un algoritmo.

Nella vicenda in esame due uomini, tra cui il sig. Rhodes, venivano fermati dalla polizia a bordo di un'automobile rubata e al cui interno venivano rinvenute massicce quantità di droga. Rhodes veniva dichiarato colpevole per il reato di spaccio e condannato dalla *Trial Court* alla pena a trent'anni di carcere. La pena veniva determinata anche sulla base del risultato – contenuto nel PSI – prodotto da

⁵⁵⁵ *Rhodes v. State*, 896 N.E.2d 1193 (2008).

LSI-R (*Level of Service Inventory-Revised*), uno strumento di *risk assessment* di terza generazione⁵⁵⁶ che produceva uno *score* di rischio alto.

A seguito della condanna il sig. Rhodes impugnava la sentenza per *abuse of discretion*, in quanto la Corte aveva considerato il risultato algoritmico quale *improper aggravating factor*.

La Corte d'Appello, pur non annullando la pena, ritenuta comunque appropriata, riscontra l'abuso di discrezionalità da parte del giudice di prime cure sostenendo in primo luogo che, se tali sistemi possono essere utili ai dipartimenti per le misure alternative alla detenzione al fine di determinare i più idonei trattamenti rieducativi, il loro utilizzo da parte del giudice «*is contrary to the essential function of the trial court in sentencing*»⁵⁵⁷.

In particolare, nel *sentencing* al giudice è assegnato il precipuo ruolo di «*create an appropriate, individualized sentence*», e quindi valutare *in autonomia* quei fattori che invece sono stati inseriti nel sistema algoritmico e che hanno contribuito a determinare il rischio di recidiva (quali ad esempio la storia criminale dell'imputato, la condizione familiare, lavorativa, sociale e così via).

Solo così è possibile evitare la deresponsabilizzazione del giudice nella determinazione della pena. Afferma ancora il giudice dell'Indiana:

*[t]he use of a standardized scoring model, such as the LSI-R, undercuts the trial court's responsibility to craft an appropriate, individualized sentence. Relying upon a sum of numbers purportedly derived from objective data cannot serve as a substitute for an independent and thoughtful evaluation of the evidence presented for consideration*⁵⁵⁸.

In conclusione, richiamando le parole espresse in *Cardwell*, il giudice dell'Indiana afferma che «*it is an abuse of discretion to rely on scoring models to determine a sentence*»⁵⁵⁹.

⁵⁵⁶ Si v. *supra* capitolo I § 3.

⁵⁵⁷ *Rhodes*, cit.,

⁵⁵⁸ *Rhodes*, cit.

⁵⁵⁹ *Ibidem*.

Come emerge da queste prime sentenze, inizialmente traspare grande preoccupazione per il possibile inaridimento della decisione discrezionale del giudice e nel valutare gli elementi rilevanti per determinare una decisione che sia individualizzata e non frutto di dati statistici e correlazioni algoritmiche. In altre parole, è solo il giudice che deve porre il proprio sguardo sugli elementi determinanti per la sentenza sulla determinazione della pena, senza che tale sguardo possa invece posarsi sul risultato finale dell'elaborazione algoritmica.

Proprio su tali conclusioni si appunteranno le critiche dei successivi arresti giurisprudenziali che, in esplicita contrapposizione a *Cardwell* e *Rhodes*, indirizzeranno il dibattito sulla legittimità dell'impiego dei *tools* nel *sentencing*.

2.2. (segue) I primi argomenti a favore dell'impiego dei tools nel sentencing

Poco tempo dopo le due sentenze del 2008, sempre in Indiana comincia a nascere un indirizzo giurisprudenziale ad esse contrario, che non riconosce un abuso di discrezionalità da parte del giudice e che anzi auspica l'impiego generalizzato dei *tools* nel *sentencing*, adottando una posizione assai permissiva sul punto.

In *Malenchik v. State*⁵⁶⁰, a seguito della dichiarazione di colpevolezza l'imputato veniva condannato a sei anni di carcere per il reato di ricettazione. Per determinare l'ammontare della pena il giudice si basava sullo *score* di due strumenti algoritmici, LSI-R (su cui si v. *supra*, capitolo I § 3) e SASSI (*Substance Abuse Subtle Screening Inventory*), uno specifico strumento di *risk assessment*, anche esso di terza generazione, che valuta il rischio di avere in futuro una dipendenza da sostanze stupefacenti.

L'imputato, richiamando a suo sostegno il caso *Rhodes*, appellava la sentenza e riteneva quindi «*improper for the trial court to take into consideration*» il risultato algoritmico, in quanto tale considerazione «*may lead to an unwise fundamental change in Indiana's sentencing system*». Inoltre, secondo l'imputato, tali sistemi mancavano di «*objective reliability*» ad essere «*unfairly discriminatory*».

⁵⁶⁰ *Malenchik v. State*, 928 N.E.2d 564, 571–73, 2010.

Lo Stato rispondeva che strumenti come LSI-R e SASSI, se impiegati nel modo corretto e con le giuste limitazioni, possono essere utilizzati come «*a legitimate and valuable contribution to the sentencing process*».

La Corte Suprema dell'Indiana dapprima precisa – ed è una precisazione significativa, come si vedrà nel corso del capitolo – che la decisione sulla determinazione della pena era basata non solo sulle risultanze dei due sistemi predittivi utilizzati ma su altri, autonomi, fattori: la Corte distrettuale pertanto «*did not rely on either the LSI-R or SASSI as an independent aggravating factor in deciding to impose more than the advisory sentence*».

Premesso che dunque la Corte distrettuale non si è totalmente affidata (*reliance*) al risultato statistico, i giudici supremi si soffermano poi sulla possibilità per la Corte circondariale di considerare (*consider*), tra i diversi fattori su cui fonda la decisione, anche gli *score* algoritmici. Dopo un'accurata analisi di quanto prescrive lo statuto dell'Indiana, che come già visto nulla dice sull'utilizzo dei *tools* nel *sentencing* (si v. *supra* § cap. I § 5) i giudici si smarcano inizialmente dalle critiche rivolte in *Cardwell* all'impiego di un «*mechanical approach*», affermando anzi che l'obiettivo insito della giustizia penale è che «*an individualized penal consequence should be efficacious in achieving the goals of reformation and minimizing recidivism*», obiettivo il cui approccio *evidence-based* dei *tools* «*has considerable promise*».

Pur riconoscendo che sistemi come quelli usati dalla Corte circondariale non sono stati progettati per «*substitute for the judicial function of determining the length of sentence appropriate for each offender*», è tuttavia

highly useful and important for trial courts to consider as a broad statistical tool to supplement and inform the judge's evaluation of information and sentencing formulation in individual cases.

Quanto poi al rischio di deresponsabilizzazione del giudice e di *abuse of discretion*, sollevati in *Rhodes* (si v. *supra*), la *Supreme Court* afferma perentoriamente: «*[w]e disagree*». All'esatto opposto, la considerazione di siffatti strumenti, pur non essendo volta a determinare la gravità della sanzione penale,

can enhance a trial judge's individualized evaluation of the sentencing evidence and selection of the program of penal consequences most appropriate for the reformation of a particular offender.

In tal modo i giudici supremi confidano nel «*sound discernment and discretion*» con cui il giudice dovrebbe assegnare agli strumenti predittivi «*proper consideration and appropriate weight*».

In conclusione, tali strumenti, nel giudizio della Corte, non sono stati usati come circostanze aggravanti rispetto alla pena da applicare, sicché

the trial court did not err in considering the LSI-R and SASSI test results in formulating the defendant's program of penal consequences.

In *J.S. v. State*⁵⁶¹, coeva a *Malenchik*, viene in risalto il medesimo profilo visto da una prospettiva opposta: la contestazione di un abuso di discrezionalità non per aver la Corte fatto proprie le risultanze dell'algoritmo, ma per averle completamente trascurate.

Nella vicenda di specie l'imputato veniva dichiarato colpevole per due distinti capi di accusa concernenti molestie su minori e gli veniva comminata una pena di otto anni per il primo e di quattro per il secondo. Anche in questo caso il *PSI report* conteneva anche il risultato prodotto dal sistema LSI-R che assegnava un punteggio di rischio particolarmente basso («*placed the defendant into the Low Risk/Needs category*»). Il giudice, tuttavia, riteneva di non dare peso allo *score* algoritmico, dando peso agli altri elementi:

that's not a good measure in cases involving sexual abuse, so I'm not going to treat that either as an aggravator or a mitigating factor. I do find that the aggravating factors outweigh the mitigating factors.

Anche in questo caso l'imputato impugnava la sentenza per *abuse of discretion*: mutuando le motivazioni spese in *Malenchik*, la Corte ha buon gioco nel rigettare l'abuso di discrezionalità in quanto il sistema algoritmico non era da considerarsi

⁵⁶¹ *J.S. v. State*, 928 N.E.2d 576 (2010).

alla stregua di una circostanza aggravante o attenuante – gli strumenti di valutazione non sono infatti preparati dal giudice che emette la sentenza, ma dai funzionari dei dipartimenti, e le loro valutazioni possono non coincidere con le conclusioni del giudice del processo – quanto piuttosto solo uno strumento per

supplement and enhance a judge's evaluation, weighing, and application of the other sentencing evidence in the formulation of an individualized sentencing program appropriate for each defendant.

In definitiva, appare chiaro l'orientamento assunto dai giudici supremi dell'Indiana: allorché il giudice nel determinare la pena debba interfacciarsi con il risultato prodotto da un *risk assessment tool*, la decisione non è impugnabile per *abuse of discretion* né quando l'algoritmo assume un peso rilevante, né quando il giudice se ne discosta, essendo quest'ultimo solo uno dei tanti elementi che il decidente deve considerare per assumere la decisione.

Su posizioni simili si assesta la successiva giurisprudenza maturata in Wisconsin, prima con il caso *Samsa* dinanzi alla Corte d'Appello e poi, finalmente, nel celebre caso *Loomis* di fronte alla Corte Suprema statale, che verrà preso autonomamente in considerazione nei paragrafi successivi per l'importanza che assume nell'economia del presente lavoro.

In *Samsa v. State*⁵⁶², l'imputato veniva condannato alla pena di cinque anni di carcere più cinque di *extended supervision* per il reato di violenza sessuale, nonostante nel PSI, anche sulla base del risultato dell'algoritmo COMPAS, venisse raccomandata l'applicazione della libertà vigilata per venire incontro ai *criminogenic needs* del soggetto (per le caratteristiche di COMPAS si v. *supra* capitolo I § 4).

Il giudice distrettuale analizzava autonomamente i risultati di COMPAS e dichiarava la propria difficoltà a capire la raccomandazione del dipartimento di applicare la libertà vigilata («*frankly, I am at a loss to figure out this agent's recommendation*»). Secondo il giudice, in effetti, seppur alcuni elementi delle scale di COMPAS indicassero un rischio basso, ciò non era sufficiente a convincerlo che

⁵⁶² *State v. Samsa*, 359 Wis.2d 580 (2014).

la soluzione della libertà vigilata invece della soluzione detentiva fosse la più adeguata («[a]nd I just – I disagree»).

Samsa appellava la sentenza lamentando una scorretta interpretazione dei risultati di COMPAS e chiedeva alla Corte la rideterminazione della pena. In particolare, l'imputato si doleva del fatto che la Corte avesse ritenuto il rischio di recidiva più alto rispetto a quello indicato da COMPAS.

Anche in tal caso la Corte d'Appello precisa che il *sentencing* deve rimanere oggetto di «*great discretion*» e che l'interpretazione delle diverse scale di COMPAS «*was perfectly within the court's discretion*», in quanto

[t]he COMPAS is merely one tool available to a court at the time of sentencing and a court is free to rely on portions of the assessment while rejecting other portions.

In definitiva:

Samsa's assertion that a sentencing court should be prohibited from considering criminogenic needs when assessing a defendant's risk of reoffence runs contrary to the principle that sentencing courts are afforded substantial discretion.

Anche i giudici del Wisconsin rigettano quindi l'argomento dell'abuso di discrezionalità, sulla base della considerazione che tali sistemi non possono e non devono catturare la discrezionalità del giudice, ma piuttosto sono elementi che il giudice ha la facoltà di considerare per assumerne una decisione più informata.

Come si vede dall'analisi di questi primi casi giurisprudenziali, fino al caso *Loomis* la partita non si gioca direttamente sul piano del rispetto delle garanzie del *due process*, quanto piuttosto sull'impatto l'ingresso di strumenti algoritmici nel processo produce sul potere discrezionale del giudice, da cui poi, come vedremo, generano le violazioni del giusto processo.

Il sindacato sull'*abuse of discretion*, comunque, non scompare nei casi successivi a *Loomis*: esso verrà piuttosto sollevato come argomento ulteriore, insieme con le ulteriori doglianze che hanno come punto focale le possibili

violazioni del *due process of law*, che hanno riscosso maggiore attenzione nella letteratura.

3. (Segue) L'insegnamento della giurisprudenza statunitense: i *risk assessment tools* alla prova del *due process of law*

La svolta decisiva sugli algoritmi impiegati nei sistemi giudiziari avviene con il celebre caso *Loomis*⁵⁶³, un caso del 2016 che ha dato vita, tanto negli Stati Uniti quanto in Europa, al dibattito sull'impiego dell'IA nei procedimenti giudiziari. Ciò in quanto, come già anticipato, per la prima volta si affronta su un piano costituzionale (*constitutional grounds*) la compatibilità dell'impiego dei *tools* con il *due process of law*.

Tale dibattito, come già anticipato nel capitolo introduttivo, ha spesso assunto in Europa delle pieghe avanguardistiche e distopiche che nulla avevano a che vedere con il caso e che hanno dato la spinta a numerosi fraintendimenti e, per lungo tempo, spostato l'attenzione sulla sostituzione del giudice con sistemi algoritmici piuttosto che sui (numerosi) problemi che il caso presentava⁵⁶⁴.

3.1. (segue) Il caso Loomis

Quanto ai fatti della vicenda, il Sig. Loomis veniva accusato di cinque distinti capi di accusa⁵⁶⁵ legati a una sparatoria in auto. Nella procedura di patteggiamento (*guilty plea*) si dichiarava colpevole delle due accuse meno gravi (tentativo di fuga da pubblico ufficiale e guida di un'auto senza il consenso del proprietario) e veniva condannato complessivamente a sei anni di reclusione e a successivi cinque anni di *extended supervision*.

⁵⁶³ *State v. Loomis*, 881 N.W.2d 749 (Wis 2016) (n 31).

⁵⁶⁴ Si veda, ad esempio, il titolo sensazionalistico apparso sul New York Times rispetto al caso *Loomis*, A. LIPTAK, *Sent to Prison by a Software Program's Secret Algorithms*, The New York Times, maggio 2017.

⁵⁶⁵ In particolare: messa in pericolo della pubblica sicurezza; tentativo di fuga da un pubblico ufficiale; guida di un'auto senza consenso del proprietario; possesso di arma da fuoco; possesso di un fucile o carabina a canna corta.

A seguito del patteggiamento, la Corte richiedeva il PSI, al cui interno veniva inclusa la valutazione del rischio di recidiva effettuata da COMPAS che indicava uno *score* elevato (*high risk*). Nel PSI si precisava inoltre che i risultati dell'algoritmo «*are not intended to determine the severity of the sentence or whether an offender is incarcerated*»⁵⁶⁶. Nella decisione sulla pena, tuttavia, la Corte utilizzava senz'altro il risultato di COMPAS per determinare l'ammontare della sanzione: «*the COMPAS report that was completed in this case does show the high risk and the high needs of the defendant. There's a high risk of violence, high risk of recidivism, high pre-trial risk; and so all of these are factors in determining the appropriate sentence*»⁵⁶⁷.

A seguito della condanna l'imputato impugnava la sentenza e chiedeva la rideterminazione della pena in quanto la considerazione di COMPAS nel *sentencing* aveva violato il suo diritto a un giusto processo.

Nell'udienza fissata per discutere la richiesta del Sig. Loomis (*post-conviction motion*), in cui l'imputato presentava altresì la perizia di un esperto il quale indicava gli usi distorti di COMPAS e le minime informazioni che i giudici avevano sul loro funzionamento, la Corte rigettava ogni argomento sostenendo che COMPAS era stato usato solo per corroborare (*corroborate*) la decisione sulla pena, la quale sarebbe stata imposta nella medesima misura anche senza la considerazione dello *score* algoritmico.

In particolare le doglianze dell'imputato riguardavano tre distinti corollari del giusto processo: l'uso di COMPAS violava il diritto ad essere giudicati sulla base di un'*accurate information* poiché la natura proprietaria dell'algoritmo privava l'imputato della possibilità di accedervi e contestarne la validità scientifica; in secondo luogo l'affidamento da parte del giudice su un sistema automatico di valutazione del rischio violava il suo diritto a una *individualized sentence*; infine, essendo presente, nei dati elaborati da COMPAS, anche il genere come fattore criminogeno, la considerazione della valutazione del rischio da parte della Corte

⁵⁶⁶ § 17

⁵⁶⁷ § 18

circondariale violava il diritto costituzionale a non essere giudicato sulla base del genere.

Quanto alla prima contestazione, per supportare le proprie argomentazioni il Sig. Loomis portava all'attenzione della Corte i casi *Gardner* e *Skaff* (si cui si v. *infra* § 7). In particolare, poiché la società produttrice del *software* – Northpointe (ora Equivant) – «*considers COMPAS a proprietary instrument and a trade secret*», essa non rivela «*how the risk scores are determined or how the factors are weighed*». Di conseguenza, l'imputato «*has been denied information which the circuit court considered at sentencing*» e quindi era impossibile verificare l'*accuracy* del sistema⁵⁶⁸.

La Corte non condivide l'impostazione della difesa, sostenendo che anche se l'imputato non aveva accesso all'algoritmo, aveva comunque la possibilità di «*review and challenge*» i risultati del test algoritmico inseriti nel PSI⁵⁶⁹. Inoltre, secondo la Corte, la guida predisposta da Northpointe (si v. *supra*, cap. I § 4) ne spiegava adeguatamente il funzionamento. Nella visione della Corte, mentre nelle sentenze portate da Loomis a supporto delle proprie posizioni non veniva consentito all'imputato di accedere alle informazioni contenute nel PSI, in questo caso il giudice e l'imputato «*had access to the same copy of the risk assessment. Loomis had an opportunity to challenge his risk scores by arguing that other factors or information demonstrate their inaccuracy*»⁵⁷⁰.

Con riguardo alla seconda doglianza – l'impossibilità per il giudice di poter pronunciare una sentenza individualizzata poiché, come già visto, la predizione del rischio non riflette una valutazione individuale quanto piuttosto l'elaborazione di dati statistici di un gruppo di persone⁵⁷¹ - la Corte precisa:

⁵⁶⁸ § 52

⁵⁶⁹ § 53

⁵⁷⁰ § 56.

⁵⁷¹ Come visto, la stessa guida COMPAS riferisce che «*[r]isk assessment is about predicting group behavior . . . it is not about prediction at the individual level*». Si v. *supra* capitolo I § 4.

*[i]f a COMPAS risk assessment were the determinative factor considered at sentencing this would raise due process challenges regarding whether a defendant received an individualized sentence*⁵⁷².

Senonché – prosegue la Corte – non vi è alcuna violazione al diritto a una sentenza individualizzata poiché la «*consideration of a COMPAS risk assessment at sentencing*» è accompagnata dalla considerazione di «*other supporting factors*»: con la conseguenza che, all’opposto, l’utilizzo della valutazione algoritmica del rischio «*is helpful in providing the sentencing court with as much information as possible in order to arrive at an individualized sentence*»⁵⁷³.

Nonostante il rigetto anche di tale argomento, la Corte afferma comunque che le caratteristiche dei *tools* – che basano le proprie predizioni su dati aggregati e tendenze di gruppo – impongono al giudice di tener conto di questo e pesare adeguatamente «*all of the factor that are relevant to sentencing an individual defendant*».

Quanto infine al terzo argomento della difesa secondo cui la violazione del giusto processo deriverebbe anche dalla considerazione di COMPAS del genere quale elemento su cui effettuare la valutazione del rischi, e quindi il diritto a non venire giudicati sulla base del genere – argomento che in questo lavoro non verrà approfondito in quanto non attiene propriamente al giusto processo per come viene inteso nel diritto costituzionale europeo quanto piuttosto al diritto alla non discriminazione – la Corte, citando alcuni studi scientifici di analisi di COMPAS, conclude che l’uso del genere nella valutazione statistica «*promotes accuracy that ultimately inures to the benefit of the justice system including defendants*»⁵⁷⁴.

Dopo aver rigettato gli specifici argomenti della difesa rispetto alle violazioni del giusto processo, la Corte suprema si sofferma sulla legittimità in generale dell’utilizzo di COMPAS nel *sentencing* e se tale utilizzo comporti un errato esercizio della discrezionalità giudiziaria⁵⁷⁵. Al riguardo ribadisce che tali sistemi

⁵⁷² § 68.

⁵⁷³ § 72.

⁵⁷⁴ §86.

⁵⁷⁵ §§ 87 ss.

non devono essere impiegati per determinare la severità della pena carceraria né tantomeno per stabilire la pena detentiva, ma solo per individuare i programmi riabilitativi più adeguati al condannato⁵⁷⁶. Ne segue che il giudice

*must explain the factors in addition to a COMPAS risk assessment that independently support the sentence imposed. A COMPAS risk assessment is only one of many factors that may be considered and weighted at sentencing*⁵⁷⁷.

In aggiunta, i giudici supremi ritengono che ogni PSI debba contenere un “libretto delle istruzioni” contenente tutti le limitazioni di COMPAS e informare il giudice di alcune cautele da tenere in considerazione, e in particolare: la natura proprietaria di COMPAS e la conseguente impossibilità di comprendere come il rischio viene determinato; l’impossibilità del sistema di fornire una valutazione di rischio individualizzata, essendo quest’ultima basata su dati statistici aggregati; la possibilità che tali sistemi non siano accurati e potenzialmente discriminatori; la necessità di un continuo monitoraggio di tali sistemi e l’adeguamento della valutazione statistica alla popolazione del Wisconsin; la considerazione che COMPAS non è stato designato per il *sentencing*, ma per adiuvarne il *Department of Corrections* nelle valutazioni rispetto al trattamento e alla supervisione del condannato.

In conclusione, poiché agli occhi della Corte Suprema il giudice del *sentencing* era parso consapevole delle suddette limitazioni e, nonostante avesse considerato il risultato di COMPAS, non gli aveva dato peso e anzi avrebbe imposto «*the same sentence*»⁵⁷⁸, la valutazione algoritmica non era stata il fattore decisivo per la determinazione della pena del Sig. Loomis. Per tale ragione la Corte conclude così la propria motivazione:

⁵⁷⁶ § 93.

⁵⁷⁷ § 99.

⁵⁷⁸ §§ 105 e 106.

*a circuit court's consideration of a COMPAS risk assessment at sentencing does not violate a defendant's right to due process and that the circuit court did not erroneously exercise its discretion here*⁵⁷⁹.

A seguito della decisione del rigetto da parte Suprema Corte del Wisconsin, l'imputato richiedeva il *Writ of Certiorari* alla Corte suprema degli Stati Uniti, uno strumento con cui si chiede una *judicial review* di una sentenza di una Corte statale⁵⁸⁰. Nel giugno del 2017 la US Supreme Court negò il *certiorari*, ma è comunque interessante segnalare che in un documento depositato per opporsi alla richiesta di revisione da parte di *Loomis*, il Governo statunitense riconosce che «*some uses of an undisclosed risk-assessment algorithm might raise due process concerns if, for example, a defendant is denied access to the factual inputs out his criminal and personal history, or if his risk scores form part of a sentencing 'matrix' or establish a 'presumptive' term of imprisonment*»⁵⁸¹.

Come si vedrà, la giurisprudenza *Loomis* è il caso di studio su cui si sono soffermati la maggior parte dei commentatori sia nordamericani sia europei; non solo, essa rappresenta anche il punto di svolta per la valutazione della legittimità dei *risk assessment tools* alla luce delle garanzie del *due process*, e per tale ragione essa viene assunta come *benchmark* dalla giurisprudenza successiva.

3.2. (segue) La giurisprudenza successiva a Loomis

Come detto, dopo *Loomis* in altre giurisdizioni le difese hanno tentato di portare all'attenzione delle Corti argomenti di carattere costituzionale afferenti alla violazione del giusto processo, senza tuttavia ottenere grande successo.

⁵⁷⁹ La giurisprudenza successiva in Wisconsin, richiamando la motivazione espressa dalla Corte suprema in *Loomis*, nega che l'impiego dei *tools* violi il *due process of law*. Si v. *State v. Olejniczak*. (2016) e *State v. Braxton*, (2019).

⁵⁸⁰ Il *writ of certiorari* è stato introdotto con il *Judiciary Act* del 1891, si attiva su istanza di parte e la scelta di emettere il *writ* da parte della Corte Suprema è del tutto discrezionale. Si cfr. C. BASSU - M. BETZU - F. CLEMENTI - G. COINU, *Diritto costituzionale degli Stati Uniti d'America. Una introduzione*, cit., pp. 82-83.

⁵⁸¹ Il documento è consultabile al seguente indirizzo: <https://www.scotusblog.com/wp-content/uploads/2017/05/16-6387-CVSG-Loomis-AC-Pet.pdf>.

In Iowa due sentenze coeve del 2018 si concentrano ancora sull'impiego dei *tools* nel *sentencing*, fornendo nuovi spunti di riflessione che verranno affrontati nel prosieguo del capitolo.

In *Gordon*⁵⁸², l'imputato veniva accusato di violenza sessuale nei confronti di una minore. Nel PSI venivano inclusi i risultati di due diversi *risk assessment tools* specificamente progettati per i *sex offenders* (STATIC-99R e SOTIPS). Nel primo veniva indicato un rischio di recidiva nella media, mentre nel secondo un rischio alto di recidiva.

Sulla base dei risultati, nel PSI si raccomandava la sospensione della pena e l'ingresso dell'imputato in una comunità riabilitativa. La Corte distrettuale disattendeva le raccomandazioni e, sulla base dei livelli di rischio, condannava l'imputato alla pena detentiva. Il Sig. Gordon appellava la sentenza ritenendo, anche sulla base delle cautele indicate in *Malenchik* e *Loomis* (si v. *supra*), che «*the district court violated his right to due process by considering and relying on the sex-offender*»⁵⁸³. E ciò in quanto la determinazione della pena assunta con la considerazione di un sistema algoritmico non del tutto conosciuto implicava che la decisione si fosse basata su un'«*inaccurate information*», in violazione del giusto processo secondo la giurisprudenza della Corte Suprema degli Stati Uniti.

La Corte d'Appello dello Iowa afferma innanzitutto di poter decidere «*on non constitutional grounds*» senza cioè affrontare il tema della violazione del *due process*, ma soffermandosi solo sull'eventuale *abuse of discretion* della Corte distrettuale, che ha considerato gli *score* algoritmici per determinare la pena.

La motivazione della Corte d'Appello ruota intorno alla presenza o meno di un'autorizzazione, negli statuti nazionali, all'utilizzo di sistemi algoritmici di valutazione del rischio. A questo riguardo, benché lo Stato asserisse che la generale formulazione dello *Iowa code* – la quale prevedeva che «*the court shall receive ... "any information" offered which is "relevant to the question of sentencing"*» – permettesse l'impiego dei *tools*, secondo la Corte non vi è nessuna specifica

⁵⁸² *State v. Gordon*, 919 N.W.2d 635, 2018.

⁵⁸³ *Ivi*, § 3.

indicazione che autorizzasse la Corte distrettuale a considerare tali sistemi nel *sentencing*.

Al riguardo, se è vero che alcuni statuti impongono l'utilizzo di sistemi di *risk assessment* in alcune ipotesi specifiche, «*Iowa's general sentencing statutes do not address risk-assessment tools such as the STATIC-99R and SOTIPS included in the psychosexual assessment attached to the PSI report in this case*»⁵⁸⁴ e aggiunge, opportunamente, che i «*risk assessment tools validated for one purpose may not be fit for another*».

Dunque secondo la Corte i programmi progettati specificamente per alcuni fini, non possono essere trapiantati in sé e per sé in altre fasi processuali senza che ci sia un'autorizzazione legislativa.

La Corte ritiene poi di affrontare un secondo argomento sollevato dallo Stato, il quale opina, sulla scia di *Loomis*, che la Corte non si è basata sugli *score* algoritmici per determinare la sentenza (*reliance upon*) ma li avrebbe semplicemente considerati (*consider*) tra i tanti elementi che una Corte è chiamata a valutare. Secondo la Corte, al contrario, non è possibile verificare il peso che ha assunto il risultato algoritmico e la differenza tra *reliance* e *consideration* è del tutto sfuggente:

[w]e are not persuaded that the difference between reliance upon and consideration of these actuarial estimates saves the sentencing process in this case. Our supreme court has long held that it cannot "speculate about the weight" that the district court has "mentally assigned" to an impermissible factor "or whether it tipped the scales to imprisonment»⁵⁸⁵.

In conclusione, poiché la Corte distrettuale ha considerato il risultato algoritmico quale circostanza aggravante per la determinazione della pena, e nessuno statuto autorizza tale impiego, la sentenza viene annullata con rinvio per una nuova sentenza.

⁵⁸⁴ § 6.

⁵⁸⁵ *Ivi*, § 9.

In *Guise*⁵⁸⁶, coeva a *Gordon*, l'imputato veniva condannato a una pena detentiva anche sulla base del risultato del *tool* IRR (Iowa Risk Revised) e appellava la sentenza sulla base di due argomenti: la violazione del *due process of law* e l'*abuse of discretion* della Corte distrettuale.

Anche in questo caso la Corte d'Appello non ritiene necessario affrontare l'argomento costituzionale della violazione del giusto processo, focalizzandosi esclusivamente sull'abuso di discrezionalità del giudice circondariale.

Secondo la Corte, come in *Gordon*, non esiste alcuna autorizzazione legislativa per l'utilizzo di sistemi di valutazione del rischio nel *sentencing*. Inoltre, a sostegno della propria decisione, con argomenti che serviranno per la nostra riflessione, i giudici di appello affermano:

*[w]e do not know what the IRR is, what factors led to the recommendation of intensive supervision, or whether the factors were appropriate for consideration in the sentencing context. It is impossible to determine whether the IRR was relevant to the question of sentencing within the meaning of section 901.2(1) [si tratta di una sezione dello Iowa Code]*⁵⁸⁷.

Inoltre, nel tentativo di valutare a fondo le caratteristiche del sistema e la sua finalità, la Corte precisa che non vi era traccia di come IRR catalogava le informazioni, come assegnava i punteggi, quali fattori venivano considerati e come venivano pesati: «*[w]e simply do not know*», affermano perentoriamente i giudici.

In altre parole:

[t]he IRR as described in Guise's PSI report was a black box, devoid of transparency

⁵⁸⁶ *State v. Guise*, 919 N.W.2d 635, 2018.

⁵⁸⁷ *Guise*, § 3.

e, mancando le informazioni essenziali che consentono alla Corte e all'imputato di valutare l'affidabilità, dell'IRR, «*the court's use of the test in the sentencing decision amounted to an abuse of discretion*»⁵⁸⁸.

Le due sentenze della Corte d'appello verranno poi ribaltate dalla Corte suprema dello Iowa per ragione più processuali che sostanziali⁵⁸⁹. Secondo la Corte Suprema gli imputati avrebbero dovuto lamentare la violazione del giusto processo già davanti alla Corte distrettuale senza poterla proporre per la prima volta in appello. Tuttavia, alcune parole dei Supremi Giudici ci saranno utili per sviluppare le nostre argomentazioni in merito alla trasparenza algoritmica (si v. *infra* alla sezione II del capitolo).

Di recente anche in Michigan, dove da alcuni anni le valutazioni di COMPAS sono entrate a far parte del PSI⁵⁹⁰, sia la Corte d'Appello sia la Corte Suprema sono state chiamate a valutare la conformità al *due process* dell'impiego dei *risk assessment tools* nel *sentencing*.

In *People v. Younglove*⁵⁹¹, i giudici di appello prendevano in considerazione i gravami di diversi imputati le cui condanne erano state determinate anche con l'utilizzo del sistema COMPAS⁵⁹². In particolare, gli imputati lamentavano la violazione del giusto processo sotto diversi profili: come in *Loomis*, lamentavano l'assenza di una sentenza individualizzata in quanto COMPAS si basa su dati statistici aggregati e non sul rischio di recidiva di un singolo imputato; inoltre, sempre sulla scia di *Loomis*, lamentavano l'assenza di studi univoci sulla sua affidabilità, la provata idoneità a produrre risultati discriminatori e la mancanza assoluta di trasparenza. Inoltre, la violazione del giusto processo si fondava, come

⁵⁸⁸ La sentenza, come accaduto in Gordon, viene poi annullata dalla *Supreme Court dello Iowa* per motivi processuali, si cfr. *State v. Guise*, 921 N.W.2d 26, 2018.

⁵⁸⁹ Si v. *State v. Gordon*, 921 N.W.2d 19, 2018 e *State v. Guise*, 921 N.W.2d 26, 2018.

⁵⁹⁰ Si cfr. la esplicativa guida prodotta dal *Department of Correction* dello Stato del Michigan al seguente indirizzo: <https://www.michbar.org/file/news/releases/archives17/COMPAS-at-PSI-Manual-2-27-17-Combined.pdf>.

⁵⁹¹ *People v. Younglove*, No. 341901, 2019.

⁵⁹² Il Sig. Younglove ricorreva in appello contro la condanna a un anno di carcere e tre anni di libertà vigilata inflitta dal tribunale per il possesso di pseudoefedrina per la produzione di metanfetamina; Il Sig. Hegler appellava la condanna da 1 anno e mezzo a 5 anni di reclusione che il tribunale comminava per guida di veicolo in stato di ebbrezza; il Sig. Wilson impugnava la decisione (dai venti ai sessanta mesi di reclusione) per frode; Il Sig. Bradford ricorreva in appello contro la sentenza da 1 anno e mezzo a 5 anni di reclusione per furto di arma da fuoco.

in *Loomis*, sul fatto che la Corte distrettuale si fosse basata su un'«*inaccurate information*» (il risultato di COMPAS), che quindi doveva essere rimossa dal PSI.

La Corte, ripercorrendo la giurisprudenza interna sul giusto processo – sostanzialmente conforme alla giurisprudenza federale – specifica che il *due process* richiede che un imputato sia giudicato sulla base di informazioni affidabili (*accurate*) e che, inoltre, l'imputato deve avere una «*reasonable opportunity at sentencing to challenge the information contained in a PSIR*».

Ciò premesso, analogamente a quanto espresso in *Loomis*, è ancora l'argomento della non esclusività della prova algoritmica a costituire la *ratio decidendi* per rigettare le doglianze afferenti alle violazioni del giusto processo. Secondo i giudici di appello, infatti

sentencing court is not bound by the recommendations included in a PSIR [...]. Nor is a sentencing court obliged to use all information contained within a PSIR in pronouncing sentence. Courts must consider the PSIR, but whether or how heavily to weigh the information contained therein remains within the court's discretion.

Inoltre, la Corte aggiunge che il risultato algoritmico inserito nel PSI non sarebbe così diverso dai pareri specialistici che sono inseriti normalmente nel *report*. Quindi, concludono i giudici d'appello: «*we are not persuaded by defendants' arguments that the inclusion of COMPAS information unfairly influences or replaces a sentencing court's individual sentencing discretion*».

Più di recente sempre in Michigan, nella decisione *People v. Canedo*⁵⁹³ della Corte suprema, di cui è disponibile esclusivamente la decisione della Corte che nega la possibilità di impugnare la sentenza del giudice di prime cure, assume grande rilevanza la *concurring opinion* della giudice Cavanagh, la quale coglie un punto chiave: se è vero, afferma la giudice, che l'imputato ha diritto a essere giudicato sulla base di informazioni accurate

⁵⁹³ *People v. Canedo*, 507 Mich. 1029, 2021.

in the context of COMPAS risk assessments, it is unclear to me what it might mean to measure the accuracy of a prediction about an individual's future conduct and how that prediction might be challenged without knowing how it was formulated.

Di fronte a un algoritmo come COMPAS che è «*proprietary, and undisclosed*», prosegue la giudice – è dunque compito arduo per un giudice esercitare correttamente il proprio potere discrezionale e adempiere *tout court* alla sua funzione.

Come risulta dalle numerose sentenze qui prese in considerazione, salvo pochi casi in cui viene messa in dubbio la legittimità di applicazione algoritmiche nel *sentencing*, la giurisprudenza che si forma all'interno delle giurisdizioni statali, con argomento spesso analoghi, rifiuta di considerare l'impiego dei *risk assessment tools* quali esercizio abusivo della discrezionalità o violazioni del giusto processo.

SEZIONE II. I DIVERSI LIVELLI DI INFLUENZA DELL'ALGORITMO SULLA DECISIONE DEL GIUDICE

4. L'autorità algoritmica nel procedimento decisionale del giudice

I casi menzionati ci forniscono un primo punto di caduta e fanno luce su una delle problematiche di cui abbiamo discusso *supra* nel capitolo VI, vale a dire l'*algorithmic authority* e i suoi effetti su chi si interfaccia con sistemi algoritmici, in questo caso il giudice.

Ed è vero che tali problematiche sono sorte in uno specifico settore del processo penale statunitense, vale a dire la fase di *sentencing* in cui il giudice è chiamato a determinare in concreto la pena, ma data la generalità del fenomeno, per come analizzato nel capitolo dedicato, potremo astrarre le considerazioni anche oltre il processo penale e oltre la fase di determinazione della pena.

4.1. Autorità algoritmica nel giudizio penale nordamericano: la pretesa non decisività dell'algoritmo come garanzia di una sentenza individualizzata

Un filo rosso che intreccia e unisce le sentenze analizzate attiene al peso e alla rilevanza che l'algoritmo assume nella valutazione discrezionale del giudice e come esso incida sia sul diritto, riconosciuto come corollario del giusto processo, ad avere una sentenza individualizzata, sia in generale sull'esercizio abusivo della discrezionalità.

Quanto al primo aspetto, la doglianza nasce dal fatto che la predizione del rischio di recidiva del singolo imputato non si fonda sulle caratteristiche individuali del soggetto condannato, quanto piuttosto su dati aggregati di gruppi di persone che costituiscono il *dataset* dell'algoritmo decisorio. Come già spiegato analizzando il funzionamento di COMPAS (si v. *supra* cap. I § 4) ciò dipende dal fatto che la valutazione del rischio non è mai una predizione individualizzata, ma si basa sull'elaborazioni di dati afferenti a gruppi di individui che hanno caratteristiche simili a quelle del soggetto sottoposto a procedimento penale. Quanto al secondo aspetto, i giudici nella maggior parte dei casi ritengono che non ci sia stato un esercizio abusivo della discrezionalità in quanto l'algoritmo era solo uno dei tanti elementi che il giudice doveva considerare per determinare la pena.

I primi giudizi sembrano riconoscere le storture dell'utilizzo dei *risk tools* nel *sentencing*. In *Cardwell e Rhodes* l'algoritmo viene dipinto come uno strumento «inevitabilmente ingiusto» in quanto privo di capacità discretiva: l'utilizzo del risultato algoritmico contrasta con la intima funzione del giudice penale che, discrezionalmente, deve poter valutare i fattori di pericolosità del soggetto sottoposto al proprio giudizio sulla cui base formulare una sentenza appropriata e individualizzata.

Nel giro di pochi anni, invece, la giurisprudenza che matura all'interno dell'Indiana e del Wisconsin, seguita di recente dalla giurisprudenza del Michigan, inizia a distinguere tra *reliance* e *consideration* del risultato standardizzato prodotto dall'algoritmo decisorio: se vi è unanime giudizio sull'impossibilità per il giudice di appiattirsi (*rely on*) esclusivamente sul risultato algoritmico quale unico fattore

rilevante per la decisione, quest'ultimo diviene uno degli elementi che il giudice ha la facoltà sia di considerare per addivenire a una sentenza individualizzata (si cfr. *Malenchik, Loomis* e *People v. Younglove*), sia di trascurare e dare peso a elementi ulteriori e diversi (si cfr. *Samsa* e *J.S.*).

Come abbiamo notato, infatti, la *ratio decidendi* che accomuna le sentenze che rifiutano le prospettazioni delle difese secondo cui l'affidamento al risultato algoritmico privi l'imputato del suo diritto a una sentenza individualizzata risiede nella valutazione che l'algoritmo sia stato solo uno dei molti fattori che il giudice ha considerato per determinare l'ammontare la pena e non, invece, l'unico fattore cui il giudice sia affidato.

Ora, l'obbligo del giudice di comminare pene individualizzate costituisce uno dei principi cardine del *due process*⁵⁹⁴, riconosciuto anche a livello federale. In una sentenza risalente, e ancora oggi a più riprese utilizzata come precedente rilevante, la Corte Suprema aveva avvertito che

A sentencing judge, however, is not confined to the narrow issue of guilt. His task within fixed statutory or constitutional limits is to determine the type and extent of punishment after the issue of guilt has been determined. Highly relevant – if not essential – to his selection of an appropriate sentence is the possession of the fullest information possible concerning the defendant's life and characteristics. And modern concepts individualizing punishment have made it all the more necessary that a sentencing judge not be denied an opportunity to obtain pertinent information... The belief no longer prevails that every offense in a like legal category calls for an identical punishment without regard to the past life and habits of a particular offender.

Ancora di recente un Giudice federale ha affermato che

⁵⁹⁴ M. BRENNER *et al*, *Constitutional Dimensions of Predictive Algorithms in Criminal Justice*, cit., p. 278.

*It has been uniform and constant in the federal judicial tradition for the sentencing judge to consider every convicted person as an individual and every case as a unique study in the human failings that sometimes mitigate, sometimes magnify, the crime and the punishment to ensue*⁵⁹⁵

Tale principio, come detto, è stato ripreso anche dalle giurisprudenze dei singoli stati ed è utile qua sottolineare come sia un principio consolidato anche nella giurisprudenza del Wisconsin che, richiamando i propri precedenti nel caso *Loomis*, prende in esame il caso *State v. Gallion*⁵⁹⁶, in cui la Corte Suprema statale, pur non riconoscendo una violazione del giusto processo, aveva rimarcato nuovamente la natura del diritto a una sentenza individualizzata.

In particolare, prendendo in considerazione il principio nel nuovo contesto delle linee guide da applicare nel *sentencing* – a cui abbiamo già fatto cenno al capitolo I – la Corte afferma:

*Although we anticipate less disparity with the advent of sentencing guidelines, that does not mean there is less of a need for the exercise of discretion. Individualized sentencing, after all, has long been a cornerstone to Wisconsin's criminal justice jurisprudence. No two convicted felons stand before the sentencing court on identical footing. . . . and no two cases will present identical factors*⁵⁹⁷.

Ciò precisato, come emerso dalle sentenze analizzate, l'utilizzo dei *tools* è stato visto da alcuni giudici come capace di financo di arricchire e completare il diritto qui in esame: in *Malenchik*, per esempio, la Corte aveva ritenuto che una predizione del rischio potesse rinforzare (*enhance*) la valutazione individuale del giudice rispetto alle conseguenze della pena e ai programmi eventualmente da selezionare per il soggetto condannato,

Altre Corti hanno invece riconosciuto apertamente che un eventuale affidamento sul risultato prodotto dall'algoritmo predittivo potrebbe sollevare frizioni con il

⁵⁹⁵ *United States v. Barnes*, 890 F.3d 910 (10th Cir. 2018).

⁵⁹⁶ *State v. Gallion*, 678 N.W.2d 197 (2004).

⁵⁹⁷ *Ivi*, § 48.

principio qui in esame: in questo senso è indicativo che in *Loomis* la Corte suprema riconosca apertamente che se l'algoritmo fosse stato elemento determinante (*determinative factor*) per l'applicazione della pena, ciò avrebbe comportato una possibile violazione del giusto processo «*regarding whether a defendant received an individualized sentence*»⁵⁹⁸.

Come si vede, da qualunque prospettiva si voglia guardare l'interpretazione dei giudici ruota intorno al peso che dovrebbe avere l'algoritmo all'interno dei tanti fattori che il giudice valuta per assumere la decisione.

Un primo punto che bisogna evidenziare per chiarire la questione è che la menzionata distinzione tra *reliance* e *consideration*, se per un verso può apparire chiara, suscita molteplici incertezze⁵⁹⁹.

La confusione sembra emergere in modo cristallino in una *concurring opinion* della Giudice Roggensack nel caso *Loomis*, in cui la stessa, pur condividendo la decisione della maggioranza rispetto alla legittimità dell'utilizzo dei *tools*, evidenzia che la Corte, in motivazione, ha erroneamente utilizzato in maniera intercambiabile i due termini (*reliance* e *consideration*), così rischiando di legittimare surrettiziamente anche decisioni che assumono le risultanze dell'algoritmo come fattore decisivo per la determinazione della pena. Afferma la Corte che:

[b]ecause at times the majority opinion interchangeably employs consider and rely when discussing a sentencing court's obligations and the COMPAS risk assessment tool, our decision could mistakenly be read as permitting reliance on COMPAS. Therefore, I write to clarify for the reader.

A conclusioni analoghe perviene, come visto, anche la Corte d'Appello nel caso *Gordon* (si v. *supra*) in cui i giudici affermano di non essere «persuasi» dalla differenza tra *reliance* e *consideration* in quanto risulta assai complicato valutare il

⁵⁹⁸ Si v. ancora *Loomis*, § 68.

⁵⁹⁹ D. KEHL - P. GUO - S. KESSLER, *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing*, cit., p. 20.

peso che il giudice del *sentencing* ha assegnato a ogni singolo fattore soggetto al suo giudizio.

Secondo la Corte, quindi, non è agevole individuare in che misura sia stata data rilevanza all'algoritmo (la Corte dice, efficacemente, «*mentally assigned*») e se esso sia stato decisivo rispetto alla decisione sulla pena.

Nella valutazione discrezionale del giudice sul *quantum* di pena, risulta in effetti arduo verificare quale peso sia stato effettivamente attribuito al risultato algoritmico, tanto più, come sovente accade, nel caso di algoritmi di *machine learning* che producono un indice di recidiva senza che risulti chiaro il procedimento che ha condotto alla produzione del risultato.

Quale che sia il risultato del sistema predittivo, se il giudice assume la propria decisione *considerando* il risultato dell'algoritmo e dunque conformando la sentenza a detto risultato, il confine con la *reliance* appare quanto mai labile, se non addirittura inesistente⁶⁰⁰.

La conclusione non muta neppure obiettando che in alcuni casi citati (si v., per esempio, *Samsa* e *J.S.*) il giudice non ha tenuto conto dell'indice di recidiva prodotto dalla macchina: e ciò sia poiché, come anche si è visto, sono casi rari rispetto a quelli in cui il giudice assume la decisione sulla base dell'algoritmo; sia avendo in mente l'impianto culturale del modello giudiziario americano, in cui le giurie sono normalmente colpevoliste e dove il *prosecutor* (il nostro pubblico ministero) possiede uno strapotere processuale, avendo nella propria disponibilità la polizia, le perizie medico legali, le prove, la possibilità di sentire i testimoni; in siffatto sistema, consegnare all'accusa un ulteriore elemento di prova (e di tale fascino, come l'algoritmo) rafforza ancora di più la disparità delle armi, ed è indicativo il fatto che i casi rinvenuti in cui il giudice ha trascurato il risultato algoritmico sono quelli in cui questi ignorava l'indice di recidiva basso o molto basso, applicando una pena di ammontare elevato (si cfr. ancora *Samsa* e *J.S.*)⁶⁰¹.

⁶⁰⁰ S. GOEL - R. SHROFF - J. L. SKEEM - C. SLOBOGIN, *The Accuracy, Equity, and Jurisprudence of Criminal Risk Assessment*, dicembre 2018.

⁶⁰¹ J. PATRONE, *Il prosecutor negli Stati Uniti. Un esempio da seguire?*, in *Questione Giustizia*, 11 marzo 2021.

Come si intuisce, quindi, i giudici americani solo raramente hanno considerato la forza dell'effetto ancoraggio, di cui abbiamo discusso *supra*. A ben vedere, infatti, pur quasi sempre determinando la pena attribuendo il peso decisivo al risultato algoritmico, hanno ritenuto non determinante il peso dell'algoritmo nella decisione.

L'algoritmo raccoglie la quasi totalità degli elementi della decisione e, per di più, non fornisce una predizione del rischio individualizzata ma basata sull'elaborazione di dati aggregati che formano dei *pattern* all'interno dei quali il soggetto sottoposto a giudizio viene classificato (si v. ancora, Guida COMPAS, *supra* capitolo I § 4).

Non è quindi revocabile in dubbio che la valutazione individualizzata di un imputato da parte del giudice venga sacrificata. Come indicato nel titolo della premessa al capitolo, in questo caso proprio l'esercizio abusivo della discrezionalità, che si appiattisce sul risultato del *tool* algoritmico, è prodromico alla violazione di una delle garanzie del giusto processo, il diritto e l'obbligo di una sentenza automatizzata.

Ciò consente di sostenere che l'esercizio della discrezionalità, a fronte del peso dell'algoritmo, subisce forti pressioni. Esso, spesso non riscontrato dai giudici che si sono rifugiati nella evanescente distinzione tra decisività e complementarità dell'algoritmo, disvela invero il mutamento dell'esercizio discrezionale del potere dinanzi ai sistemi di predizione del rischio e trova un suo primo punto di ricaduta nella impossibilità quasi immanente di distinguere tra un obbligo – affidarsi integralmente a un risultato automatizzato – e, all'opposto, una mera facoltà di attenervisi.

4.2. L'influenza dell'autorità algoritmica nelle altre ipotesi di giustizia algoritmica

Abbiamo finora analizzato uno specifico campo in cui il pregiudizio da automazione si è manifestato nella pratica, vale a dire la giurisprudenza americana.

Ciò nonostante, avendo delineato l'impianto teorico dell'*algorithmic authority* e avendone messo in luce una manifestazione evidente in uno specifico campo, è forse possibile azzardare l'estensione di tale effetto anche alle altre ipotesi di giustizia predittiva oggi in fase di sperimentazione in Europa e ai possibili sviluppi della giustizia predittiva, delineati nei capitoli precedenti.

Pensiamo per esempio alle ipotesi di giustizia predittiva, in cui i giudici si trovano a dover decidere la controversia anche sulla base di strumenti algoritmici in grado di predire (*rectius* prevedere) il risultato della controversia.

Già nella Carta Etica approvata dalla CEPEJ si afferma che si corre il rischio che – senza alcune cautele necessarie, di cui discuteremo nell'ultimo capitolo – «*[t]he effect of these tools may be not only to provide incentives but to be almost prescriptive, creating a new form of normativity, which could supplement the law by regulating the sovereign discretion of the judge*». Pur senza nominarlo (mentre invece lo nomina esplicitamente, come visto, l'*AI Act*), la Carta Etica sembra ben consapevole del rischio dell'*effetto ancoraggio* sulla discrezionalità del giudice.

Quali dunque le conseguenze di un appiattimento della decisione discrezionale del giudice su strumenti di giustizia predittiva?

Già la stessa Carta Etica indica quale possibile conseguenza la possibilità che la percezione dell'effetto prescrittivo di strumenti di giustizia predittiva potenzialmente potrebbe condurre a una «*standardisation of judicial decisions based no longer on case-by-case reasoning by the courts, but on a pure statistical calculation linked to the average compensation previously awarded by other courts*».

Ora, senza qui mettere in dubbio la sempre maggiore rilevanza dei precedenti giurisprudenziali anche nei sistemi con una tradizione di *civil law* come quello italiano e come la maggior parte degli ordinamenti dell'Europa continentale⁶⁰², appaiono diverse le situazioni in cui da una parte il giudice, con gli strumenti tradizionali, esegue la propria ricerca autonoma sui precedenti rilevanti per decidere e dall'altra, il giudice si affida a un algoritmo di *machine learning* (non da lui

⁶⁰² Si cfr. i contributi pubblicati in A. CARLEO (a cura di), *Il vincolo giudiziale del passato*, Bologna 2018.

prodotto né alimentato, vedremo con quali conseguenze) che gli suggerisce l'esito della controversia.

Tali tematiche sono tutt'ora al centro della letteratura europea e italiana e riflettono sostanzialmente le preoccupazioni manifestate in tempi non sospetti – in cui l'IA non era ancora nemmeno all'orizzonte – da Roscoe Pound, giurista statunitense che segnalava già agli inizi del '900 i rischi di basarsi solo sulla giurisprudenza passata, criticando quella che lui definiva “giurisprudenza meccanica”⁶⁰³. Qualche anno più tardi lo stesso Pound in un suo scritto successivo scriveva: «*law must be stable and yet it cannot stand still*»⁶⁰⁴.

In Europa la dottrina francese ha parlato di «*effet moutonnier*»⁶⁰⁵, traducibile in italiano con «effetto gregge»: la tendenza dell'algoritmo a catturare la decisione del giudice si tradurrà in un inesorabile impoverimento dei giudici, che preferiranno adattarsi a ciò che il sistema suggerisce piuttosto che impegnarsi nel caso e sviluppare nuove soluzioni.

In Italia, il problema del conformismo giudiziario ha trovato spazio anche all'interno della magistratura: nella recente relazione sullo stato della giustizia telematica, il CSM, segnalando le potenzialità e i rischi dell'implementazioni di banche dati pubbliche con lo scopo di alimentare sistemi predittivi di IA ha precisato che «mentre un tale sistema di conoscibilità delle decisioni per un verso appare certamente funzionale alla costruzione di un pensiero attivo e consapevole, anche nell'ottica della prevedibilità delle decisioni e del controllo di coerenza da

⁶⁰³ R. POUND, *Mechanical Jurisprudence*, in *Columbia Law Review*, 8 605-623, 1908.

⁶⁰⁴ R. POUND, *Interpretations of legal history*, The Macmillan Company, New York, 1923, p. 1.

⁶⁰⁵ A. GARAPON, J. LASSÈGUE, *Justice digitale. Révolution graphique et rupture anthropologique*, cit. Rispetto all'impiego di strumenti di giustizia predittiva gli A. affermano che: «*C'est l'effet moutonnier de la justice prédictive: elle pousse au conformisme et réclame plus d'indépendance d'esprit aux juges qui estiment qu'ils doivent aller à contre-courant, c'est-à-dire qui veulent simplement faire leur métier*». In un'intervista rilasciata a *Questione Giustizia* a cura di Emanuela Fronza e Corrado Caruso Garapon spiega così le ragioni del possibile effetto gregge: «I giudici si troveranno ad essere sotto pressione per via della doppia lettura di questa triste radiografia delle istituzioni. Per forza di cose. Ad esempio, una start up francese propone di dire al giudice, a fronte di un caso concreto, come 100 dei suoi colleghi deciderebbero quel caso. Ciò può rendere il giudice soggetto all'opinione della maggioranza, nel caso in cui sappia, ad esempio, che il 70% dei suoi colleghi ha deciso in una data maniera. Ciò può portare il giudice a seguire la maggioranza come pecore (*effet moutonnier*), con un rischio di deresponsabilizzazione dello stesso». Si v. E. FRONZA - C. CARUSO (a cura di), *Ti faresti giudicare da un algoritmo. Intervista ad Antoine Garapon*, in *Questione Giustizia*, trimestrale, n. 4, 2018, pp. 189-199.

parte dei cittadini del servizio Giustizia, per altro verso, potrebbe anche favorire effetti potenzialmente negativi, facilitando, per *mero produttivismo*, la passiva ricezione di apparati motivazionali, con conseguente *conformismo giudiziario* ed *impoverimento della stessa qualità della giurisdizione* (corsivi nostri)⁶⁰⁶.

Anche la dottrina costituzionalistica italiana ha sottolineato tali pericoli⁶⁰⁷, segnalando altresì la possibilità di un tale conformismo possa comportare il sacrificio di uno dei principi chiave della giurisdizione, la soggezione del giudice soltanto alla legge previsto dall'art. 101, secondo comma, Cost⁶⁰⁸.

A questo riguardo ci sembra di poter dire che, come la legge ha negli ultimi tempi valorizzato la considerazione dei precedenti, soprattutto quelli della Cassazione, è la stessa legge che potrebbe prevedere la possibilità, da parte del giudice, di considerare nella decisione anche il risultato di un algoritmo di giustizia predittiva, così salvaguardando il principio dell'art. 101 Cost. Certo è che se, come abbiamo visto, anche considerando l'insegnamento della giurisprudenza americana, l'indicazione dell'algoritmo diventa una *prescrizione* quasi obbligata per il giudice, il principio della soggezione esclusiva alla legge ne potrebbe soffrire. Anche qui le soluzioni, sia tecniche che giuridiche non mancano, ma verranno discusse nel capitolo conclusivo dopo aver segnalato tutte le criticità che la giustizia algoritmica ha dimostrato di provocare rispetto alle garanzie costituzionali.

Tutt'altro caso sarebbe se la stessa legge imponesse l'obbligo del giudice di conformarsi sempre al risultato algoritmico. La Carta etica afferma in proposito che nel caso in cui si fornisse a un giudice, tramite algoritmi, il contenuto delle decisioni prodotte da tutti gli altri giudici, sarebbe totalmente da rifiutare il fatto di «pretendere che la sua scelta futura debba essere necessariamente compresa nella

⁶⁰⁶ CSM, Relazione sullo stato della Giustizia Telematica 2024, già citata *supra* alle note 213 p. 54.

⁶⁰⁷ Si cfr., *ex multis*, C. CASONATO, *Intelligenza artificiale e giustizia: potenzialità e rischi*, in *DPCE online*, 2020/3, p. 3388; C. SALAZAR, *Judex ex machina? Note su giustizia, giudici e intelligenza artificiale*, in *ConsultaOnline*, 18 novembre 2021, p. 927.

⁶⁰⁸ SALAZAR, *Judex ex machina? Note su giustizia, giudici e intelligenza artificiale*, cit.; M. LUCIANI, *La decisione giudiziaria robotica*, cit., p. 889.

massa di tali “precedenti” ... in quanto tale numero elevato non può costituire un’aggiunta alla legge né può sostituirsi a essa»⁶⁰⁹.

Terminando il discorso sull’autorità dell’algoritmo, nel nostro approccio induttivo abbiamo altresì segnalato alcune ipotesi futuribili (e in alcuni casi già accadute) di impiego generalizzato di strumenti di intelligenza artificiale generativa all’interno del processo. Non abbiamo qui né dottrina né giurisprudenza in grado di illuminare il discorso: quello che ci sembra, tuttavia, è che anche in quei casi si può ritenere che l’impiego di ChatGPT, senza le necessarie guarentigie, potrebbe divenire uno strumento che influisce enormemente sulla decisione del giudice, catturando la decisione. Ciò è anche desumibile dalla significativa affermazione del giudice inglese che ha ammesso di aver utilizzato lo strumento per la risoluzione di un processo (si v. *supra* cap. III § 2): egli era già giunto a una decisione del caso e aveva sperimentato ChatGPT per verificarne l’attendibilità *ex post*, ma aveva significativamente messo in guardia sull’utilizzo di tali strumenti per decidere la causa, sia ovviamente per la nota idoneità a commettere errori, sia probabilmente per la tendenza del giudice, una volta utilizzati, ad assumerne il risultato come esatto, senza verificarne l’affidabilità.

In definitiva, si può affermare che l’algoritmo, di qualunque tipo esso sia, tende e influire in modo decisivo sulla decisione discrezionale del giudice. I casi americani hanno messo bene in luce tale fenomeno, anche quando i giudici hanno negato tale influenza affermando di aver solo considerato l’algoritmo tra una serie di fattori, evidentemente ignorando il fenomeno dell’*automation bias* che abbiamo spiegato nel capitolo VI.

Ora, se il peso dell’algoritmo diviene determinante nella decisione giudiziaria, in assenza di un’adeguata regolamentazione forse si correrebbe il rischio di assistere a un passaggio di potere. E non, come spesso frettolosamente si adduce, un passaggio da una intelligenza umana (il giudice) a una artificiale (l’algoritmo), bensì da una umana (pubblica) e una altrettanto umana (ad oggi quasi sempre privata), che trascina con sé, come si vedrà, notevoli e ulteriori preoccupazioni.

⁶⁰⁹ Carta Etica, cit., p. 44 (traduzione nostra).

5. Opacità dell’algoritmo ed esternalizzazione della funzione giudicante

Come si è cercato di spiegare nel capitolo VI, l’opacità algoritmica attraversa ogni discorso sulla compatibilità degli algoritmi nel procedimento giurisdizionale ed è di conseguenza alla base della normativa di riferimento della giustizia algoritmica. In questa sezione, tuttavia, si sta discutendo, prima di affrontare nel dettaglio le ricadute sulle garanzie costituzionali, di come influiscono le caratteristiche degli algoritmi sulla decisione del giudice. È quindi sotto questa lente che qui prendiamo in considerazione l’opacità degli algoritmi.

Anche qui occorre prendere esempio dalla giurisprudenza americana. Come affermato dalla difesa nel caso *People v. Younglove* (si v. *supra* § 3),

*using COMPAS to make a sentencing decision risk transferring
sentencing discretion from judges to software developers*

Una delle contestazioni sollevate dagli imputati in *People v. Younglove* concerneva proprio il rischio che la discrezionalità nella decisione da assumere nel *sentencing* si «trasferisse» dal giudice al progettista del *software* utilizzato nel processo. Nonostante i giudici di appello non prendano seriamente tale doglianza, se non con affermazioni che ricalcano le distinzioni che abbiamo evidenziato nei precedenti paragrafi tra decisività e non decisività dell’algoritmo, essa ci consente di introdurre un punto decisivo della riflessione fin qui svolta.

Come abbiamo visto la progettazione degli algoritmi oggi è riservata non all’amministrazione pubblica (o comunque mai all’amministrazione giudiziaria), ma viene invece, come visto, *esternalizzata* a soggetti privati⁶¹⁰. Anche se ci si arrestasse qui, ci si potrebbe chiedere se già tale fenomeno possa ritenersi

⁶¹⁰ Si cfr. D.K. CITRON, *Technological Due Process*, cit., p. 1296, in cui l’A. afferma che «[a]utomation produces a new delegation of ... power to automated systems and their designers – a delegation that needs to be addressed». Si v. anche *mutatis mutandis*, con riguardo all’amministrazione per algoritmi, A. DI MARTINO, *L’amministrazione per algoritmi ed i pericoli del cambiamento in atto*, in *Il diritto dell’economia*, vol. 3, 2020, pp. 627 ss.

compatibile con i principi dello stato di diritto e della legittimazione democratica, seppure peculiare, della magistratura.

Ma il fenomeno della esternalizzazione della progettazione di algoritmi al di fuori del perimetro della magistratura si lega a doppio filo con un secondo elemento che abbiamo analizzato, vale a dire l'opacità algoritmica.

Muovendo ancora dall'insegnamento della giurisprudenza americana, il Giudice supremo dello Iowa, con notevole enfasi, si chiede: «*How are we to determine the due process implications of the district court's use of risk assessment tools, when we do not know anything about the tools and Gordon failed to object to their use?*»

Tali preoccupazioni inducono a indugiare sulla trasparenza non solo dal punto di vista dei diritti dell'imputato ad essa connessi – su cui ci soffermeremo in seguito – ma dalla prospettiva del giudicante che deve “fare i conti” con sistemi algoritmici di cui non conosce pressoché nulla e che fondano (parzialmente) il suo giudizio.

Se assumiamo che l'algoritmo assume un peso preponderante o financo decisivo nella decisione del giudice, ma tale algoritmo non è controllabile perché segreto o incomprensibile, le implicazioni vanno oltre la semplice esternalizzazione della progettazione algoritmica.

In effetti, se giustapponiamo il fenomeno dell'opacità con il fatto che l'impiego dell'algoritmo tende ad assorbire la stessa decisione discrezionale del giudice, il quale aderisce, anche non del tutto coscientemente, a quanto l'algoritmo predittivo gli “suggerisce”, non sembrerebbe azzardato asserire che l'esternalizzazione della progettazione dell'algoritmo si traduce, a ben vedere, in una parziale esternalizzazione, quasi sempre a un privato, della stessa funzione giudicante.

A questo riguardo, l'esternalizzazione del potere discrezionale, nell'accezione che qui si è sostenuta, non è che una conseguenza del mancato controllo che il giudice dovrebbe poter esercitare sull'algoritmo predittivo, in quanto sostanzialmente inaccessibile; detto altrimenti, l'esercizio del potere discrezionale non può venire correttamente esercitato se il giudice non ha piena contezza della “fonte” da cui la sua decisione promana, come accade, appunto, di fronte ad algoritmi opachi: si assiste dunque a una “relazione biunivoca” tra esternalizzazione

e opacità, dove quest'ultima da un lato è una concausa della esternalizzazione e a sua volta ne aggrava gli effetti.

Come provocatoriamente affermato da alcuni Autori: «*who are the decision-makers – the government officials involved or software programs they are relying on, which are designed by companies for profit?*»⁶¹¹ .

L'attribuzione formale del potere in capo al giudice penale diventa quindi un guscio vuoto che fatica a confinare una decisione assunta, di fatto, da chi ha progettato il *software*. Come correttamente affermato riguardo alla sentenza emessa nel caso *Loomis*, «*the Court in effect outsourced its decision-making, at least to some extent, to algorithms that are questionable in quality and insensitive to the fundamental norms in the US legal system, consequently undermining its public accountability*»⁶¹² e, soprattutto, «*[w]ithout understanding and addressing the 'black box' challenge and its legal ramifications, the Loomis decision impedes the role of the judiciary as the final gatekeeper to protect individuals' rights and to provide effective remedies*»⁶¹³.

SEZIONE III. LE RICADUTE SULLE GARANZIE DEL GIUSTO PROCESSO

Una volta verificati e dimostrati i vari livelli di influenza e impatto che gli algoritmi hanno sul processo decisionale del giudice e sul suo potere discrezionale occorre verificare come questa influenza si riflette sulle garanzie del giusto processo.

Tutto quello che abbiamo ricostruito finora, dalle premesse introduttive, alle indicazioni esperienze in corso, all'enucleazione dei tentativi di regolamentazione dell'IA e della giustizia algoritmica fino alle modalità di influenza sulla decisione del giudice è finalizzato a indagare, ora con la consapevolezza necessaria, le ricadute sul processo, su quella parte del sistema costituzionale che oggi ancor più di ieri appare decisiva per la tutela dei diritti fondamentali.

⁶¹¹ H. LIU - C. LIN - Y. CHEN, *Beyond State v Loomis: artificial intelligence, government algorithmization and accountability*, cit., p. 137.

⁶¹² *Ibidem*

⁶¹³ *Ibidem*, p. 135.

Questa sezione, quindi, costituisce la parte centrale del lavoro, e cercherà di unire i punti di quanto finora scritto. Potrebbe apparire alla lettura una visione particolarmente pessimistica sulla compatibilità della giustizia algoritmica con le garanzie del giusto processo: in realtà abbiamo dedicato questa sezione autonoma a cercare di capire quali sono oggi le preoccupazioni costituzionali più pressanti, in astratto e in concreto, dedicando il capitolo finale a domandarci quali soluzioni e quali guarentigie occorre applicare per guidare tale cambiamento (alcune delle quali sono già state implementate dai legislatori).

Ci occuperemo, quindi, delle garanzie che più appaiono implicate dalla giustizia algoritmica, sia quelle istituzionali sia quelle procedurali (seguendo la classificazione della Corte europea dei diritti dell'Uomo).

6. La giustizia algoritmica alla prova dell'indipendenza e dell'imparzialità del giudice

Consapevoli che indipendenza e imparzialità sono concetti distinti, ai fini del presente capitolo appare più opportuno trattarli congiuntamente. È la stessa Corte europea dei diritti ad affermare a più riprese che i concetti di indipendenza e imparzialità del giudice «*are closely linked and, depending on the circumstances, may require joint examination*»⁶¹⁴ e che anzi l'indipendenza e l'imparzialità delle Corti è un prerequisito per la garanzia della *rule of law* e una *conditio sine qua non* per garantire il giusto processo ai sensi dell'art. 6 CEDU⁶¹⁵.

Inoltre, per le implicazioni che hanno le applicazioni algoritmiche rispetto a tali principi (in particolare chi progetta gli algoritmi e quali dati e regole vengono in essi inseriti), si è affermato che nel contesto delle applicazioni di IA nel processo, «imparzialità e indipendenza si equivalgono»⁶¹⁶.

⁶¹⁴ Corte Edu, *Ramos Nunes de Carvalho e Sá v. Portugal* [GC], 2018, §§ 150 e 152. Anche nella giurisprudenza dell'art. 47 CDFUE spesso i due requisiti sono trattati congiuntamente dalla Corte di giustizia dell'Unione europea.

⁶¹⁵ Corte Edu, *Grzęda v. Poland* [GC], 2022, §§ 298 e 301.

⁶¹⁶ J. NIEVA-FENOLL, *Intelligenza artificiale e processo*, cit., pp. 121-122.

L'indagine ha messo in luce l'importanza di questi due principi nel nuovo contesto della giustizia algoritmica: nella giurisprudenza americana, le doglianze hanno spesso riguardato l'abuso di discrezionalità, da cui discendevano decisioni arbitrarie, e per ciò non imparziali. Inoltre, quasi tutte le esperienze normative (compreso, da ultimo, *l'AI Act*), hanno menzionato l'indipendenza e l'imparzialità come principi irrinunciabili e da salvaguardare nel nuovo contesto algoritmico. Significativamente, il primo principio della Carta etica, il rispetto dei diritti fondamentali (si v. *supra* capitolo V § 6), richiede che gli strumenti di intelligenza artificiale «*should also be used with due respect for the principles of the rule of law and judges' independence in their decision-making process*»⁶¹⁷. Nella stessa Carta etica si specifica che tali strumenti potrebbero provocare «*effects on the independence and impartiality of the judiciary, particularly in systems where the independence of the judiciary is not fully achieved. In these systems, we cannot rule out the risk that such norms will place indirect pressure on judges when decisions are taken and prompt their approval, or that the executive will monitor those who depart from the norm*»⁶¹⁸.

Ora, nonostante il trattamento congiunto di imparzialità e indipendenza del giudice, è utile premettere un breve cappello introduttivo sull'impianto multilivello di tali garanzie.

A livello sovranazionale tali principi sono previsti dagli art. 6 § 1 CEDU e dall'art. 47 § 2 CDFUE – che hanno una formulazione pressoché analoga⁶¹⁹ – i quali stabiliscono che «*everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal established by law*».

Nonostante la stretta connessione che spesso caratterizza i due principi nella giurisprudenza della Corte europea dei diritti (e nella giurisprudenza della Corte di Giustizia dell'Unione europea), il termine “*independent*” si riferisce principalmente all'indipendenza nei confronti degli altri poteri (in particolare il potere legislativo

⁶¹⁷ Carta Etica, cit., p. 8.

⁶¹⁸ Carta Etica, p. 48.

⁶¹⁹ Nell'art. 6 della Convenzione si aggiunge che tali principi valgono «*In the determination of his civil rights and obligations or of any criminal charge against him*» (art. 6, § 1) CEDU.

e il potere giudiziario, cd. indipendenza esterna)⁶²⁰ e nei confronti dei colleghi (cd. indipendenza interna)⁶²¹. Tra i criteri che la Corte utilizza per verificare l'indipendenza dei giudici compaiono, tra gli altri – e come vedremo sarà interessante nel nostro discorso – l'esistenza di garanzie contro pressioni ricevute dall'esterno e se l'organo presenta nel suo complesso un'apparenza di indipendenza⁶²².

L'imparzialità invece denota per la Corte europea dei diritti «*the absence of prejudice or bias*» nei confronti delle parti in causa⁶²³. Come noto, la verifica dell'imparzialità deve seguire due differenti *test*, uno *soggettivo*, avuto riguardo delle convinzioni personali del singolo giudice; e uno *oggettivo*, spesso sovrapponibile al concetto di indipendenza, che richiede di accertare se l'organo giudiziario nel suo complesso (e la sua composizione) presenta sufficienti garanzie per escludere dubbi sull'imparzialità dei giudici⁶²⁴. Il requisito oggettivo ha quindi a che vedere con l'apparenza di imparzialità (e di indipendenza) che deve circondare l'esercizio della funzione giurisdizionale: secondo le parole della Corte: «*justice must not only be done, it must also be seen to be done. What is at stake is the confidence which the courts in a democratic society must inspire in the public*»⁶²⁵.

A livello nazionale il principio di indipendenza esterna è previsto dagli art. 104, primo comma, Cost., secondo cui «La magistratura costituisce un ordine autonomo e indipendente da ogni altro potere» e dall'art. 101, secondo comma, Cost. ai sensi del quale «i giudici sono soggetti soltanto alla legge». Già dalle prime sentenze la Corte costituzionale, nell'interpretare il combinato degli artt. 101 e 104 Cost., ha affermato che mentre l'art. 101, enunciando il principio della indipendenza del singolo giudice, ha inteso indicare che questi nell'esercizio della sua funzione non

⁶²⁰ Corte EDU, *Beaumartin v. France*, 1994, § 38.

⁶²¹ Corte EDU, *Agrokompleks v. Ukraine*, 2011, § 137.

⁶²² Corte EDU, *Kleyn and Others v. the Netherlands* [GC], 2003, § 190.

⁶²³ Corte EDU, *Micallef v. Malta* [GC], 2009, § 93.

⁶²⁴ Tali test sono stati indicati in diverse sentenze della Grande Camera tra cui si cfr. Corte EDU *Morice v. France* [GC], 2015, §§ 73-78; e *Denisov v. Ukraine* [GC], 2018, §§ 61-65.

⁶²⁵ Corte EDU, *Micallef v. Malta* [GC], 2009, § 98. Sul tema si cfr. B. RANDAZZO, *Imparzialità del giudice e fiducia nella magistratura nella prospettiva costituzionale europea*, cit.

ha altro vincolo che quello della legge, l'art. 104, intende, invece, porre il principio della indipendenza della organizzazione giudiziaria nel suo complesso, nel senso che l'ordine della magistratura non deve dipendere da altro potere⁶²⁶e, in particolare, ovviamente, da organi politici⁶²⁷. Tale principio, come noto, è strettamente connesso al principio della separazione dei poteri⁶²⁸ ed è salvaguardato da una serie di disposizioni costituzionali e legislative concernenti il reclutamento dei magistrati, la loro formazione, l'istituzione di un organo di autogoverno ecc...⁶²⁹. Ai nostri fini assume particolare importanza la cd. indipendenza esterna funzionale, che si fonda sull'art. 101, secondo comma, Cost., e che implica un'autonoma potestà decisionale del singolo giudice, che non deve essere condizionata da alcuna interferenza proveniente da autorità di governo o altri soggetti: è solo la legge⁶³⁰ che può indicare al giudice le regole da applicare al giudizio⁶³¹. A ciò strettamente connesso è dunque la garanzia del libero convincimento del giudice nella valutazione dei fatti e delle prove rilevanti in un giudizio, che l'ordinamento deve in ogni caso assicurare⁶³².

Come già visto, l'indipendenza è inoltre *strumento* per garantire l'imparzialità del giudice, vale a dire ne costituisce il *presupposto*⁶³³. Senza qui ricostruire le numerose indagini dottrinali sul fondamento costituzionale dell'imparzialità, tale principio non era esplicitamente previsto nell'articolato originario della Costituzione, ma era ricavato dagli artt. 3 e 24 Cost, e quindi nel principio di uguaglianza di fronte alla legge e nel diritto di difesa. La diffusione poi della

⁶²⁶ Corte Cost., sent. n. 22 del 1959.

⁶²⁷ N. ZANON - F. BIONDI, *Il sistema costituzionale della magistratura*, Torino, 2024, p. 103

⁶²⁸ Si v., *ex multis*, N. ZANON - F. BIONDI, *Il sistema costituzionale della magistratura*, cit. p. 96 e G. FERRI, *Autonomia e indipendenza della magistratura tra 'vecchio' e 'nuovo' ordinamento giudiziario*, in *Rivista AIC*, n. 4, 2017, pp. 1-52.

⁶²⁹ In generale sul tema dell'indipendenza della magistratura si v. ancora N. ZANON - F. BIONDI, *Il sistema costituzionale della magistratura*, cit.

⁶³⁰ Sull'interpretazione da dare al termine legge si v., tra i molti, N. ZANON - L. PANZERI, *Art 101*, in R. BIFULCO - A. CELOTTO - M. OLIVETTI (a cura di), *Commentario alla Costituzione*, Utet giuridica, Torino, 2006, p. 1961.

⁶³¹ Si v. *ex multis*, Corte cost., sent. n. 8 del 1962.

⁶³² Sul libero convincimento nella valutazione delle prove si v. ad esempio, Corte cost, sent. n. 255 del 1992 in cui la Corte ha definito il libero convincimento come «libertà del giudice di valutare la prova secondo il proprio prudente apprezzamento, con l'obbligo di dare conto in motivazione dei criteri adottati e dei risultati conseguiti» (punto n. 3 del *Considerato in diritto*).

⁶³³ Così, *ex multis*, Corte cost., sent. n. 128 del 1974.

nozione di “giusto processo”, derivata soprattutto dalla giurisprudenza della Corte EDU ha dapprima indotto la Corte costituzionale a ricondurre l'imparzialità alla più ampia formula del “giusto processo”⁶³⁴ e poi il legislatore costituzionale a introdurre, mediante l. cost. 2 del 1999, all'art. 111 cinque nuovi commi, i cui primi due oggi prevedono che «La giurisdizione si attua mediante il giusto processo regolato dalla legge»⁶³⁵ e che «Ogni processo si svolge nel contraddittorio tra le parti, in condizioni di parità, davanti a giudice terzo e imparziale. La legge ne assicura la ragionevole durata».

Ora, tali coordinate generali, il cui approfondimento ulteriore esula dalle finalità del presente scritto, sono utili al fine di calare i principi espressi nel contesto del giusto processo europeo alla nuova (*rectius* alle nuove) realtà della giustizia algoritmica⁶³⁶.

Dobbiamo ovviamente considerare le esperienze che abbiamo analiticamente considerato nei primi tre capitoli e i problemi comuni che abbiamo segnalato nel capitolo VI.

Le esperienze più rilevanti di giustizia algoritmica (soprattutto l'esperienza americana e quella francese) hanno dimostrato che la digitalizzazione prima, e l'impiego dell'IA poi, producono un'attrazione di alcune dinamiche processuali al di fuori dalla disponibilità della organizzazione giudiziaria e il loro affidamento, spesso, all'esecutivo che a sua volta si avvale di società private per progettare gli algoritmi (sono i dipartimenti di giustizia negli Stati Uniti a stipulare i contratti con le società private per lo sviluppo dei *risk tools* ed è stato il Ministero della Giustizia in Francia che ha contratto la licenza di un *software predittivo* da un privato per poi fornirlo ai magistrati). L'affidamento della progettazione degli algoritmi e della scelta dei dati al di fuori del potere giudiziario (quella che qui abbiamo chiamato

⁶³⁴ Si v. *ex multis*, Corte cost., sentt. nn. 131 e 177 del 1996.

⁶³⁵ Si cfr. A. ANDRONIO, *Art. 111*, in R. BIFULCO - A. CELOTTO - M. OLIVETTI (a cura di), *Commentario alla Costituzione*, Utet giuridica, Torino, 2006, pp. 2111 ss.

⁶³⁶ Sul tema si v. ancora A. PAJNO, *Intelligenza artificiale e autonomia del giudice. il ruolo del fatto e il valore del precedente*, cit.

“esternalizzazione”, si v. *supra* § 5) «si tradurrebbe in una questione di potere ... e si rischierebbe di passare al “magistrato bocca del governo”»⁶³⁷.

Ora, se associamo i problemi legati all'*automation bias* e alla trasparenza che comportano l'*esternalizzazione* di parte della funzione giurisdizionale al nuovo contesto algoritmico in cui è l'organo esecutivo che governa e gestisce l'impiego degli algoritmi nel processo, senza controlli da parte dell'organo giudiziario, è evidente che il principio di indipendenza ne viene compromesso⁶³⁸. Se a ciò si aggiunge che, come visto, il ruolo dei privati è oggi preponderante nella progettazione dei sistemi algoritmici, che oppongono, come visto negli Stati Uniti, segreti commerciali alle richieste di accesso all'algoritmo, ecco che il giudice si ritrova ad assumere una decisione basata su influenze che derivano da soggetti *estranei* all'organo giudiziario e che per di più egli non ha la possibilità di sindacare⁶³⁹. Se quindi gli strumenti di supporto finiscono per costituire, come visto, un'influenza decisiva sulla decisione del giudice, e in ultimo sull'espletamento della funzione giurisdizionale, viene meno il principio per cui la giurisdizione deve essere libera da fattori esterni secondo il principio di indipendenza⁶⁴⁰.

In definitiva, se l'attuazione delle riforme tecnologiche e dell'intelligenza artificiale rimane responsabilità dell'esecutivo, l'autonomia decisionale dei giudici ne verrebbe profondamente pregiudicata. Il *design* algoritmico potrebbe compromettere tale indipendenza se alla magistratura non fosse affidato un contributo informato ed efficace, né una supervisione sul suo funzionamento (di queste soluzioni si discuterà nell'ultimo capitolo).

⁶³⁷ A. FORMISANO, *L'impatto dell'intelligenza artificiale in ambito giudiziario sui diritti fondamentali*, in *Federalismi*, n. 22/2024, pp. 112-149, spec. p. 121.

⁶³⁸ Così E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit., p. 330. In Italia, salvo alcuni esperimenti che coinvolgono il Ministero della giustizia, esistono, come visto, esperienze in cui i protagonisti rimangono all'interno del sistema giudiziario, anche se ancora sono lontane dall'idea di impiegare l'IA nel processo decisionale del giudice (si v. *supra* capitolo II).

⁶³⁹ M. ZALNIERUTE - F. BELL, *Technology and the Judicial role*, in G. APPLEBY - A. LYNCH, *The Judge, the Judiciary and the Court: Individual, Collegial and Institutional Judicial Dynamics in Australia*, Cambridge University Press, 2021, pp. 116-142, spec. pp. 130 ss. Sul tema si v. anche T. SOURDIN, *What if Judges Were Replaced by AI?*, in *Turkish Policy Quarterly*, 2022.

⁶⁴⁰ E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit., p. 33.

Tali considerazioni si riflettono poi sul principio di imparzialità del giudice, proprio perché il giudice si ritrova a impiegare algoritmi di cui non conosce il funzionamento e nemmeno l'elaborazione dei dati al suo interno. Come abbiamo visto (capitolo VI), il controllo e la *governance* dei dati sono decisivi per il corretto impiego degli algoritmi nel campo della giustizia: nell'esempio dei *tools* i dati inseriti nei sistemi, già riflettenti un contesto carcerario discriminatorio, producevano altrettanti risultati discriminatori (*garbage in, garbage out*). Se il processo decisionale del giudice, quindi, si appiattisce sul risultato di un algoritmo che è discriminatorio e che peraltro non fornisce spiegazioni, anche l'imparzialità del giudice viene inevitabilmente pregiudicata⁶⁴¹.

Su queste problematiche è più facile ragionare prendendo l'esempio dei *risk assessment tools* impiegati negli Stati Uniti, essendo utilizzati in un settore come quello penale dove le implicazioni sul mancato rispetto delle garanzie costituzionali sono particolarmente delicate. In realtà problemi di indipendenza e imparzialità sono estendibili ad applicazioni algoritmiche di varia natura, proprio per i nodi irrisolti che ad oggi adombrano la giustizia algoritmica.

Se ragioniamo di giustizia predittiva è vero che senz'altro, come visto, sarebbe incostituzionale una norma che imponesse al giudice di vincolare la decisione alla "predizione" dell'algoritmo. E tuttavia, come si è visto, l'autorità dell'algoritmo unita alla mancanza di trasparenza delle procedure potrebbero facilmente comportare una 'delega' della decisione al risultato prodotto dalla macchina.

In questo senso, un sistema di giustizia predittiva, gestito dall'organo esecutivo con o senza la collaborazione di aziende private, rischierebbe di trasformarsi in uno strumento «*for controlling judges and interfering with judicial independence*»⁶⁴². Infatti, le predizioni – gestite oggi al di fuori dell'ordine giudiziario – potrebbero essere utilizzate per determinare quanto un caso sia controverso e, se la probabilità che un caso specifico porti a un determinato esito è alta, il giudice potrebbe

⁶⁴¹ Sul tema si v. tra gli altri, A. SIMONCINI, *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*, in *Rivista di diritto processuale*, 2, 2024, pp. 389-423. Si v. inoltre, J. ULENAERS, *The Impact of Artificial Intelligence on the Right to a Fair Trial: Towards a Robot Judge?*, in *Asian J Law Econ*, 11 (2) 2020.

⁶⁴² F. GALLI - G. SARTOR, *AI Approaches to Predictive Justice: A Critical Assessment*, cit. p. 184

concludere che il caso non sia controverso e di conseguenza non trattarlo con la dovuta attenzione⁶⁴³.

Proprio perché il processo decisionale giudiziario è un'attività altamente specializzata che richiede formazione ed esperienza significative, l'uso di strumenti, quali quelli di giustizia predittiva, basati sui dati in sostituzione della ricerca giuridica dei giudici, potrebbe compromettere la capacità di un giudice di impegnarsi nella risoluzione del caso e assumere la conseguente decisione, indebolendo per esempio la capacità di un giudice di determinare quale sia la prova rilevante e influenzare negativamente la sua capacità di valutare i fatti.

Per concludere, le esperienze oggi di giustizia algoritmica, sia negli Stati Uniti sia in Europa, sembrano sollevare alcune preoccupazioni per il rispetto delle garanzie di indipendenza e imparzialità del giudiziario e del singolo giudice. Ciò impone una seria riflessione sull'accettabilità costituzionale di supportare la decisione del giudice con strumenti preparati al di fuori dal controllo pubblico e, più specificamente, al di fuori del controllo dello stesso ordine giudiziario (su questo tema si cfr. *infra*, capitolo IX).

7. La *fairness* del processo giurisdizionale

Indagati i profili istituzionali del giusto processo e la disamina delle ragioni che conducono a ritenere *a rischio* i fondamentali principi di indipendenza e imparzialità che devono circondare la funzione giurisdizionale, occorre ora nei prossimi paragrafi spostare l'attenzione su quelli che la Corte europea dei diritti chiama "requisiti procedurali" e che insieme devono sussistere perché possano dirsi soddisfatte le garanzie del giusto processo. Ricordiamo fin subito che secondo l'interpretazione della Corte la valutazione della *fairness* del procedimento deve essere valutata «*as a whole*»: non è sufficiente cioè una singola violazione se ad essa viene posto rimedio nel corso del procedimento⁶⁴⁴, per esempio a seguito della

⁶⁴³ *Ibidem*.

⁶⁴⁴ Corte Edu, *Centro Europa 7 S.r.l. and Di Stefano v. Italy* [GC], 2012, § 197; *Regner v. the Czech Republic* [GC], 2017, §§ 151 e 161.

possibilità di impugnare la decisione dinanzi a un altro organo giudiziario che possieda tutti i requisiti dell'art. 6 § 1 CEDU⁶⁴⁵. Vedremo nel corso delle pagine che le violazioni a cui si assiste in diverse applicazioni di giustizia algoritmica difficilmente consentono di ripristinare le garanzie del giusto processo e, anzi, conducono inevitabilmente alla compressione del diritto fondamentale alla difesa e alla tutela giurisdizionale effettiva (si v. *infra*).

7.1. La giustizia algoritmica alla prova del contraddittorio e della parità delle armi

Mentre nella giurisprudenza americana non si fa direttamente riferimento ai principi di indipendenza e imparzialità del giudice, assorbiti nelle contestazioni sull'abuso di discrezionalità, i "requisiti procedurali" del *due process* sono invece espressamente invocati nei casi americani sull'impiego dei *tools* nel processo. Muovere dalla giurisprudenza americana, ancora una volta, aiuterà a fare luce sulla compatibilità degli algoritmi alla luce delle garanzie del giusto processo europeo.

In particolare, nelle controversie statunitensi si invoca il diritto costituzionale a venire giudicati sulla base di *accurate information*, strettamente legato al diritto di conoscere le informazioni contenute nel PSI prima del giudizio sulla determinazione della pena. Come visto, in *Loomis* si lamenta l'intrinseca violazione di tali principi in quanto la natura proprietaria di COMPAS non consentiva la verifica delle informazioni; in *Gordon e Guise* si lega la violazione di tale principio alla *reliance* che il giudice ha posto su un sistema algoritmico di cui non si conosce il funzionamento; in *People v. Younglove* si lamenta la mancanza di studi univoci sull'accuratezza del sistema algoritmico impiegato; nella *concurring opinion* allegata al caso *People v. Canedo*, al pari di *Loomis*, la giudice Cavanagh si chiede come fare a misurare l'accuratezza di un sistema il cui accesso è impedito dalla sua natura proprietaria (si v. *supra* § 3).

⁶⁴⁵ Corte Edu, *Oerlemans v. the Netherlands*, 1991, §§ 53-58.

La giurisprudenza federale e statale degli Stati Uniti riconosce apertamente tale diritto, ed è rilevante sinteticamente evidenziarla proprio perché fonda le contestazioni sollevate dalle difese nei casi nordamericani.

Punto di riferimento per l'affermazione di siffatto principio è la sentenza *Townsend v. Burke*⁶⁴⁶ della Corte Suprema, che aveva annullato una sentenza di condanna basata su prove poi dimostrate false. Si legge nella sentenza: «*we conclude that, while disadvantaged by lack of counsel, this prisoner was sentenced on the basis of assumptions concerning his criminal record which were materially untrue. Such a result, whether caused by carelessness or design, is inconsistent with due process of law, and such a conviction cannot stand*⁶⁴⁷».

Sempre la Corte Suprema qualche anno più tardi, in *United States v. Tucker*⁶⁴⁸, pur riconoscendo che il «*trial judge in the federal judicial system generally has wide discretion in determining what sentence to impose*» tale discrezionalità trova il suo limite nel caso in cui sia stata pronunciata una sentenza «*founded at least in part upon misinformation of constitutional magnitude*».

Ancora, in un caso degli anni '80, una corte federale, citando la giurisprudenza *Townsend* ha specificato che: «*The requirements of due process are not suspended with the pronouncement of guilt, but continue to operate in the sentencing process. Thus, a sentencing judge may not rely on mistaken information or baseless assumptions. However, courts must be concerned not merely when a sentencing judge has relied on demonstrably false information, but when the sentencing process created a significant possibility that misinformation infected the decision*»⁶⁴⁹.

Anche se la giurisprudenza non è stata costante negli anni – e sia a livello statale sia a livello federale ci sono casi in cui si mantiene solida la distinzione tra *trial*, in cui la protezione del *due process* è molto forte, e *sentencing*, caratterizzato invece da una più ampia discrezionalità da parte del giudice – la giurisprudenza *Townsend* e la sua progenie hanno mantenuto una notevole importanza e vengono citate come

⁶⁴⁶ *Townsend v. Burke*, 334 U.S. 736 (1948).

⁶⁴⁷ *Ibidem*, 740-741.

⁶⁴⁸ *United States v. Tucker*, 404 U.S. 443

⁶⁴⁹ *United States v. Lemon*, 723 F.2d 922, (D.C. Cir. 1983).

precedenti rilevanti dai giudici che devono vagliare la legittimità dell'utilizzo dei *tools* da parte del giudice. Ciò è dimostrato anche dall'affermazione del medesimo principio anche all'interno della giurisprudenza dei singoli stati: nello Stato del Wisconsin la Corte d'Appello nel caso *Skaff*⁶⁵⁰, citato in *Loomis*, aveva per la prima volta riconosciuto che:

defendant has a federal due process right to timely read his PSI to ensure its accuracy.

Del pari, in Michigan nel caso *People v. Malkowski*⁶⁵¹, citato in *People v. Younglove* insieme a *Townsend*, si precisa:

[i]t is vitally important to the defendant and to the ends of justice that the sentence be based upon accurate information. The great majority of defendants plead guilty. For those defendants, probation or determination of the length of time they will be in prison are the only significant decisions to be made by the judge.

E ancora, la Suprema Corte del Michigan ha più volte affermato che «*sentence is invalid if it is based on inaccurate information*»⁶⁵².

Come anticipato, strettamente legata all'accuratezza delle informazioni da poter considerare nel *sentencing* è la possibilità di verificare le informazioni contenute nel PSI (che quindi non dovrebbero restare segrete) e avere di conseguenza la possibilità di confutarle dinanzi a un giudice⁶⁵³.

A questo riguardo la Corte Suprema degli Stati Uniti, in *Gardner v. Florida*⁶⁵⁴, precedente citato in *Loomis*, aiuta a chiarire la portata del principio. Nella sentenza la Corte, dopo aver stabilito che la *due process clause* si applica anche nel

⁶⁵⁰ *State v. Skaff*, 152 Wis. 2d 48 (1989).

⁶⁵¹ *People v. Malkowski*, 385 Mich. 244 (1971).

⁶⁵² Si cfr., per esempio, *People v. Miles*, 454 Mich. 90 (1997), citata in *People v. Younglove*.

⁶⁵³ M. BRENNER *et al*, *Constitutional Dimensions of Predictive Algorithms in Criminal Justice*, cit. pp. 282 ss; Si v. anche J. VILLASENOR - V. FOGGO, *Artificial intelligence, due process, and criminal sentencing*, cit. pp. 214 ss. K. FREEMAN, *Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis*, cit., pp. 91-95.

⁶⁵⁴ *Gardner v. Florida.*, 430 U.S. 349 (1977).

sentencing, afferma che la mancata divulgazione delle informazioni all'imputato costituisce una violazione del giusto processo. In particolare, concludono i giudici:

that petitioner was denied due process of law when the death sentence was imposed, at least in part, on the basis of information which he had no opportunity to deny or explain.

Ora, come è stato fatto notare, *Gardner* riguardava una sentenza capitale, che ha le sue peculiarità e a cui si applicano requisiti più stringenti, diversi da quelli richiesti nelle condanne ordinarie⁶⁵⁵. Ciò nondimeno, la sua importanza nel contesto di decisioni algoritmiche rimane intatta, venendo anch'essa citata, come visto, da imputati condannati sulla base di algoritmi predittivi.

Seguendo *Gardner*, in Wisconsin nel caso *Skaff*, già menzionato *supra*, la Corte aveva stabilito che

denial of access to the PSI denied the defendant an essential factor of due process, i.e., a procedure conducive to sentencing based on correct information

quindi sostenendo che il diritto a essere giudicati su informazioni affidabili porta con sé il diritto di verificare i dati contenuti nel PSI, su cui la Corte fonda la propria decisione.

Non dissimili, e anzi ancora più specifiche, appaiono le garanzie previste dal giusto processo nella prospettiva europea multilivello.

Nella giurisprudenza della Corte europea dei diritti uno degli aspetti fondamentali di in ogni processo, tanto civile quanto penale⁶⁵⁶, è che sia svolto in contraddittorio tra le parti (*adversarial proceeding*). Tale diritto, inoltre, è strettamente connesso e spesso sindacato congiuntamente alla parità delle armi (*equality of arms*)⁶⁵⁷, principio che implica che a ciascuna parte venga data una ragionevole opportunità di presentare il proprio caso – comprese le prove – in

⁶⁵⁵ J. VILLASENOR - V. FOGGO, *Artificial intelligence, due process, and criminal sentencing*, cit., p. 324.

⁶⁵⁶ Corte Edu, *Werner v. Austria*, 1997, § 66.

⁶⁵⁷ Corte Edu, *Regner v. the Czech Republic* [GC], 2017, § 146.

condizioni che non la pongano in una situazione di «*substantial disadvantage*» rispetto all'altra parte⁶⁵⁸.

Ed è proprio sul diritto a conoscere gli elementi (le prove) su cui si basa la decisione della controversia che si concentra il giudizio della Corte sul rispetto dei principi dell'*adversarial proceeding*. Le parti devono avere infatti l'opportunità di essere a conoscenza di tutte le prove presentate o gli elementi che possono influire sulla decisione del giudice⁶⁵⁹.

A questo riguardo il *leading case* in materia è il caso *Brandstetter c. Austria*, in cui la Corte ha precisato che la parità delle armi e il diritto al contraddittorio implicano che venga assicurata alle parti l'effettiva conoscenza degli elementi che hanno influenzato la decisione del giudice nonché della concreta possibilità di contestarle e falsificarle⁶⁶⁰. Infatti, «*an indirect and purely hypothetical possibility for an accused to comment on prosecution arguments included in the text of a judgment*»⁶⁶¹ non soddisfa l'art. 6 CEDU⁶⁶².

Questa garanzia – e per i nostri fini ciò è particolarmente rilevante – non si limita agli elementi presentati dalle parti nel procedimento, ma comprende le prove e le informazioni assunte su iniziativa dello stesso giudice (come lo sono i risultati dei *risk tools* e come, ragionevolmente, saranno i sistemi di giustizia predittiva)⁶⁶³.

Ebbene, la necessaria *disclosure* degli elementi rilevanti del giudizio e delle prove è ovviamente finalizzata a verificarne l'accuratezza o, nell'espressioni della Corte europea, l'autenticità. Sia nei processi civili sia in quelli penali la Corte, per quanto abbia costantemente ripetuto che, in linea di principio, le regole sull'ammissibilità della prova sono devolute alle discipline nazionali, nondimeno

⁶⁵⁸ Corte Edu, *Kress v. France* [GC], 2001, § 72.

⁶⁵⁹ *Ibidem*, §§ 65 e 44.

⁶⁶⁰ Corte Edu., *Brandstetter v. Austria*, 1991, § 67.

⁶⁶¹ Corte. Edu., *Brandstetter v. Austria*, cit. § 68.

⁶⁶² Il caso riguardava un procedimento penale ma le considerazioni spese dalla Corte sono riprese dalla stessa e calate anche in procedimenti civili. Si v. per esempio Corte Edu, *McMichael v. the United Kingdom*, 1995, § 80 e *K.S. v. Finland*, 2001, §§ 23-24.

⁶⁶³ Di recente, Corte Edu, *Vorotnikova v. Latvia*, 2021, § 26.

l'art. 6 CEDU impone che alle parti venga data la possibilità di contestare il contenuto e l'autenticità delle prove con modalità *adeguate*⁶⁶⁴.

Se queste sono le coordinate generali delle garanzie costituzionali imposte dal giusto processo come interpretato dalle Corti europee e nordamericane, occorre ora calare questi principi processuali nel contesto della giustizia algoritmica.

Anche qui consideriamo i nodi problematici esposti nel capitolo VI per orientare la riflessione sulle evidenti compressioni dei diritti di parità delle armi e del contraddittorio⁶⁶⁵.

Innanzitutto occorre valutare la compatibilità con le garanzie esposte nel presente capitolo per quanto concerne i *risk tools* impiegati negli Stati Uniti, poiché abbiamo già una giurisprudenza che ne ha sindacato l'impiego.

Il punto è ben individuato, come già visto, dalla giudice Cavanagh nel caso *People v. Canedo* (si v. *supra*, § 3), la quale ha sottolineato che di fronte a un algoritmo proprietario nemmeno ha senso chiedersi se il sistema sia da considerarsi accurato, posto che non vi è modo di conoscere il codice e il *dataset* dell'algoritmo. Afferma ancora la giudice

The secretive nature of the algorithm raises questions. Without knowing what the algorithm is, it is difficult to know whether and how race, class, and other personal factors influence a potentially biased score.

Come rilevato in dottrina, i protagonisti della dinamica processuale, per valutare l'accuratezza delle informazioni contenute nell'algoritmo dovrebbero quindi poter esaminare ed eventualmente contestare il procedimento seguito dallo stesso nonché i dati contenuti all'interno⁶⁶⁶.

In questo senso le soluzioni adottate nel caso *Loomis* in cui la Corte ha ritenuto sufficiente che l'impiego di COMPAS sia sorretto da alcune specifiche cautele da

⁶⁶⁴ Per il processo civile si cfr. Corte EDU, *Krčmář and Others v. the Czech Republic*, 2000, § 42; per quello penale si cfr., di recente, *Yüksel Yalçınkaya v. Türkiye* [GC], 2023, § 324.

⁶⁶⁵ Sulla violazione di tali diritti nel processo penale si cfr. le interessanti riflessioni svolte in S. QUATTROCOLO, *Equità del processo penale e automated evidence alla luce della convenzione europea dei diritti dell'uomo*, in *Revista Ítalo-Española de Derecho Procesal*, Vol. 1, 2019 pp. 17 ss.

⁶⁶⁶ M. BRENNER *et al*, *Constitutional Dimensions of Predictive Algorithms in Criminal Justice*, cit., p. 282.

indicare nel PSI tra cui la natura proprietaria dell'algoritmo e l'impossibilità di accedervi, anche se perlomeno manifesta la consapevolezza della Corte sui problemi che tale segretezza comporta⁶⁶⁷, non convincono né sotto sotto la lente del *due process* statunitense, né tantomeno alla luce del *fair trial* europeo.

E nemmeno sembra accettabile la conclusione della Corte del Wisconsin quando, per rigettare l'argomento della segretezza delle informazioni contenute nel PSI, afferma che l'imputato, pur non potendo contestare il funzionamento dell'algoritmo *tout court*, poteva, con l'aiuto delle spiegazioni contenute nella Guida COMPAS, almeno «*review and challenge the resulting risk scores set forth in the report attached to the PSI*»⁶⁶⁸ e quindi, in definitiva, aveva «*an opportunity to challenge his risk scores by arguing that other factors or information demonstrate their inaccuracy*»⁶⁶⁹.

Del pari, non coglie nel segno nemmeno la decisione assunta in *People v. Younglove* (si v. *supra*, § 3), in cui la Corte rigetta l'argomento della non accuratezza delle informazioni sulla base del fatto che gli imputati hanno mancato di indicare specificatamente quali valutazioni di COMPAS fossero «*inaccurate*».

Innanzitutto, poiché le società produttrici sono parti interessate che, come visto, stipulano contratti particolarmente onerosi per fornire la licenza all'utilizzo dei *tools*, i documenti da esse prodotti (come per esempio la guida COMPAS) non possono essere considerati elementi sufficienti per confermare l'affidabilità del

⁶⁶⁷ In questo senso, di interesse è anche il caso *State v. Buford* (2019), che offre un esempio interessante di come l'atteggiamento dei giudici americani di fronte a strumenti di analisi algoritmica non trova riscontro in pronunce che, non occupandosi direttamente dell'utilizzo di COMPAS o di altri strumenti di analisi algoritmica ma di situazioni senz'altro analoghe, sembrano contraddire apertamente la consolidata giurisprudenza americana sui *risk assessment tools*. Nella pronuncia in esame la Corte d'appello del Wisconsin annulla la sentenza di primo grado e rinvia per una nuova determinazione della pena da applicare in concreto in quanto la Corte distrettuale aveva basato la propria decisione su un'«informazione inaccurata» e in particolare su uno studio statistico sull'incidenza degli abusi sessuali sui minori, dal quale emergeva un indice particolarmente alto. In questo caso la Corte d'appello, condividendo le doglianze del condannato, riteneva che quel dato non fosse presente nello studio utilizzato dal giudice di primo grado e quindi non poteva costituire un elemento su cui fondare l'applicazione della pena. Alla controdeduzione dello Stato secondo cui, anche se la Corte si era basata su quella informazione erronea, la decisione sarebbe stata la medesima, la Corte replica che «*it is not clear beyond a reasonable doubt that the circuit court would have imposed the same sentence without this inaccurate information*».

⁶⁶⁸ *Loomis*, § 53.

⁶⁶⁹ *Ivi*, § 56.

sistema utilizzato⁶⁷⁰. E in effetti gli studi citati nella Guida che certificano l'*accuracy* del sistema sono spesso condotti da persone che lavorano nella società produttrice e non da soggetti indipendenti⁶⁷¹.

In secondo luogo, anche conoscendo gli *input* che vengono inseriti nell'algoritmo, le teorie criminologiche alla base, il campione utilizzato per addestrare l'algoritmo, le tecniche di *machine learning* impiegati e l'*output* finale – tutti elementi che sono presenti nella guida di COMPAS (si v. *supra* capitolo I § 4) – senza l'accesso al procedimento algoritmico è impossibile valutare realmente come l'algoritmo ha assunto il proprio *score* di rischio⁶⁷².

In definitiva, rispetto ai *risk assessment tools*, sembra potersi affermare che per le loro caratteristiche e per le modalità in cui vengono impiegati all'interno delle giurisdizioni statunitensi non risultano compatibili, ad oggi, con le garanzie costituzionali del contraddittorio e della parità delle armi.

Tali garanzie, peraltro, sembrano poter essere poste sotto pressione anche nelle altre ipotesi di giustizia algoritmica che sembrano oggi essere all'orizzonte.

Quanto agli strumenti di giustizia predittiva, qualora fossero prodotti esclusivamente e senza controlli da imprese private, avrebbero in ogni caso il problema che le parti, senza adeguate guarentigie normative di trasparenza, non potrebbero accedervi, minando così la possibilità di contestare un elemento che, come visto, potrebbe diventare decisivo per la definizione della controversia.

Inoltre, pensando ancora a strumenti impiegati dai giudici che “prevedono” l'esito di una controversia dove la progettazione è affidata parzialmente all'esecutivo (come visto è spesso accaduto, vedi *supra* capitolo II), qualora la

⁶⁷⁰ K. FREEMAN, *Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis*, cit., p. 95.

⁶⁷¹ K. B. FORREST, *When machines can be judge, jury, and executioner: justice in the age of artificial intelligence*, cit., p. 77.

⁶⁷² Ciò viene spiegato da M. BRENNER *et al*, *Constitutional Dimensions of Predictive Algorithms in Criminal Justice*, cit., p. 282, con un esempio particolarmente efficace: «consider the following mathematical example: $2 \times 2 = 4$. We know the inputs: 2 and 2. We know the output: 4. What we do not know is the function of the computation, “ \times ”. Could it be that the equation is $2+2$? 2×2 ? 2^2 ? This may not be of mathematical significance because the output, 4, remains constant no matter the computation. But it could be of legal significance if, for example, multiplication was constitutionally impermissible. Just as defendants have an interest in understanding and challenging the factors that determined their sentencing offense level, they also have an interest in understanding and challenging the inputs and calculations that an algorithm uses to estimate risk».

controversia conduca di fronte al giudice una parte privata e lo Stato o un'articolazione di esso è evidente che si creerebbe un'asimmetria conoscitiva incolmabile tra le parti in giudizio⁶⁷³.

Ma anche in un caso di giudizio tra privati l'impiego di mezzi di intelligenza artificiale impiegati dal giudice, quand'anche fossero posti nella disponibilità conoscitiva delle parti, potrebbe creare delle compressioni del diritto al contraddittorio e alla parità delle parti. Qualunque sarà il modo in cui si potrà accedere agli algoritmi, essi saranno probabilmente sempre più complessi e richiederanno competenze sempre maggiori per comprenderli: tralasciando qui il ruolo del giudice di fronte a tale complessità (già vista e di cui ancora si discuterà) l'uso di mezzi algoritmici potrebbe facilitare i procedimenti per alcune parti, come aziende con particolari risorse o persone esperte di informatica e, d'altro canto, creare notevoli difficoltà per altre⁶⁷⁴ (a ciò è strettamente collegato il tema dell'alfabetizzazione digitale, si v. *infra*, capitolo IX).

Parrebbe dunque sempre più fragile l'equazione tra uno strumento di giustizia predittiva e i metodi di ricerca oggi impiegati sui precedenti giurisprudenziali, anche ovviamente con mezzi tecnologici: se in quest'ultimo caso le parti potrebbero, a seguito della motivazione del giudice basata anche sui precedenti, in condizioni di parità, contestare la decisione, nel caso di strumenti di giustizia predittiva una delle parti potrebbe diventare l'unica a conoscere quell'elemento che è risultato decisivo per la decisione stessa⁶⁷⁵.

Epitomando quanto affermato in questo paragrafo, secondo le prospettive oggi realistiche di giustizia algoritmica sembra a rischio anche la *fairness* del procedimento, principalmente per i rischi che corrono i diritti al contraddittorio e alla parità delle armi in giudizio. Ed è vero che in diverse occasioni la Corte europea

⁶⁷³ In questo senso A. PAJNO *et. al.*, *AI: profili giuridici: Intelligenza Artificiale: criticità emergenti e sfide per il giurista*, in *BioLaw Journal – Rivista di BioDiritto*, n. 3/2019, p. 228. Si cfr. anche, in campo penale, V. MANES, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in *Discrimen*, 15 maggio 2020, p. 15.

⁶⁷⁴ J. ULENAERS, *The Impact of Artificial Intelligence on the Right to a Fair Trial: Towards a Robot Judge?*, in *Asian J Law Econ*, 11 (2) 2020, pp. 15-26.

⁶⁷⁵ Sull'inopportunità di tale equazione dal punto di vista del giudice si cfr. G. DE MINICO, *Giustizia e intelligenza artificiale: un equilibrio mutevole*, in *Rivista AIC*, n. 2/2024, p. 96.

dei diritti ha affermato che il diritto a un *adversarial proceeding* non è assoluto, soprattutto quando la mancata *disclosure* degli elementi rilevanti non ha avuto un impatto sulla decisione del giudice⁶⁷⁶, ma gli interessi che dovrebbero bilanciare tale mancata *disclosure* dovrebbero riguardare motivi di sicurezza nazionale o specifici interessi pubblici⁶⁷⁷: difficilmente nel bilanciamento con tale diritto potrebbero entrare, come invece è successo negli Stati Uniti, motivazioni legate al segreto industriale degli algoritmi, a nostro avviso recessive di fronte al diritto al contraddittorio. In ogni caso, secondo la Corte europea sono ammesse solo le misure che limitano i diritti di una parte a conoscere gli elementi rilevanti nel procedimento che non compromettono l'essenza stessa del *fair trial* ai sensi dell'art. 6 CEDU⁶⁷⁸: quando le prove non sono state mostrate al ricorrente per motivi di interesse pubblico, la Corte esamina attentamente il processo decisionale del giudice per verificare che comunque abbia rispettato i requisiti di un processo equo e della parità delle armi nel suo complesso⁶⁷⁹.

Anzi, proprio nel caso degli strumenti algoritmici di elevata complessità tecnica e come visto, per diverse ragioni, estranei alla conoscenza del giudice (il che provoca l'ormai noto *automation bias*), dovrebbe essere garantita la possibilità di contestare tale prova, vista l'influenza sulla decisione della controversia. Ed è la stessa Corte europea che talvolta ha riconosciuto che quando il giudice si affida a pareri tecnici di cui ha poca contezza, tali pareri acquisiscono un ruolo determinante nella decisione del caso, di talché ancor più ne dovrebbe essere assicurata la possibilità di contestazione⁶⁸⁰, soprattutto quando il parere tecnico (nel nostro caso l'algoritmo) diventa la prova decisiva o l'unica prova del processo⁶⁸¹.

In definitiva, l'utilizzo di algoritmi all'interno del sistema giudiziario pone notevoli interrogativi di compatibilità con le garanzie costituzionali considerate.

⁶⁷⁶ Vedi, in particolare, Corte EDU *Janáček v. the Czech Republic*, 2023, § 48.

⁶⁷⁷ Corte Edu, *Regner v. the Czech Republic* [GC], 2017.

⁶⁷⁸ Di recente, *Adomaitis v. Lithuania*, 2022, §§ 68-74.

⁶⁷⁹ Corte Edu, *Regner c. Repubblica Ceca* [GC], 2017, §§ 147-49.

⁶⁸⁰ V. per esempio Corte EDU, *Mantovanelli v. France*, 1997, § 36; *Storck v. Germany*, 2005, § 135.

⁶⁸¹ Corte Edu, *Devinar v. Slovenia*, 2018, §§ 49-50.

In questo senso, anche se, come detto, la Corte europea non si è ancora pronunciata su ipotesi di giustizia algoritmica (poiché ancora non in uso in Europa) una recente sentenza della Grande Camera può fornire, quantomeno sul diritto all'*adversarial proceeding* e alla parità delle armi, qualche insegnamento per un'eventuale futura giurisprudenza sull'impiego di algoritmi da parte del giudice⁶⁸². Il caso riguardava l'impiego in un procedimento penale da parte dell'autorità pubblica di una prova elettronica (non algoritmica quindi, ma pur sempre digitale) su cui il giudice si era affidato per determinare la condanna e di cui non si poteva conoscere la validità e l'accuratezza. La Corte ha riconosciuto la violazione dell'art. 6 § 1 per violazione del diritto al contraddittorio, volendo altresì specificare alcune questioni riguardanti la prova utilizzata, che ci sembrano estendibili anche all'impiego futuro di algoritmi decisionali da parte dei giudici.

La Corte nota dapprima che «*electronic evidence differs in many respects from traditional forms of evidence, including as regards its nature and the special technologies required for its collection, securing, processing and analysis. Most significantly, it raises distinct reliability issues as it is inherently more prone to destruction, damage, alteration or manipulation. The Court also reiterates that the use of untested electronic evidence in criminal proceedings may involve particular difficulties for the judiciary as the nature of the procedure and technology applied to the collection of such evidence is complex and may therefore diminish the ability of national judges to establish its authenticity, accuracy and integrity*»⁶⁸³. Precisa inoltre che la particolarità delle prove elettroniche non significa che «*in the abstract call for the safeguards under Article 6 § 1 to be applied differently, be it more strictly or more leniently*»⁶⁸⁴. Inoltre se è vero che prove di questo tipo possono essere utili in diversi campi (in quel caso si trattava di lotta al terrorismo) «*the right to a fair trial, from which the requirement of the proper administration of justice is to be inferred, applies to all types of criminal offence, from the most straightforward to the most complex. The right to the fair administration of justice*

⁶⁸² Corte Edu, *Yüksel Yalçınkaya v. Türkiye* [GC], 2023.

⁶⁸³ *Ivi*, § 312.

⁶⁸⁴ *Ivi*, § 313.

holds so prominent a place in a democratic society that it cannot be sacrificed for the sake of expediency and the evidence obtained, whether electronic or not, may not be used by the domestic courts in a manner that undermines the basic tenets of a fair trial»⁶⁸⁵.

In sintesi, estendendo tali considerazioni a una futura giustizia algoritmica in Europa, la Corte europea sembra essere consapevole che nasceranno tecnologie all'avanguardia che si differenzieranno dalle modalità tradizionali di assunzione delle decisioni; e ciò nondimeno, le garanzie del giusto processo sono perfettamente adattabili al nuovo contesto senza bisogno di interpretazioni più restrittive o estensive. Inoltre, e ciò sembra racchiudere il nucleo della decisione, la corretta amministrazione della giustizia e i principi del giusto processo hanno un ruolo tanto determinante in una società democratica che non dovrebbero essere mai sacrificati per ragioni di convenienza (e, potremmo dire, di efficienza o celerità).

A questo punto un quesito sorge spontaneo: è lecito chiedersi infatti se le problematiche segnalate potrebbero in parte alleviarsi tramite un utilizzo accorto dello strumento principe di verifica critica, da parte tanto delle parti quanto della pubblica opinione, dell'esercizio della funzione giurisdizionale, vale a dire la motivazione.

Ciò anche perché le norme a difesa della parità delle armi sono strettamente connesse all'autonomia del giudice e al dovere di motivare, «quanto meno nel senso che tale autonomia può esistere in quanto il contraddittorio è pieno e paritario, dal momento che ogni squilibrio del contraddittorio inevitabilmente si risolverebbe in una oggettiva situazione di mancanza di equilibrio del giudice, che si vedrebbe proiettato nell'orbita e nel potere di una delle parti»⁶⁸⁶.

In questi termini, come si dirà nel prossimo paragrafo, traluce un'insufficienza intrinseca dello sforzo motivazionale, sia per le caratteristiche proprie dei moderni sistemi algoritmici – come già ampiamente evidenziati – sia per il divario, oggi ancora ampio, tra la complessità di siffatti sistemi e la formazione del giudice.

⁶⁸⁵ *Ivi*, § 344.

⁶⁸⁶ A. PAJNO, *Intelligenza artificiale e autonomia del giudice. il ruolo del fatto e il valore del precedente*, cit., p. 403.

7.2. La giustizia algoritmica alla prova della motivazione della decisione

Come visto, negli Stati Uniti la motivazione delle sentenze è uno dei corollari del *due process of law* (si v. *supra*, capitolo 7 § 1) e nel caso *Loomis* i giudici affermano che «[a]n erroneous exercise of discretion occurs when a circuit court imposes a sentence “without the underpinnings of an explained judicial reasoning process”»⁶⁸⁷. Quando mancano i “sostegni” per la motivazione, quindi, non può che ravvisarsi un esercizio fallace del potere discrezionale, privato degli elementi essenziali per il suo costruirsi e, conseguentemente, la violazione delle garanzie processuali. La soluzione cui giungono spesso in modo non del tutto convincente i giudici statunitensi, come visto, recupera tuttavia la legittimità dell’esercizio motivazionale nella *non esclusività* del *tool* utilizzato, chiarendo che comunque il giudice dovrebbe spiegare e motivare anche sulla rilevanza e il peso degli elementi diversi dall’algoritmo⁶⁸⁸.

Che la motivazione costituisca, anche nel nostro sistema costituzionale, una garanzia ordinamentale irrinunciabile tanto endoprocessuale quanto extraprocessuale è cosa nota. Nella nostra Costituzione l’obbligo motivazionale, previsto dall’art. 111, sesto comma, Cost. («tutti i provvedimenti giurisdizionali devono essere motivati»), costituisce un principio ineludibile del giusto processo. La motivazione, infatti, assolve alla funzione sia di garantire l’esercizio legittimo

⁶⁸⁷ *Loomis*, § 30.

⁶⁸⁸ Si afferma per esempio in *Loomis* che «a circuit court must explain the factors in addition to a COMPAS risk assessment that independently support the sentence imposed», § 99. Ma questa fiducia nella capacità del giudice di motivare in base a tutti fattori in gioco si afferma in diversi contesti. Già la Guida di COMPAS incoraggia gli utilizzatori dell’algoritmo – tra cui quindi il giudice – a impiegare il proprio giudizio professionale per disattendere il risultato dell’algoritmo quando esso non appare conforme alle circostanze specifiche del caso concreto (si v. *supra*, cap. I § 4). Nel *Model Penal Code* si precisa che i sistemi di valutazione del rischio di recidiva debbano sempre essere sotto il controllo del giudizio discrezionale del giudice umano (si v. *supra*, cap. I § 5). Nel respingere le doglianze dell’imputato in *Malenchik* la Corte afferma che i sistemi algoritmici sono in ogni caso sottoposti al *sound discernment* del giudice che è in grado di assegnare agli strumenti predittivi la giusta considerazione e il giusto peso (si v. *supra* in questo capitolo).

della giurisdizione da parte di un giudice che sia imparziale e indipendente, sia di consentire il rispetto dei diritti delle parti, in particolar modo il diritto di difesa⁶⁸⁹.

Ancora la Corte EDU ha contribuito in modo decisivo a delineare i requisiti della motivazione dei provvedimenti giurisdizionali.

Nella giurisprudenza della Corte in molteplici occasioni si afferma che la *fairness* prevista dall'art. 6 § 1 CEDU impone alla Corte di fornire *sufficient reasons* delle proprie ragioni. In una recente sentenza la Corte sintetizza le caratteristiche che deve avere la motivazione per non incorrere in violazioni del giusto processo⁶⁹⁰. In detta sentenza precisa che «*in view of the principle that the Convention is intended to guarantee not rights that are theoretical or illusory but rights that are practical and effective, the right to a fair trial cannot be seen as effective unless the requests and observations of the parties are truly "heard", that is to say, properly examined by the tribunal*». Inoltre: «*Judgments of courts and tribunals should adequately state the reasons on which they are based. Without requiring a detailed answer to every argument advanced by the complainant, this obligation to give reasons presupposes that parties to judicial proceedings can expect to receive a specific and explicit reply to the arguments which are decisive for the outcome of those proceedings*» Infine specifica altresì che «*the extent to which this duty to give reasons applies may vary according to the nature of the decision and must be determined in the light of the circumstances of the case*».

Quindi la motivazione è necessaria affinché il provvedimento non appaia *arbitrario*⁶⁹¹ e deve essere tale che le parti possano ottenere un rimedio non illusorio e, quindi, impugnare la sentenza⁶⁹². Soprattutto, per quanto il grado di dettaglio della motivazione può variare rispetto alle circostanze del singolo processo o del

⁶⁸⁹ E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit., p. 342.

⁶⁹⁰ Corte Edu, *Zayidov v. Azerbaijan* (no. 2), 2022, § 91.

⁶⁹¹ La Corte talvolta valuta espressamente la legittimità della motivazione sotto il profilo dell'*arbitrariness*, si v. per esempio Corte Edu, *Aykhan Akhundov v. Azerbaijan*, 2023, §§ 105 ss.

⁶⁹² Corte Edu, *Hirvisaari v. Finland*, 2001, § 30.

caso concreto, quantomeno il giudice deve motivare sugli elementi che sono risultati *decisivi* per la decisione⁶⁹³.

Se queste sono le coordinate, occorre verificare se la motivazione possa in un certo senso contenere il fenomeno della esternalizzazione del potere giudicante e, di conseguenza, la violazione delle garanzie del giusto processo, o se anche questa conclusione sia in realtà solo un ulteriore velo di Maya che nasconde una realtà più complessa.

Ora, un discorso sulla possibilità per il giudice di motivare adeguatamente sull'impiego che ha fatto del risultato algoritmo non può che prendere le mosse da un fondamentale *distinguo*: altro è delineare il perimetro dell'obbligo motivazionale a fronte di algoritmi del tutto trasparenti – o, perlomeno, facilmente comprensibili dalle parti e dal giudice – e sotto il controllo del giudiziario, altro è considerare legittima una motivazione che si fonda, parzialmente o esclusivamente, su un'informazione prodotta dagli algoritmi che rimangono tuttora oscuri agli occhi del giudice.

Nella prima ipotesi nulla osta a che, attraverso un impiego accurato della motivazione – che tuttavia potrà raggiungersi solo allorché il giudice sarà pienamente consapevole dello strumento utilizzato (si v. *infra*, cap. IX) – la discrezionalità venga ad esercitarsi in modo legittimo: è facilmente immaginabile un giudice che, ben conscio dei dati inseriti all'interno dell'algoritmo e delle modalità attraverso cui esso produce il punteggio di rischio, consideri quest'ultimo alla stregua degli altri elementi di fatto emersi in sede processuale, confermando ovvero disattendendo il risultato algoritmico attraverso la motivazione.

Quanto invece ad algoritmi, come la maggior parte di quelli utilizzati nelle Corti americane e di quelli più evoluti che oggi stanno venendo sperimentati nelle giurisdizioni europee, che mancano di trasparenza per gli ostacoli giuridici o tecnici di cui abbiamo discusso (si v. *supra*, capitolo VI § 2), non può che riconoscersi l'insufficienza dello sforzo motivazionale dei giudici per due ordini di ragioni: in primo luogo l'inaccessibilità degli algoritmi più evoluti non consente al

⁶⁹³ In generale si v. Corte Edu, *Ruiz Torija v. Spain*, 1994, § 30, nel processo penale si v. Corte Edu, *Moreira Ferreira v. Portugal* (no. 2) [GC], 2017, § 84.

giudice di rendere comprensibile – e quindi legittima – la propria decisione, e ciò sia nel caso in cui decida di disattendere il risultato automatico, assegnando peso decisivo ad altri fattori indipendenti, sia che invece intenda attenervisi; in secondo luogo, anche nei casi di impiego di algoritmi astrattamente idonei a consentire l’accesso al codice e di rendere noti i dati e il funzionamento del sistema, sembra ad oggi mancare al giudice, come meglio si chiarirà, la formazione professionale adeguata per governare al meglio una decisione che rischia di risultare intrinsecamente immotivata e quindi illegittima.

Del resto ciò è anche stato riconosciuto da una *concurring opinion* della giudice Abrahamson nella stessa sentenza *Loomis*, nella quale afferma che «*this court’s lack of understanding of COMPAS was a significant problem in the instant case*», che rischia di infirmare il «*basic requirement that a circuit court explain its exercise of discretion at sentencing*»⁶⁹⁴.

A ciò si aggiunga che, come visto, l’algoritmo assume nella decisione un *peso decisivo*, e l’impossibilità di motivare proprio sul fattore che ha determinato la decisione costituirebbe una patente violazione dell’obbligo motivazionale secondo la menzionata giurisprudenza della Corte europea dei diritti dell’uomo.

Il destino incerto della motivazione di fronte a sistemi algoritmici impiegati dai giudici è stato evidenziato sia dalla dottrina nordamericana sia da quella europea e italiana.

Quanto all’utilizzo dei *tools* nel processo penale, si è detto che se un punteggio di rischio “alto” può portare un tribunale a emettere una sentenza più grave rispetto a quella che altrimenti avrebbe emesso (come si è visto), non è sostenibile affermare che tali sistemi non abbiano alcun impatto sulla gravità della sanzione penale⁶⁹⁵, come invece più volte sostenuto dalle Corti. In altre parole, parrebbe irrealistico aspettarsi che un Giudice, dopo aver letto il *risk score* prodotto dall’algoritmo,

⁶⁹⁴ *Loomis*, § 141.

⁶⁹⁵ J. VILLASENOR - V. FOGGO, *Artificial intelligence, due process, and criminal sentencing*, in *Mich. St. L. Rev.*, 2020, p. 311.

eserciti il proprio potere motivazionale senza alcun pregiudizio nei confronti dell'imputato⁶⁹⁶.

Spostandoci in generale sugli algoritmi utilizzati nel processo, quali anche strumenti di giustizia predittiva, secondo alcuni il giudice che impiegherà tali strumenti sarà indotto a motivare seguendo modelli preesistenti⁶⁹⁷. Siffatta motivazione, se si limitasse a indicare: «ho deciso così perché l'algoritmo predittivo ha indicato tale esito» sarebbe, prima ancora che illegittima, tautologica⁶⁹⁸ e fintanto che la decisione sarà catturata da sistemi il cui accesso è impedito alle parti⁶⁹⁹, difficilmente potrà essere accettata come ragionevole dalle stesse⁷⁰⁰.

Sembrerebbe quindi che l'assolvimento di un vero e legittimo onere motivazionale, che permetta di addivenire a un esercizio discrezionale che sia coerente con la funzione del giudice e – in ultimo – che garantisca il rispetto del giusto processo, potrà raggiungersi solo allorché gli algoritmi saranno realmente accessibili e che gli stessi giudici acquisiscano gli strumenti adeguati al fine di ponderare correttamente tutti gli elementi di fatto rilevanti nel processo.

⁶⁹⁶ H. LIU - C. LIN - Y. CHEN, *Beyond State v Loomis: artificial intelligence, government algorithmization and accountability*, cit., p. 130.

⁶⁹⁷ E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit., p. 343.

⁶⁹⁸ Si cfr. G. DE MINICO, *Giustizia e intelligenza artificiale: un equilibrio mutevole*, cit., p. 96. Segnalano le problematiche relative alla motivazione anche C. CASONATO, *Giustizia e intelligenza artificiale: considerazioni introduttive*, in *BioLaw Journal – Rivista di BioDiritto*, 2, 2021, p. 362; ID, *Intelligenza artificiale e giustizia: potenzialità e rischi*, cit., p. 3380. A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, cit., pp. 81.

⁶⁹⁹ In questo senso A. Simoncini rilegge la giurisprudenza costituzionale sul divieto di automatismi legislativi irragionevoli. Se infatti il giudice sarà vincolato, per la sua «forza pratica» al risultato algoritmico, ecco che tale vincolo condurrebbe a una compressione del libero convincimento del giudice e della conseguente motivazione. In questo senso si v. A. SIMONCINI, *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*, in *Rivista di diritto processuale*, 2, 2024, p. 416. Ancor prima l'A. aveva espresso tale pensiero in A. SIMONCINI - S. SUWEIS, *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, cit., p. 101. Rispetto agli automatismi legislativi, la Corte costituzionale ha, anche di recente, affermato che «in un sistema equilibrato di tutele, la discrezionalità del giudice riveste un ruolo cruciale», si cfr. da ultimo Corte Cost., sent. n. 59 del 2021.

⁷⁰⁰ K. D. ASHLEY, *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*, in *Cambridge University Press*, 2017.

8. Nota conclusiva: la giustizia algoritmica alla prova del diritto di difesa e a una tutela giurisdizionale effettiva

Nel capitolo VI abbiamo tentato di delineare i problemi che oggi attengono all'impiego di strumenti algoritmici e di intelligenza artificiale nel processo, legati soprattutto alla progettazione degli algoritmi, all'opacità giuridica e tecnica degli stessi e alla tendenza dell'essere umano, e quindi del giudice, ad affidarsi alla tecnologia.

Sulla base di tali problematiche, dopo aver delineato le caratteristiche del *due process of law* e del giusto processo nella prospettiva europea (capitolo VII) abbiamo dedicato quest'ultimo capitolo a tentare di comprendere quali sono oggi le conseguenze più evidenti sulle garanzie costituzionali del giusto processo: per perseguire tale obiettivo abbiamo tratto insegnamento dalla giurisprudenza statale nordamericana, in cui la giustizia algoritmica ha avuto, negli ultimi anni, un'estensione notevole nella forma di algoritmi predittivi impiegati nel processo penale. Sulla base di tale esperienza si è tentato di ragionare sui vari livelli in cui gli algoritmi, in generale, possono influenzare la decisione del giudice e il suo potere discrezionale.

Dimostrata tale influenza, abbiamo spiegato come essa influisce sulle garanzie del "giusto processo" nella prospettiva multilivello, segnalata al principio di questo lavoro: in tal senso, abbiamo preso in considerazione sia le garanzie istituzionali della magistratura (in particolare i principi di indipendenza e imparzialità) sia le garanzie procedurali legate alla *fairness* del procedimento giurisdizionale (in particolare, il diritto al contraddittorio, alla parità delle armi, e alla motivazione dei provvedimenti).

Ora, arrivati a questo punto ci sembra di poter racchiudere l'intero discorso sulla compatibilità della giustizia algoritmica con le garanzie costituzionali in un diritto che a sua volta funge da guardiano di tutti i principi e diritti che abbiamo delineato: il diritto a una tutela giurisdizionale *effettiva*, che rappresenta un'espressione in grado di ricomprendere tutti i requisiti del giusto processo che abbiamo indicato nei precedenti paragrafi e che sono implicati dalle ipotesi di giustizia algoritmica.

Tale principio, come noto, è strettamente connesso il diritto di difesa in giudizio, come reso ben chiaro dalla giurisprudenza della Corte europea dei diritti dell'uomo, la quale nel caso *Golder v. the United Kingdom* ha per la prima volta riconosciuto che «*[i]t would be inconceivable*», che l'art. 6 § 1 «*should describe in detail the procedural guarantees afforded to parties in a pending lawsuit and should not first protect that which alone makes it in fact possible to benefit from such guarantees, that is, access to a court*»⁷⁰¹.

Se il giusto processo garantisce il diritto di accesso al giudice, quest'ultimo deve garantire che le parti abbiano la possibilità di tutelare i propri diritti in maniera *effettiva*.

È qui che sta il cuore del ragionamento della Corte e il collegamento tra tutela giurisdizionale e diritti di difesa (in senso lato): perché abbiano un senso le garanzie previste dall'art. 6 CEDU è necessario che alle parti sia garantita non solo la possibilità di adire un giudice, ma è necessario che tale diritto sia «*practical and effective*»⁷⁰². Ciò significa che una volta avuto accesso al giudice, un individuo deve avere «*a clear, practical opportunity to challenge an act that is an interference with his rights*»⁷⁰³.

Non dissimile, peraltro, è la formulazione dell'art. 47 CDFUE che, come già visto, prevede il ricorso effettivo davanti a un giudice (§ 1) la cui effettività deve essere vagliata alla luce dei requisiti dell'equità del procedimento previsti nel § 2 del medesimo articolo.

Infine, il diritto alla tutela giurisdizionale e alla difesa, come noto, sono anche garantiti all'art. 24 della Costituzione che ai primi due commi stabilisce che «Tutti possono agire in giudizio per la tutela dei propri diritti e interessi legittimi» (primo comma) e che «La difesa è diritto inviolabile in ogni stato e grado del procedimento» (secondo comma).

⁷⁰¹ Corte EDU, *Golder v. the United Kingdom*, 1975, §§ 28-36.

⁷⁰² Corte EDU, *Zubac v. Croatia* [GC], 2018, §§ 76-79.

⁷⁰³ Corte EDU, *Bellet v. France*, 1995, § 36.

Senza qui voler dare conto di tutte le interpretazioni e evoluzioni di tale articolo⁷⁰⁴, è utile solo vedere come la Corte costituzionale ha interpretato tali tutele nel tempo e come ha riconosciuto la necessaria *effettività* della tutela, anche grazie all'influenza proprio della giurisprudenza della Corte europea dei diritti.

Già con la sentenza n. 98 del 1965 la Corte ha riconosciuto che la tutela giurisdizionale è un diritto «tra quelli inviolabili dell'uomo, che la Costituzione garantisce all'art. 2, come si arguisce anche dalla considerazione che se ne è fatta nell'art. 6 della Convenzione europea dei diritti dell'uomo»⁷⁰⁵.

Qualche anno dopo, con la sentenza n. 18 del 1982 la Corte ha ascrivito la tutela giurisdizionale «tra i principi supremi del nostro ordinamento costituzionale, in cui è intimamente connesso con lo stesso principio di democrazia l'assicurare a tutti e sempre, per qualsiasi controversia, un giudice e un giudizio»⁷⁰⁶.

Finalmente, in una prospettiva di effettività della tutela, derivata anche dall'ormai prolifica giurisprudenza della Corte EDU, la Corte ha precisato che «[a]l riconoscimento della titolarità di diritti non può non accompagnarsi il riconoscimento del potere di farli valere innanzi a un giudice in un procedimento di natura giurisdizionale [...] L'azione in giudizio per la difesa dei propri diritti, d'altronde, è essa stessa il contenuto di un diritto, protetto dagli articoli 24 e 113 della Costituzione e da annoverarsi tra quelli inviolabili, riconducibili all'art. 2 della Costituzione un diritto che non si lascia ridurre alla mera possibilità di proporre istanze o sollecitazioni, foss'anche ad autorità appartenenti all'ordine giudiziario, destinate a una trattazione fuori delle garanzie procedurali minime costituzionalmente dovute, quali la possibilità del contraddittorio, la stabilità della decisione e l'impugnabilità con ricorso per cassazione»⁷⁰⁷.

⁷⁰⁴ Assumiamo qui l'interpretazione *ut magis valeat*, come suggerito da A. SIMONCINI, *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*, cit. p. 406. In generale sull'art. 24 Cost. e le sue diverse interpretazioni si v. A. POLICE, *Art. 24*, in R. BIFULCO - A. CELOTTO - M. OLIVETTI (a cura di), *Commentario alla Costituzione*, Utet giuridica, Torino, 2006, pp. 502 ss.

⁷⁰⁵ Corte Cost., sent. n. 98 del 1965, Punto 2 del *Considerato in diritto*.

⁷⁰⁶ Corte cost., sent. n. 18 del 1982, punto 4 del *Considerato in diritto*.

⁷⁰⁷ Corte cost., sent. n. 26 del 1999, punto 3.1 del *Considerato in diritto*. Nel medesimo senso, più recentemente, Corte cost., sentt. nn. 120 del 2014, 386 del 2004 e 29 del 2003.

La Corte ha infine sostenuto che «il diritto al giudice ed a una tutela giurisdizionale effettiva dei diritti inviolabili è sicuramente tra i grandi principi di civiltà giuridica in ogni sistema democratico del nostro tempo»⁷⁰⁸.

E se è vero che il diritto a una tutela giurisdizionale effettiva non è, come del resto tutti i diritti, sottratta al bilanciamento nell'interpretazione della Corte costituzionale⁷⁰⁹, né a possibili limitazioni nell'interpretazione della Corte europea dei diritti⁷¹⁰ e della Corte di Giustizia dell'Unione europea⁷¹¹, le esperienze oggi condotte di giustizia algoritmica non sembrano oggi poter contrapporre al diritto fondamentale alla tutela giurisdizionale diritti o interessi di forza tale da contrarre tale diritto.

In definitiva, non basta garantire alle parti la possibilità di accedere alla tutela giurisdizionale – che comunque non sembra essere compromessa nelle ipotesi che abbiamo delineato di giustizia algoritmica, in quanto si tratta di strumenti impiegati dal giudice già all'interno del processo – ma tale tutela non deve essere illusoria.

Al tal riguardo, per concludere sembra potersi affermare, dopo aver messo in luce le possibili interferenze della giustizia algoritmica con il *due process of law* e con il *fair trial*⁷¹², che allo stato attuale il diritto a una tutela giurisdizionale *effettiva*, e cioè la possibilità di difendere in giudizio i propri diritti (compreso ovviamente il diritto all'impugnazione della decisione), sia seriamente compromesso. L'impossibilità di comprendere il funzionamento degli algoritmi (sia da parte del giudice che delle parti), l'autorità algoritmica e la sostanziale esternalizzazione della funzione giurisdizionale a soggetti estranei alla giurisdizione, il mancato

⁷⁰⁸ Corte cost., sent. n. 238 del 2014, punto n. 3.4 del *Considerato in diritto*. Di recente, Corte cost., sent. n. 48 del 2021, punto n. 4.3 del *Coniserato in diritto*.

⁷⁰⁹ Corte cost., sent. n. 85 del 2013.

⁷¹⁰ Si cfr. *ex multis*, Corte EDU, *Ali Riza v. Switzerland*, 2021, §§ 94-96 e *Stanev v. Bulgaria* [GC], 2012, § 230.

⁷¹¹ Si cfr. *ex multis*, Corte di Giustizia, sent. 9 novembre 2010, C-92/09 e C-93/09.

⁷¹² Avendo deciso di affrontare la questione della compatibilità della giustizia algoritmica con il processo *in generale*, abbiamo deciso di non affrontare direttamente le garanzie del processo penale, quali la presunzione di innocenza e il diritto al silenzio, che come noto hanno una loro specificità. Per un approfondimento di tali problematiche si rinvia, *ex multis*, a S. QUATTROCOLO, *Equità' del processo penale e automated evidence alla luce della convenzione europea dei diritti dell'uomo*, cit.; V. MANES, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, cit., p. 15; G. CONTISSA - G. LASAGNI - G. SARTOR, *Quando a decidere in materia penale sono (anche) algoritmi e IA: alla ricerca di un rimedio effettivo*, in *Diritto di internet*, n. 4/2019, pp. 619-634.

controllo sui dati raccolti e implementati negli algoritmi da parte di chi utilizza i sistemi algoritmici e la carenza intrinseca della motivazione di un provvedimento assunto con l'ausilio di tali sistemi sembrano pregiudicare l'effettività della tutela giurisdizionale e impongono, quindi, grande cautela e prudenza nell'accogliere soluzioni di giustizia algoritmica negli ordinamenti europei e nell'ordinamento italiano.

Ciò detto, nella ormai lunga elaborazione legislativa (di cui abbiamo parlato in termini generali nella seconda parte del lavoro, giurisprudenziale (non solo negli Stati Uniti ma anche in Italia, sebbene in altri contesti) e dottrinale, si possono individuare già delle tutele e rimedi idonei ad accompagnare con gli strumenti del diritto l'attuazione di tale profondo cambiamento che sia compatibile con le garanzie del giusto processo.

Nell'ultimo capitolo di questo lavoro cercheremo di delineare tali possibili soluzioni, quelle già messe in atto dalla legislazione e quelle che riteniamo necessarie per raggiungere quello che abbiamo definito il "giusto processo algoritmico".

PARTE IV

CONCLUSIONI

Nota introduttiva: sciogliere i nodi

Non è compito facile, dopo la trattazione svolta, tentare di ricucire il discorso in modo tale da renderlo organico, lineare. Non è facile perché la giustizia algoritmica è un tema complesso, ricco di sfaccettature, indagabile da numerose prospettive, e mai in attesa che qualcuno lo studi prima di mutare, di assumere forme ancora diverse. Anche nella presente ricerca si è potuta toccare con mano l'urgenza di adattare il discorso al mutato panorama tecnologico.

Se prima la riflessione era esclusivamente basata su strumenti algoritmici e di intelligenza artificiale che erano progettati per svolgere un singolo compito (per esempio i *risk assessment tools* o gli strumenti di giustizia predittiva), nel 2022 è comparsa nel mercato l'IA generativa *con finalità generali* che ha inevitabilmente sparigliato le carte e ci ha indotto a riflettere anche sull'impiego di tale IA nel processo. Ancora, la ricerca è cominciata quando mancava del tutto una regolazione negli Stati Uniti e, in Europa, era appena stata approvata la Proposta di Regolamento sull'intelligenza artificiale (aprile 2021). Alla fine del percorso di ricerca, solo limitandoci a guardare l'ordinamento europeo e quello statunitense, da una parte vediamo gli Stati Uniti che hanno approvato un *Executive order* sull'IA e stanno piano piano ragionando su una legge federale sull'intelligenza artificiale, dall'altra un'Europa che non solo ha approvato un Regolamento vincolante per tutti gli Stati membri, ma ha persino elaborato una Convenzione quadro sull'IA che abbia al centro il rispetto dello stato di diritto, oggi sottoscritta anche dagli Stati Uniti.

Tutto ciò per esporre una verità banale: la presente ricerca tenterà, nell'ultimo capitolo, di mettere un punto all'indagine, muovendo dalle problematiche emerse nel corso del lavoro. L'obiettivo sarà quello di fornire alcune prospettive realistiche di impiego di algoritmi e di intelligenza artificiale all'interno dei sistemi giudiziari

che siano compatibili con i principi del giusto processo e in generale con le garanzie costituzionali. Ciò nonostante, giustizia e algoritmi vivono un rapporto di *equilibrio mutevole*⁷¹³, ed è facile pensare che in breve tempo la giustizia algoritmica avrà assunto forme così diverse da rendere persino obsoleto il presente lavoro.

A questo rischio si oppongono almeno due cinture di sicurezza: la prima è la speranza che la costituzionalizzazione delle problematiche emerse sia riuscita a stimolare una riflessione che vada oltre le singole esperienze di giustizia algoritmica. In tal senso, anche se in breve tempo toccheremo nuove frontiere dell'IA, ci sembra di poter affermare che almeno alcuni dei rischi per la tenuta delle garanzie costituzionali siano formulati in modo tale da reggere le possibili evoluzioni della tecnologia. La seconda è legata a come, almeno in Europa, la giustizia algoritmica si sta affermando nelle aule giudiziarie: con grande *prudenza*. Al contrario dell'impiego di algoritmi in altri settori, probabilmente per il confluire di diversi fattori – consapevolezza dei diritti in gioco, mancanza di risorse, mancanza di aggiornamento professionale dei giudici – in campo giurisdizionale si percepisce ancora molto timore, spesso giustificato, di integrare l'attività del giudice con strumenti algoritmici, il che ha condotto a non investire particolari risorse (economiche e intellettuali) su tale impiego. In realtà, come si dirà nell'ultimo capitolo, esistono dei sentieri sicuri per guidare la giustizia algoritmica all'interno del sistema giustizia, per poter sfruttare le potenzialità di una tecnologia che in un modo o nell'altro pervaderà ogni aspetto delle nostre vite.

⁷¹³ G. DE MINICO, *Giustizia e intelligenza artificiale: un equilibrio mutevole*, cit.

CAPITOLO IX

IL GIUSTO PROCESSO ALGORITMICO E IL RINNOVATO RUOLO DEL GIUDICE

SOMMARIO: 1. Algoritmi e giusto processo: un intreccio possibile - 2. Controllo dell'algoritmo e dei dati da parte del giudiziario - 3. La trasparenza algoritmica come tassello irrinunciabile per il rispetto del giusto processo - 3.1. Obblighi di trasparenza per un *alghoritmic due process* negli Stati Uniti - 3.2. Obblighi di trasparenza per un giusto processo algoritmico in Europa e in Italia - 4. Alfabetizzazione e autonomia del giudice - Un tentativo (parziale ma ambizioso) di conclusione: la giustizia algoritmica come opportunità

1. Algoritmi e giusto processo: un intreccio possibile

Partendo dall'insegnamento della giurisprudenza americana sulla compatibilità degli algoritmi con il *due process of law* e, in generale, con il potere discrezionale del giudice, abbiamo indagato le problematiche ad oggi più sotto scrutinio della giustizia algoritmica alla luce delle garanzie del giusto processo in chiave costituzionalistica europea. In quest'ultimo capitolo seguiremo un'impostazione simile: seppure con strumenti diversi, derivati dalle differenze tra i due ordinamenti, la dottrina e la giurisprudenza statunitensi si sono interrogate su come addivenire a un *technological due process*⁷¹⁴, tentando di risolvere i problemi che abbiamo delineato in questo lavoro, dall'opacità algoritmica al ruolo dei privati fino all'*automation bias*.

Partendo dalle soluzioni studiate oltreoceano, proveremo a riflettere su quali strumenti e guarentigie sarebbero a nostro avviso necessarie per accompagnare la giustizia algoritmica negli ordinamenti europei e renderla compatibile con le garanzie costituzionali che abbiamo considerato nel corso del lavor.

Alcune di tali soluzioni sono state già poste sul tavolo dalla regolazione europea, in particolar modo dal Regolamento europeo sull'IA, che ha adottato delle strategie normative (non sempre convincenti) per risolvere il problema dell'opacità

⁷¹⁴ D.K. CITRON, *Technological Due Process*, cit.

algoritmica e per imporre a chi impiega i sistemi di intelligenza artificiale, come nel nostro caso l'organo giudiziario, di adottare delle procedure di garanzia per chi sarà soggetto a una decisione assunta anche attraverso procedure automatizzate.

Inoltre, seppure in altro campo (quello delle decisioni amministrative), in Italia la giurisprudenza del Consiglio di Stato ha dovuto confrontarsi con l'impiego di algoritmi da parte della pubblica amministrazione, delineando uno *statuto costituzionale* dell'algoritmo. Le decisioni dei supremi giudici amministrativi si spingono decisamente oltre il *thema decidendum*, fornendo delle coordinate interpretative che non possono essere ignorate se si indaga il tema dell'impiego di algoritmi da parte del giudice.

Su questi dati di realtà, quindi, proveremo a delineare quella che oggi appare una possibile via verso il *giusto processo algoritmico*.

Ciò porterà con sé, inevitabilmente, una riflessione sul rinnovato ruolo da parte del giudice, che oggi non sembra ancora avere, come vedremo, gli strumenti per essere egli stesso pilota (o quantomeno copilota) di un cambiamento tanto incisivo, anche se non sappiamo ancora quanto imminente. Ciò non significa, peraltro, che attraverso un'elaborazione teorica, che si traduca poi in riforme pratiche, il giudice non possa essere all'altezza di un tale processo. Anzi, dato il ruolo sempre più importante assunto dalla giurisdizione nel sistema giuridico disorganico di oggi, ci sembra che sia proprio il giudice a poter essere protagonista di un impiego costituzionalmente legittimo della giustizia algoritmica, realizzando un auspicato intreccio tra intelligenza umana e intelligenza artificiale⁷¹⁵.

L'impostazione dei paragrafi seguenti seguirà il percorso sin qui delineato, muovendo quindi dai problemi principali segnalati: il controllo sull'algoritmo e sui dati, l'opacità algoritmica e l'*automation bias*.

2. Controllo del giudiziario sull'algoritmo e sui dati

⁷¹⁵ Su tale visione ottimistica si cfr. A. Z. HUQ, *A Right to a Human Decision*, in *105 Virginia Law Review*, 2020.

Chi dovrebbe controllare la progettazione dell'algoritmo e selezionare i dati con cui alimentare il *dataset* di algoritmi impiegati nella giustizia? Pensare a una soluzione “*in house*” in cui tutto viene progettato e gestito all'interno dei sistemi giudiziari pare, se non utopico, quantomeno ancora lontano dalla realtà, per la mancanza di risorse economiche e professionali.

Allo stato attuale di giustizia algoritmica abbiamo analizzato diverse ipotesi: i *risk tools* impiegati nel processo statunitense, progettati da imprese private (le quali selezionano il *dataset* e stipulano contratti con l'esecutivo) e utilizzati senza alcun controllo da parte dei giudici; strumenti di giustizia predittiva progettati da società private a cui l'esecutivo aveva fornito i dati (come l'esempio francese testato nelle Corti di Rennes e Douai e l'esperimento italiano di Prodigit); strumenti di giustizia predittiva progettati e implementati solo dal Ministero della Giustizia (esperimento *Datajust* in Francia); progetti ed esperimenti in collaborazione tra singoli uffici giudiziari e Università (quasi tutti quelli oggi in studio in Italia, si v. *supra* capitoli I e II).

Ora, in tutti i casi abbiamo visto che è facile che si creino dei cortocircuiti rispetto ai principi di indipendenza e imparzialità del giudice, sia quando i sistemi sono progettati da società private, sia quando i dati sono controllati esclusivamente dall'esecutivo (in particolare dal Ministero della Giustizia).

Indipendentemente dal soggetto estraneo alla magistratura che interviene nella progettazione degli algoritmi e nell'inserimento dei dati, sembra irrinunciabile mantenere un controllo costante e significativo da parte di soggetti interni alla magistratura per poter impiegare strumenti algoritmici che incidono sulla decisione del giudice.

Una siffatta prospettiva è peraltro coerente con alcune delle considerazioni svolte dagli stessi giudici sul tema della giustizia algoritmica. Come si è visto nell'esempio francese, dopo l'esperimento fallito di impiego di strumenti predittivi in due Corti d'Appello (si v. *supra* capitolo II), nei successivi colloqui tra magistrati e avvocati, i primi hanno segnalato l'importanza che sia la stessa autorità giudiziaria a organizzare le banche dati e, in generale, i dati giudiziari che potenzialmente

servono a nutrire sistemi di giustizia predittiva⁷¹⁶. Non dissimile la posizione del CSM espressa nell'ultima relazione sulla giustizia (si v. *supra* capitolo II § 2.3), in cui l'organo di autogoverno ha espresso la necessità che la diffusione dei dati giudiziari debba essere monitorata con attenzione in modo da valutarne di volta in volta le importanti ricadute organizzative e l'impatto che possano avere sul lavoro dei giudici.

Sembra quindi che da parte della magistratura ci sia la volontà di mantenere un ruolo di controllo su strumenti che essa stessa sarà chiamata ad utilizzare.

In Italia, occorre chiedersi se il controllo della magistratura – e in particolare del CSM – sull'impiego di strumenti algoritmici sia coerente col testo del disegno di legge oggi in discussione in Parlamento (si v. *supra* capitolo V § 4.4), il quale prevede che sia il Ministero della Giustizia, e solo lui, a disciplinare l'impiego di sistemi di intelligenza artificiale da parte della magistratura ordinaria (art. 14). Probabilmente una scelta maggiormente coerente e più adatta al contesto algoritmico sembra essere quella di coinvolgere la magistratura nella disciplina di tali sistemi, proprio per le problematiche che abbiamo segnalato nei capitoli precedenti. Non si può prevedere come verrà riorganizzata l'amministrazione della giustizia quando i sistemi automatizzati faranno il loro ingresso in maniera strutturale nei panorami giudiziari, ma la speranza è che siffatto ingresso non si trasformi in un nuovo terreno di scontro tra Ministro della Giustizia e CSM e che, anzi, i loro rapporti in un contesto che sarà particolarmente delicato per garantire l'efficienza del sistema giudiziario siano guidati dal principio, più volte espresso dalla Corte costituzionale, di leale collaborazione⁷¹⁷.

Per riassumere, una possibile soluzione al problema della progettazione degli algoritmi e dell'inserimento dei dati è garantire che tali attività rimangano soggette

⁷¹⁶ B. LOUVEL, *Communication à l'occasion du colloque pour le bicentenaire de l'ordre des avocats au Conseil d'État et à la Cour de Cassation*, in *Ordre des avocats au Conseil d'État et à la Cour de cassation, La Justice prédictive*, cit.

⁷¹⁷ Si cfr., tra le tante, Corte cost. nn. 379 del 1992 e 380 del 2003. In dottrina si cfr., *ex multis*, G. SILVESTRI, *Consiglio superiore della magistratura e sistema costituzionale*, in *Questione Giustizia*, trimestrale n. 4 del 2017, pp. 19-29.

a un controllo pubblico⁷¹⁸, che non sembra poter essere esercitato esclusivamente dall'Esecutivo, per i problemi di indipendenza che abbiamo tentato di delineare in questo lavoro, ma che deve coinvolgere necessariamente soggetti appartenenti alla magistratura.

Le modalità pratiche in cui ciò potrebbe avvenire non sono ad oggi facilmente ipotizzabili: certamente non è peregrino pensare all'istituzione di *pool* di funzionari pubblici con garanzie di indipendenza, che operino a stretto contatto con i giudici, altamente preparati e in grado di progettare, in autonomia o con l'ausilio di soggetti e investimenti privati, algoritmi trasparenti e diretti esclusivamente a un impiego giurisdizionale al cui interno inserire i dati controllati e valutati dalla stessa magistratura.

Un'impostazione di tal guisa sembra peraltro coerente con le indicazioni previste nella Carta etica, il cui terzo principio di qualità e sicurezza dei dati, come visto (*supra* capitolo V § 6), richiede un approccio quanto più possibile interdisciplinare e plurale che coinvolga anche le competenze dei giudici, che non devono rimanere estranei alla progettazione di algoritmi e scelta dei dati.

In ogni caso, qualunque siano le modalità che verranno attuate in Europa e in Italia, il Regolamento sull'IA impone alcune guarentigie utili per la *governance* dei dati dei sistemi ad alto rischio, come quelli che si utilizzeranno nei sistemi giudiziari.

Secondo l'art. 10 del Regolamento, i dati impiegati per l'addestramento, la convalida, e la prova di sistemi di IA dovranno essere, tra le altre cose, «sufficientemente rappresentativi e, nella misura del possibile, esenti da errori e completi nell'ottica della finalità prevista. Essi possiedono le proprietà statistiche appropriate anche, ove applicabile, per quanto riguarda le persone o i gruppi di persone relativamente ai quali il sistema di IA ad alto rischio è destinato a essere usato» (§ 3). Inoltre, tali dati devono tenere conto, «nella misura necessaria per la finalità prevista, delle caratteristiche o degli elementi particolari dello specifico

⁷¹⁸ Così G. CARLOTTI, *La giustizia predittiva e le fragole con la panna*, in *Giustizia amministrativa*, 2023, p. 11.

ambito geografico, contestuale, comportamentale o funzionale all'interno del quale il sistema di IA ad alto rischio è destinato a essere usato» (§ 4).

Qualora il controllo dei dati, come qui sostenuto, fosse di competenza (anche) di un'articolazione dell'ordine giudiziario, il Regolamento specifica che «nella misura in cui esercita il controllo sui dati di input, il deployer garantisce che tali dati di input siano pertinenti e sufficientemente rappresentativi alla luce della finalità prevista del sistema di IA ad alto rischio» (art. 26 § 4).

Tali cautele, anche se ancora non sappiamo come verranno implementate nei singoli Stati membri, potrebbero sicuramente venire incontro ai segnalati problemi sia in riferimento alla natura non rappresentativa dei dati impiegati nei sistemi di giustizia algoritmica negli Stati Uniti⁷¹⁹, sia con riguardo alla scelta dei dati per addestrare sistemi di giustizia predittiva⁷²⁰. Ovviamente non è sufficiente e, anzi, il controllo pubblico sui dati e sugli algoritmi – per come qui delineato – non risolve del tutto i problemi di opacità algoritmica segnalati né la garanzia che permanga l'autonomia decisionale del giudice.

3. La trasparenza algoritmica come tassello irrinunciabile per il rispetto del giusto processo

La trasparenza algoritmica è stata definita come la “pietra angolare” (*cornerstone*) dello Stato di diritto, fondamentale per garantire il rispetto delle garanzie del giusto processo⁷²¹. Non considerando tutti gli altri nodi che abbiamo delineato in questo lavoro, già solo decisioni giudiziarie basate su informazioni non conoscibili non sarebbero, intrinsecamente, conformi alle garanzie che reggono la

⁷¹⁹ Spiegando il funzionamento di COMPAS, abbiamo segnalato che i dati utilizzati non erano rappresentativi, perché addestravano un sistema impiegato in diverse giurisdizioni con lo stesso *dataset*, senza distinzioni che oggi sono richieste dall'*AI Act* (si cfr. *supra* capitolo I § 4).

⁷²⁰ Abbiamo visto che la scelta dei dati da implementare in sistemi di giustizia predittiva deve essere adeguata, e cioè riflettere gli orientamenti maggioritari in una determinata materia (capitolo VI § 1) Anche per questo sembra che un controllo da parte di soggetti appartenenti alla magistratura sia particolarmente importante per garantire l'adeguatezza dei dati impiegati.

⁷²¹ M. EBERS - P. K. TUPAY, *Due Process, Fair Trial, Transparency, and Explainability*, in *ID, Artificial Intelligence and Machine Learning Powered Public Service Delivery in Estonia: Opportunities and Legal Challenges*, 2023, p. 105.

giurisdizione. La legittimità di una decisione del giudice, in altre parole, dipende in gran parte dalla trasparenza del processo decisionale che ha portato a tale decisione, soprattutto quando essa incide sui diritti e sugli obblighi degli individui.

La centralità della trasparenza, del resto, è restituita dal fatto che quasi ogni strumento normativo che abbiamo analizzato nella seconda parte del lavoro contempla la trasparenza algoritmica come principio chiave nella possibile introduzione di sistemi algoritmici nella giurisdizione, anche se quasi mai è esplicitata la soglia di trasparenza richiesta.

In questo senso, nei prossimi paragrafi, partendo dalle soluzioni elaborate dalla giurisprudenza e dottrina statunitensi, cercheremo di indagare alcune soluzioni possibili rispetto a un tema così importante nell'articolazione di questo lavoro.

3.1. Obblighi di trasparenza per un alghoritmico due process negli Stati Uniti

Abbiamo visto che la contestazione più frequente sollevata dagli imputati soggetti alla valutazione del rischio di recidiva è quella di non poter accedere all'algoritmo predittivo e in definitiva di conoscere come quest'ultimo produce il proprio *output*. Ciò principalmente per l'opposizione, da parte della società produttrice dell'algoritmo, del segreto commerciale.

La necessità di conoscere il funzionamento dell'algoritmo utilizzato è stata di recente riconosciuta, nel caso *Flores v. Stanford*⁷²², da una Corte federale che ha tentato di bilanciare gli interessi della società produttrice dell'algoritmo con le necessità di pubblicità e di trasparenza inerenti alla giustizia penale. Sebbene non si tratti di un caso in cui COMPAS veniva impiegato nel *sentencing* ma nella fase esecutiva della pena, le conclusioni cui giunge la Corte assumono notevole importanza perché mostrano come le Corti statunitensi stiano gradualmente tentando di bilanciare gli interessi in gioco nell'ambito di decisioni assunte tramite algoritmi⁷²³, bilanciamento che sarà interessante confrontare con le soluzioni

⁷²² *Flores v. Stanford*, WL 4441614, 2021

⁷²³ Ci sono diversi casi in cui le Corti americane hanno sindacato l'impiego di algoritmi in settori più disparati, principalmente quando essi sono stati utilizzati dalle articolazioni del Governo

proposte dalla giurisprudenza e dalla dottrina europee (si v. *infra*). In tale sentenza, la Corte distrettuale federale di New York era chiamata a valutare la richiesta di Northpointe (società produttrice di COMPAS, ora Equivant), di impedire la divulgazione di alcuni elementi proprietari di COMPAS al perito nominato dai ricorrenti. Il caso seguiva una precedente decisione della stessa Corte del 12 febbraio 2021 che, in un procedimento tra alcuni ricorrenti contro lo stato di New York aveva ordinato a Northpointe di produrre in giudizio alcune informazioni protette da *trade secret* riguardanti COMPAS (i «*Compelled Materials*»⁷²⁴), il quale nello stato di New York viene correntemente utilizzato dalle commissioni costituite all'interno delle Corti per concedere agli ergastolani la libertà condizionale (cd. *parole decision*). Per evitare di sacrificare del tutto i diritti proprietari di Northpointe, la Corte ordinava che le informazioni venissero prodotte tramite quello che viene definito un *protective order*⁷²⁵, e quindi svelate solo ed esclusivamente a un esperto nominato dalle parti.

Nonostante le condizioni concordate del *protective order*, la società non consentiva l'accesso alle informazioni e si rivolgeva nuovamente al tribunale, affermando che i materiali da svelare costituivano le più importanti informazioni relative a COMPAS e che «*permitting Dr. Rudin [il perito] access to [it] ... would threaten [Northpointe's] very existence*». Ciò in quanto – secondo la difesa di Northpointe – il perito era un professore di *computer science* e *computer engineering* che a lungo aveva criticato l'impiego di COMPAS nelle decisioni pubbliche e aveva pubblicato numerosi articoli sull'opacità di siffatto algoritmo.

La Corte precisa dapprima che non vi è un privilegio assoluto contro il disvelamento di informazioni confidenziali e segreti commerciali⁷²⁶ e anzi

federale. Per un'indagine quantitativa e qualitativa di siffatta giurisprudenza si cfr. G. G. CUSENZA, *Litigating Governmental Use Of AI*, in *Ohio State Technology Law Journal*, 2024.

⁷²⁴ In particolare veniva ordinato di svelare il *dataset* utilizzato per addestrare COMPAS e il modello algoritmico alla base del funzionamento delle scale di rischio (per tali concetti si v. *supra* capitolo I § 4).

⁷²⁵ Il *protective order* è uno strumento che consente di bilanciare i diritti di accesso delle parti all'algoritmo con i diritti di privacy delle società, imponendo lo svelamento dell'algoritmo solo ad alcuni soggetti, con l'obbligo di questi ultimi di non rivelare ad altri il contenuto.

⁷²⁶ *Flores v. Stanford*, cit., § 4.

if a court finds that commercially sensitive materials are relevant to a party's claim, the court balances that party's need to have its chosen expert review the materials against the producing party's interest in protecting the materials from a potential competitor.

Il bilanciamento, secondo la Corte, deve considerare i seguenti fattori: 1) se la persona destinataria delle informazioni è coinvolta in ambiti concorrenziali con il materiale svelato; 2) il rischio di una divulgazione involontaria delle informazioni; 3) le difficoltà che eventualmente scaturiscono per la società dal *protective order*; 4) il tempismo e l'utilità dello svelamento (*time and scope of the remedy*).

Secondo la Corte, le informazioni da svelare sono «*relevant to the merits of Plaintiffs' constitutional claims*»⁷²⁷: ciò in quanto la Commissione (*Parole Board*) è chiamata a considerare per le proprie decisioni anche lo *score* generato da COMPAS che tuttavia alcuni studi hanno dimostrato essere discriminatorio, in quanto talvolta considera la giovane età quale un fattore aggravante e quindi priva condannati ergastolani giovani della possibilità che venga concessa loro la libertà condizionale dimostrando la propria maturità e il successo del percorso riabilitativo. Ciò costituirebbe, secondo la Corte, una violazione del divieto di applicare pene crudeli e inusitate (VIII emendamento) e, ovviamente, del giusto processo (V e XIV emendamento).

Proprio perché i ricorrenti lamentano il peso che determinati fattori assumono per la generazione del rischio di recidiva, senza tuttavia avere accesso a tali informazioni proprietarie, «*more information is necessary to support Plaintiffs' claim that COMPAS's risk models violate young offenders' constitutional rights*» e quindi «*Dr. Rudin's analysis of these documents may help Plaintiffs substantiate their allegations that COMPAS punishes juvenile offenders for their youth, such that Defendants' reliance on this tool is constitutionally problematic*»

Alla contestazione di Norhpointe secondo cui le informazioni sarebbero irrilevanti in quanto solo uno tra i ricorrenti aveva ricevuto un punteggio di rischio alto, la Corte risponde, al contrario, che

⁷²⁷ *Ivi*, § 6.

*all of the named Plaintiffs maintain that Defendants' reliance on COMPAS is constitutionally inappropriate because COMPAS's algorithm inherently discriminates against young offenders and is based on inaccurate statistics,*⁷²⁸

e quindi non occorre che sia stato determinato per ogni singolo ricorrente un rischio di recidiva alto, posto che altri fattori considerati da COMPAS potrebbero aver diminuito lo *score* finale.

Infine, la Corte considera l'interesse della società produttrice ad evitare danni di natura concorrenziale che deriverebbero dalla divulgazione al pubblico delle informazioni riservate ma conclude che tale il rischio di tale divulgazione sia basso per diverse ragioni: il *protective order* vieta la divulgazione delle informazioni e ne limita l'uso alla controversia in corso; il perito ha accettato sotto giuramento di essere vincolato agli *orders* e di sottoporsi alla giurisdizione della Corte per il loro rispetto e per qualsiasi altro provvedimento relativo agli stessi; si è inoltre impegnato a adottare ulteriori misure per prevenire la divulgazione involontaria e ha anche in passato gestito informazioni riservate quando è stato chiamato in qualità di perito in altri procedimenti penali. In definitiva, non vi sono informazioni rilevanti che mettano in discussione l'integrità del perito⁷²⁹.

In conclusione, la Corte non accoglie la richiesta di Northpointe ma invita le parti (i ricorrenti e la stessa Northpointe) a presentare alla Corte un'ulteriore domanda di *protective order* che abbia i seguenti requisiti: a) i termini del *protective order* sono vincolanti per la il perito; b) quest'ultimo non può condividere o discutere le informazioni o il loro contenuto con nessuno tranne che con gli individui specificamente autorizzati; c) le informazioni riservate devono essere fornite al perito che deve restituirle a Northpointe al termine della causa; d) il perito non può copiare, scannerizzare o altrimenti duplicare o manipolare i materiali consegnatele; e) al termine del procedimento deve distruggere ogni altro documento e/o appunto

⁷²⁸ *Ivi*, § 7.

⁷²⁹ *Ivi*, §§ 8-11.

legato alle informazioni riservate; f) il perito non potrà usare le informazioni in nessun altro progetto e per nessun altro scopo che non sia il procedimento in corso.

La sentenza qui commentata è significativa in quanto mostra le modalità con cui una Corte ha tentato di alleviare le tensioni tra diritti contrapposti impiegando uno strumento, il *protective order*, il cui utilizzo per la divulgazione protetta dei segreti commerciali relativi agli algoritmi era in verità già stato paventato da alcuni autori⁷³⁰. Esso consente da un lato di proteggere la società produttrice del *software*, che riesce a mantenere il proprio vantaggio competitivo nei confronti dei concorrenti, dall'altro consente quanto meno alle parti interessate nel procedimento di conoscere il funzionamento dell'algoritmo attraverso l'esame da parte di un esperto.

Già prima di questa sentenza, in un caso del 2017 che non aveva a che vedere con l'uso dei *tools* nel processo ma con l'uso di un algoritmo proprietario da parte della pubblica amministrazione nel contesto scolastico, una Corte era stata chiamata a verificare la conformità della *legal black box* con le garanzie del giusto processo. È utile segnalare il caso poiché ci mostra ancora, seppur in altro contesto, il bilanciamento scelto dalle corti statunitensi tra gli interessi in gioco; inoltre, il caso suscita interesse in quanto – come si vedrà nel prosieguo – un caso simile che concerne decisioni algoritmiche nel contesto scolastico è sorto anche in Italia.

In *Houston Fed'n of Teachers, Local 2415 v. Houston Indep. School*⁷³¹ una *preliminary decision* della Corte distrettuale federale del Texas era chiamata a giudicare la legittimità dell'uso di algoritmi da parte di un distretto scolastico di Houston per valutare il rendimento dei propri insegnanti e per fornire al distretto una base conoscitiva per licenziare gli insegnanti che ottenevano valutazione insufficiente⁷³².

Il distretto scolastico – organismo pubblico – si era affidato a un fornitore esterno per sviluppare e gestire il suo sistema algoritmico, che considerava i suoi algoritmi

⁷³⁰ J. VILLASENOR - V. FOGGO, *Artificial intelligence, due process, and criminal sentencing*, cit., p. 296. Si v. anche D. WON, *The Missing Algorithm: Safeguarding Brady Against the Rise of Trade Secrecy in Policing*, in *Michigan Law Review*, 2021.

⁷³¹ *Hous. Fed'n of Teachers, Local 2415 v. Hous. Indep. Sch. Dist.*, 251 F. Supp. 3d 1168 (S.D. Tex. 2017).

⁷³² Sul caso si cfr. C. COGLIANESE - D. LEHR, *Transparency and Algorithmic Governance*, cit.

e il suo *software* come segreti commerciali, rifiutandosi di divulgarli sia al distretto sia agli insegnanti stessi⁷³³. Nove insegnanti e il sindacato locale contestavano quindi l'affidamento del distretto scolastico sugli *output* algoritmici per diversi motivi, tra cui violazione del giusto processo in quanto senza «*access to the computer algorithms and data necessary to verify the accuracy of their scores*», sarebbero violati i loro diritti costituzionali⁷³⁴.

Rigettando la richiesta di un *summury judgement* da parte del distretto scolastico – cioè una decisione assunta sulle prove già presentate e senza un processo dinanzi alla giuria – la Corte afferma che

*without access to... proprietary information – the value-added equations, computer source codes, decision rules, and assumptions – [the teachers'] scores will remain a mysterious 'black box,' impervious to challenge*⁷³⁵.

A differenza del caso *Flores* sopra citato, qui la Corte non pone in discussione la protezione del *trade secret* e anzi condivide l'argomentazione del distretto scolastico per cui the «*[d]ue Process Clause does not empower Plaintiffs to put [the vendor] out of business' by requiring the disclosure of its trade secret*». Invece, la Corte afferma che è incompatibile con il *due process* una *policy* che consenta a organismi pubblici di impiegare algoritmi segreti.

Il rimedio, quindi, non sarebbe lo svelamento del segreto commerciale quanto l'obbligo per il distretto scolastico di cercare un'alternativa che non contempli l'utilizzo di algoritmi segreti. In questo senso, la Corte afferma che

*When a public agency adopts a policy of making high stakes employment decisions based on secret algorithms incompatible with minimum due process, the proper remedy is to overturn the policy, while leaving the trade secrets intact*⁷³⁶.

⁷³³ *Ibidem*, § 1177.

⁷³⁴ *Ibidem*, § 1176.

⁷³⁵ *Ibidem*, § 1179.

⁷³⁶ *Ibidem*.

In tal caso, dunque, la soluzione per contemperare gli interessi in gioco non è stata quella di trovare le modalità con cui svelare alle parti il codice sorgente dell'algoritmo protetto da *trade secret*, quanto piuttosto l'esclusione *tout court* di algoritmi segreti dalla procedura pubblica.

Sul disvelamento del *trade secret* per venire incontro a diritti legati al giusto processo è utile qui rapidamente segnalare anche una recentissima sentenza che non ha a che vedere con algoritmi o intelligenza artificiale – e anzi ha come protagonista una delle forme più semplici di “tecnologia”, la matita – ma che da autorevole dottrina è stata individuata come indicativa di una tendenza che potrebbe affermarsi anche in decisioni, come *Flores v. Stanford*, il cui oggetto è il disvelamento dell'algoritmo⁷³⁷.

Nel caso deciso nella sentenza *Royal Brush Manufacturing v. United States*⁷³⁸, la Corte d'Appello federale, soffermandosi sul *Trade Secret Act*, legge che proibisce la diffusione di segreti commerciali se non autorizzati dalla legge, la Corte afferma perentoriamente: «*[w]e have no doubt that a release of information is “authorized by law” within the meaning of the Trade Secrets Act if that release is required as a matter of constitutional due process, as is the case here*». Anzi:

because the Constitution authorizes, and indeed requires, the release of confidential business information in this case, the Trade Secrets Act does not stand in the way of such release.

Se quindi il giusto processo, come afferma la Corte, include il diritto di sapere su quali prove si è fondata la decisione, allora «*the evidence used to prove the [g]overnment's case must be disclosed to the individual so that he has an opportunity to show that it is untrue*».

Su tali premesse la Corte, come in *Flores*, conclude indicando che le informazioni riservate dovevano venire dischiuse alla parte mediante un *protective order* che non ne avrebbe consentito la divulgazione al di fuori del processo giudiziario.

⁷³⁷ C. COGLIANESE, *AI, Due Process, and Trade Secrets*, in *The Regulatory Review*, 2023.

⁷³⁸ *Royal Brush Mfg. v. United States*, 75 F.4th 1250 (2023).

In definitiva, sebbene in quest'ultima sentenza analizzata non si discuta di disvelamento di un algoritmo, le conseguenze potrebbero essere significative per chiunque tenti di contestare, sulla base della protezione fornita dalla *due process clause*, l'uso di algoritmi di intelligenza artificiale da parte dei soggetti pubblici, tra cui i giudici.

Ora, se decisioni come quelle qui prese in esame sono significative in quanto mostrano una certa consapevolezza da parte dei giudici rispetto alle complicazioni che la natura proprietaria degli algoritmi provocano al rispetto del *due process*, non per questo dissipano i dubbi relativi all'impiego di algoritmi opachi all'interno dei procedimenti giudiziari, sia negli Stati Uniti sia in Europa.

Quanto al meccanismo del *protective order*, se è vero che è uno strumento utile per venire incontro alle esigenze di tutte le parti in causa, difficilmente potrebbe essere risolutivo nella giustizia penale "contaminata" da strumenti algoritmici.

Ciò in quanto non di rado le cautele richieste da un tale strumento – che di norma è utilizzato in controversie civili dove le parti in causa sono società con disponibilità economica elevata – richiedono un'ingente quantità di risorse che difficilmente potrebbero essere sostenute da un singolo imputato⁷³⁹. Tornando a *Flores*, basti pensare ai costi che un imputato dovrebbe sostenere per il pagamento dei servizi di un esperto in grado di analizzare e comprendere il contenuto di un codice algoritmico nonché i relativi costi di accesso ai *software* e di successiva distruzione: costi che, come è stato detto, sarebbero sostenibili piuttosto da organizzazioni non-profit e associazioni sindacali che si occupano di diritti umani nell'ambito della giustizia penale⁷⁴⁰. Tali organizzazioni potrebbero fornire le risorse necessarie a coprire i costi di ottemperanza al *protective order* e, inoltre, un'eventuale vittoria in giudizio sulla base di violazioni del giusto processo potrebbe giovare non solo ai singoli imputati rappresentati ma costituire un precedente anche per futuri imputati sottoposti alle medesime valutazioni di rischio.

⁷³⁹ J. VILLASENOR - V. FOGGO, *Artificial intelligence, due process, and criminal sentencing*, cit., p. 344.

⁷⁴⁰ *Ibidem*, p. 345.

È sulla scia di problemi di questo tipo che altri Autori ritengono che la *litigation strategy* più adeguata a garantire il rispetto del *due process* in decisioni assunte tramite algoritmi non sia da ricercare sul piano individuale, quanto piuttosto su quello collettivo. È solo attraverso azioni collettive – come le *class-action* – che il rimedio processuale potrebbe avere successo sia nel senso di precludere l’uso di strumenti di *machine learning* in assenza di dati verificati e appropriati, sia nel senso di catalizzare processi di analisi e riprogettazione in cui l’algoritmo non viene abbandonato ma piuttosto migliorato e reso conforme alle garanzie del giusto processo⁷⁴¹.

Ma vi è un altro, assorbente, motivo per cui singoli obblighi di *disclosure* – siano essi legislativi o, come i *protective order*, determinati da provvedimenti assunti in sede giurisdizionale – non forniscono da soli soluzioni soddisfacenti per fronteggiare la complessità procedimentale dell’intelligenza artificiale e, in ultimo, per scongiurare l’applicazione di sentenze basate su informazioni segrete e quindi non ammissibili.

Lo abbiamo definito come il problema della *technical black box*, e cioè delle caratteristiche dei moderni sistemi di IA il cui funzionamento non è facilmente comprensibile⁷⁴². In effetti, se venissero utilizzati sistemi di *machine learning* il cui accesso non consente comunque di verificare le modalità di assunzione della decisione, difficilmente sarebbe sufficiente consentire tale accesso per conformare siffatto impiego con le garanzie processuali.

In ogni caso, la dottrina statunitense più attenta ritiene che per raggiungere uno *standard* adeguato di giusto processo il primo obiettivo è quello di garantire la trasparenza *tout court* del processo decisionale. Alcuni ritengono che ciò andrebbe perseguito attraverso l’apertura del codice algoritmico (e quindi la soccombenza dei diritti di privativa) e la predisposizione di continui e indipendenti test sugli algoritmi impiegati, anche con la partecipazione attiva del pubblico⁷⁴³. Altri

⁷⁴¹ A. Z. HUQ, *Constitutional Rights in the Machine Learning State*, in *Cornell Law Review*, Vol. 105, 2020, pp. 50–52.

⁷⁴² Si v. *supra* capitolo VI § 2.

⁷⁴³ D.K. CITRON, *Technological Due Process*, cit., pp. 1308 ss. e anche D. K. CITRON - F. PASQUALE, (2014). *The scored society: Due process for automated predictions*, cit.

sottolineano che vi sono comunque opzioni intermedie tra la “*complete algorithmic secrecy*” e la “*complete public disclosure*” come, per esempio, lasciare a esperti indipendenti la valutazione del codice protetto da segreto (in fase processuale sarebbe lo strumento del *protective order*)⁷⁴⁴.

Altre proposte includono, se non l’apertura del codice, almeno la spiegazione alle parti in linguaggio comprensibile di come l’algoritmo ha assunto la decisione e la conseguente possibilità di impugnarla⁷⁴⁵ e altre che, indipendentemente dal mezzo, richiedono che i soggetti della decisione (quindi le parti) la “comprendano” (*understand*)⁷⁴⁶. Soprattutto quest’ultimo aspetto è messo in crisi dai problemi di comprensibilità che pongono gli algoritmi più evoluti, che paiono un ostacolo insormontabile al loro impiego nella giustizia.

Come si vede, rispetto all’ordinamento europeo in cui è intervenuto il legislatore (si v. *infra*) la tutela del *due process of law* negli Stati Uniti ad oggi è lasciata sostanzialmente nelle mani della giurisprudenza⁷⁴⁷, da cui la letteratura ha tratto spunto per estendere le tutele ritenute necessarie a garantire un *algorithmic due process of law*.

3.2. Obblighi di trasparenza per un giusto processo algoritmico in Europa e in Italia

Come ampiamente specificato, in Europa e in Italia non esistono ancora reali esperienze di giustizia algoritmica, sicché naturalmente manca una giurisprudenza

⁷⁴⁴ F. PASQUALE, *Secret Algorithms Threaten the Rule of Law*, in *MIT Technology Review*, 1 giugno 2017.

⁷⁴⁵ A. SELBST, S. BOROCAS, *The Intuitive Appeal of Explainable Machines*, cit.

⁷⁴⁶ M. BRENNER *et al.*, *Constitutional Dimensions of Predictive Algorithms in Criminal Justice*, cit. pp. 283 ss.

⁷⁴⁷ Occorre comunque segnalare l’imponente riforma in corso negli Stati Uniti attuata, sul piano federale, con il First Step Act del 21 dicembre 2018, che, tra le altre cose, ha autorizzato il Procuratore Generale degli Stati Uniti a progettare un nuovo tool. Quest’ultimo, annunciato nel 2019 e nominato PATTERN, è uno strumento predittivo istituzionale e quindi non coperto da segreto ma, ciò nonostante, non mancano le critiche e gli interrogativi rispetto al suo funzionamento sollevati sia in dottrina sia in giurisprudenza. Si cfr. B. L. GARRET - J. MONAHAN, *Judging Risk*, cit., pp. 484-486. Prima si v. anche T. SOURDIN, *Judge v Robot? Artificial intelligence and judicial decision-making*, in *UNSW Law Journal*, 41, 2018, pp. 1128 ss.

che ha vagliato la compatibilità degli algoritmi impiegati nella giustizia, in particolare sotto il profilo della loro opacità.

Ciò nonostante, sia le disposizioni di legge sia alcune decisioni giurisprudenziali ci possono offrire un quadro per comprendere i rimedi giuridici che potrebbero condurre verso un *fair trial algoritmico*.

Alcune prime indicazioni si possono trarre, con un cauto adattamento, dalla giurisprudenza che si è formata in seno alla giustizia amministrativa italiana – di primo grado e di appello – da cui è possibile rinvenire un coacervo di principi assunti a veste costituzionale, soprattutto sul tema della trasparenza.

La vicenda, molto nota⁷⁴⁸ e su cui non ci soffermeremo, ha tratto origine dalla legge 13 luglio 2015, n. 107 “*Riforma del sistema nazionale di istruzione e formazione e delega per il riordino delle disposizioni legislative vigenti*” (la cd. “Buona Scuola”), in forza della quale il Ministero dell’Istruzione dell’Università e della Ricerca ha disposto che le procedure di assegnazione e trasferimento dei docenti alle sedi di servizio fosse demandata a un sistema informatico automatizzato, progettato da HPE Service Italia S.r.l., società privata risultata vincitrice nella procedura di gara bandita dall’amministrazione.

La giurisprudenza amministrativa che ne è seguita ha meritoriamente ricostruito il rapporto tra decisore umano e algoritmo. In particolare, secondo le ultime pronunce del Consiglio di Stato la legittimità dell’impiego dell’algoritmo in sede decisionale è subordinata alla *conoscibilità* dello stesso «secondo una declinazione rafforzata del principio di trasparenza, che implica anche quello della piena conoscibilità di una regola espressa in un linguaggio differente da quello giuridico»⁷⁴⁹, a ciò non essendo sufficiente il mero accesso al codice sorgente

⁷⁴⁸ Tra i numerosi commenti si cfr. D.U. GALETTA, *Algoritmi, procedimento amministrativo e garanzie*, cit.; B. MARCHETTI, *La garanzia dello human in the loop alla prova della decisione amministrativa algoritmica*, in *BioLaw Journal – Rivista di BioDiritto*, n. 2/2021; L. MUSSELLI, *La decisione amministrativa nell’età degli algoritmi: primi spunti*, in *MediaLaws-Rivista di diritto dei media*, 1/2020; S. VERNILE, *Verso la decisione amministrativa algoritmica*, in *MediaLaws- Rivista di diritto dei media*, 2/2020. Sui profili costituzionali dell’amministrazione algoritmica si cfr. A. SIMONCINI, *Profili costituzionali della amministrazione algoritmica*, in *Riv. Trim. Dir. Pubbl.*, 4, 2019, pp. 1149 ss. e ID, *Amministrazione digitale algoritmica. Il quadro costituzionale*, in R. CAVALLO PERIN - D.U. GALETTA (a cura di), *Il diritto dell’amministrazione pubblica digitale*, Torino 2020.

⁷⁴⁹ Cons. Stato, Sez. VI, sent. n. 2270 del 2019.

dell'algoritmo – come aveva sostenuto il T.A.R. nel 2017⁷⁵⁰e poi ancora nel 2018⁷⁵¹– ed essendo invece necessaria la predisposizione di spiegazioni che traducano «la formula algoritmica in una regola giuridica che sia «leggibile e comprensibile, sia per i cittadini che per il giudice»⁷⁵².

Detta regola algoritmica, quindi, deve rimanere soggetta alla piena cognizione del giudice amministrativo, quale inevitabile «declinazione diretta del diritto di difesa del cittadino, al quale non può essere precluso di conoscere le modalità (anche se automatizzate) con le quali è stata in concreto assunta una decisione destinata a ripercuotersi sulla sua sfera giuridica»⁷⁵³. In continuità con quanto affermato, i Supremi Giudici amministrativi, in alcune sentenze successive⁷⁵⁴, mettono ulteriormente a punto lo 'statuto costituzionale' dell'algoritmo: innanzitutto la conoscibilità è da intendersi come *piena comprensibilità*, non potendosi dare rilievo alla «riservatezza delle imprese produttrici dei meccanismi informatici utilizzati i quali, ponendo al servizio del potere autoritativo tali strumenti, all'evidenza ne accettano le relative conseguenze in termini di necessaria trasparenza»⁷⁵⁵.

Inoltre, in armonia con quanto affermato dai giudici americani, si afferma il principio di non esclusività della decisione algoritmica, ricavato dall'art. 22 del GDPR⁷⁵⁶ cui si ricollega la necessaria imputabilità della decisione algoritmica al

⁷⁵⁰ T.A.R. Lazio, sent., Sez. III-bis, n. 3742 e 3769 del 2017. In commento alle sentenze v. E. PROSPERETTI, *Accesso al software e al relativo algoritmo nei procedimenti amministrativi e giudiziari. Un'analisi a partire da due pronunce del TAR Lazio*, in *Dir. e inf.*, 4-5, 2019, pp. 979 e ss.

⁷⁵¹ T.A.R. Lazio, Sez. III-bis, sentt. nn. 9924-9230 del 2018 su cui si v. *ex multis*, A. SIMONCINI, *Profili costituzionali della amministrazione algoritmica*, in *Riv. Trim. Dir. Pubbl.*, 4, 2019, p. 1162.

⁷⁵² Cons. Stato, Sez. VI, sent. n. 2270 del 2019.

⁷⁵³ *Ibidem*.

⁷⁵⁴ Cons. Stato, Sez. VI, sent. n. 8472 del 2019 e sent. n. 881 del 2020.

⁷⁵⁵ *Ibidem*.

⁷⁵⁶ Che al § 1 prevede che «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». I paragrafi successivi introducono invece eccezioni rilevanti alla regola che per parte della dottrina è la stessa regola della non esclusività a trasformarsi in eccezione. Si cfr. A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, cit., p. 80. Tale articolo, di recente, è stato oggetto per la prima volta di un'interpretazione della Corte di Giustizia dell'Unione europea (sentt. *Schufa Holding*, *Case C-634/21 e Joined Cases C-26/22 and C-64/22*). La Corte, in un caso in cui la decisione era stata assunta *sulla base* di valutazioni automatizzate ha ricondotto tale decisione al divieto dell'art. 22 GDPR. Secondo la Corte, infatti, «sussisterebbe un rischio di

funzionario pubblico⁷⁵⁷. Qui però, rispetto alle decisioni dei giudici americani, il bilanciamento tra diritti di privativa delle società produttrici degli algoritmi e il diritto di conoscere (comprendere) l'algoritmo è più netto: i primi devono soccombere quando gli algoritmi vengono utilizzati nell'ambito dell'amministrazione e, aggiungiamo noi, *a fortiori* nell'ambito del procedimento giurisdizionale.

Solo alla trasparenza così intesa e alla conseguente permanenza dell'imputabilità della decisione all'essere umano può coesistere un reale obbligo di motivazione che si rafforza, avendo l'amministrazione l'obbligo di rendere comprensibile la propria decisione⁷⁵⁸.

Come si vede, quindi, secondo la giurisprudenza amministrativa il tassello irrinunciabile per verificare la legittimità ovvero l'illegittimità di un provvedimento amministrativo assunto (anche) da un algoritmo è la *comprensibilità* dello stesso, secondo una declinazione *rafforzata* del principio di trasparenza che rende la decisione compatibile con l'imputazione diretta della decisione al funzionario pubblico nonché con le garanzie processuali del destinatario del provvedimento (in particolare il diritto di difesa il diritto a una tutela giurisdizionale).

elusione dell'articolo 22 del RGPD e, di conseguenza, una lacuna nella protezione giuridica qualora fosse accolta un'interpretazione restrittiva di tale disposizione, secondo la quale il calcolo del tasso di probabilità deve essere considerato soltanto un atto preparatorio e solo l'atto adottato dal terzo può, se del caso, essere qualificato come «decisione», ai sensi dell'articolo 22, paragrafo 1, di tale regolamento. Infatti, in tale ipotesi, il calcolo di un tasso di probabilità come quello di cui trattasi nel procedimento principale sfuggirebbe ai requisiti specifici previsti all'articolo 22, paragrafi da 2 a 4, del RGPD, sebbene tale procedura si basi su un trattamento automatizzato e produca effetti che incidono significativamente sull'interessato, in quanto l'azione del terzo, al quale tale tasso di probabilità è trasmesso, è condizionata in modo decisivo da quest'ultimo (§ 71). Su tali sentenze si cfr. F. PALMIOTTO, *'Scoring' for Data Protection Rights: The Court of Justice's First Judgment on Article 22 GDPR (Case C-634/21 and Joined Cases C-26/22 and C-64/22)*, in *EULAWLive*, 2024.

⁷⁵⁷ Anche se non rileva direttamente ai fini del presente lavoro, i giudici amministrativi enunciano altresì il principio di non discriminazione algoritmica, secondo cui, anche di fronte ad algoritmi perfettamente conoscibili e non costituenti l'esclusiva fonte della decisione finale, essi non devono produrre effetti discriminatori nei confronti dei destinatari, si cfr. ancora le sentenze da ultimo citate.

⁷⁵⁸ In una recente decisione del T.A.R. Campania si legge che «il fatto che il provvedimento venga emanato sulla scorta di una complessa operazione di calcolo produce l'opposto effetto di rafforzare, per certi versi, l'obbligo motivazionale in capo all'Amministrazione, la quale dovrà rendere la propria decisione finale non solo conoscibile, ma anche comprensibile», cfr. T.A.R. Campania, Sez. III, sent. n. 7003 del 14 novembre 2022.

Le esigenze di trasparenza segnalate nel presente lavoro sono state considerate anche a livello legislativo: ricordando gli atti di *soft law* che, sia negli Stati Uniti sia in Europa, hanno indicato la trasparenza algoritmica come un principio irrinunciabile, qui ci soffermiamo sugli strumenti vincolanti. Già il GDPR conteneva delle previsioni riconducibili all'esigenza di conoscere i dati contenuti in un sistema automatizzato, anche se ancora non è sopito il dibattito sull'esistenza o meno nel GDPR di un reale *diritto alla spiegazione*⁷⁵⁹. Il nuovo Regolamento sull'IA sembra invero fare un passo in avanti verso garanzie di trasparenza (per i sistemi ad alto rischio, che a noi interessano), anche se sembra emergere dal testo dell'articolato una trasparenza "informativa", molto diversa, quindi, da quella indicata dalla giurisprudenza amministrativa, e che produce una evidente asimmetria tra chi progetta e pone sul mercato il sistema di IA e il soggetto destinatario delle decisioni⁷⁶⁰.

La natura "informativa" e non tecnica della trasparenza emerge già dalla rubrica dell'art. 13 § 1 *AI Act* "Trasparenza e fornitura di informazioni ai deployer", il quale obbliga il fornitore (*provider*)⁷⁶¹ a sviluppare i sistemi di IA ad alto rischio «in modo tale da garantire che il loro funzionamento sia *sufficientemente trasparente*». Ciò per consentire al *deployer*⁷⁶² – definito nel Regolamento «una persona fisica o giuridica, un'autorità pubblica, un'agenzia o un altro organismo che utilizza un sistema di IA sotto la propria autorità, tranne nel caso in cui il sistema di IA sia utilizzato nel corso di un'attività personale non professionale» (e quindi, ai nostri

⁷⁵⁹ Sul tema si cfr., compiutamente, A. FORMISANO, *L'impatto dell'intelligenza artificiale in ambito giudiziario sui diritti fondamentali*, in *Federalismi*, n. 22/2024, pp. 112-149. Sull'esistenza del diritto alla spiegazione nel GDPR si cfr. A. SELBST, *Meaningful information and the right to explanation*, in *7(4) International Data Privacy Law*, 2017, pp. 233 ss. Sulla trasparenza nel GDPR si cfr. F. DONATI, *Intelligenza artificiale e giustizia*, in *Rivista AIC*, 1/2020, pp. 415-436, spec. pp. 425-428.

⁷⁶⁰ Si cfr. G. LO SAPIO, *L'Artificial Intelligence Act e la prova di resistenza per la legalità algoritmica*, cit., pp. 285 ss.

⁷⁶¹ Ai sensi dell'art. 3 del Regolamento, il fornitore è «una persona fisica o giuridica, un'autorità pubblica, un'agenzia o un altro organismo che sviluppa un sistema di IA o un modello di IA per finalità generali o che fa sviluppare un sistema di IA o un modello di IA per finalità generali e immette tale sistema o modello sul mercato o mette in servizio il sistema di IA con il proprio nome o marchio, a titolo oneroso o gratuito» (n. 3).

⁷⁶² Il termine è rimasto in lingua inglese anche nella versione italiana del regolamento. Nella proposta della commissione, invece, la parola inglese usata era *user*, tradotta in italiano con *utente*.

fini, potrebbe essere sia l'ordine giudiziario sia il singolo giudice)⁷⁶³, di «interpretare l'output del sistema e utilizzarlo adeguatamente».

L'estrema vaghezza della previsione (non è chiaro il significato di *sufficientemente*)⁷⁶⁴ viene almeno in parte concretizzata nei paragrafi seguenti, dove si precisa che i sistemi di IA *high risk* siano accompagnati da istruzioni per l'uso, predisposte dai fornitori, «che comprendono informazioni concise, complete, corrette e chiare che siano pertinenti, accessibili e comprensibili per i *deployer*» (art. 13 § 2), tra cui l'identità e i dati del fornitore, le caratteristiche e le capacità del sistema di IA (accuratezza, dati di addestramento ecc...) nonché le misure tecniche impiegate per facilitare l'interpretabilità degli *output* da parte del *deployer* (art. 13 § 3).

L'incerta estensione di tali obblighi non è chiarita nemmeno nei *considerando* del Regolamento, che come noto, anche se non vincolanti, contribuiscono a gettare luce interpretativa sulle norme: per il *considerando n. 27*, infatti, per «trasparenza» si intende che «i sistemi di IA sono sviluppati e utilizzati in modo da consentire un'adeguata tracciabilità e spiegabilità, rendendo gli esseri umani consapevoli del fatto di comunicare o interagire con un sistema di IA e informando debitamente i *deployer* delle capacità e dei limiti di tale sistema di IA e le persone interessate dei loro diritti». Anche qui, l'*adeguatezza* della spiegabilità è un concetto tutt'altro che definito e quindi foriero di diverse possibili interpretazioni.

In sostanza, quindi, è il progettista (*provider*) a poter determinare ciò che deve essere reso trasparente: ciò sembra venire incontro a due diverse ragioni.

La prima è che l'*AI Act*, a differenza della giurisprudenza amministrativa interna, non pare voler affrontare le logiche proprietarie di coloro che progettano gli algoritmi. E anzi, gli obblighi di trasparenza variamente imposti dal regolamento fanno salva la necessità di rispettare e proteggere i diritti di proprietà intellettuale, le informazioni commerciali riservate e i segreti commerciali conformemente al

⁷⁶³ Bisognerà capire come verrà interpretata la parola *deployer*. Il regolamento parla sia di persona fisica sia giuridica, sia di autorità pubblica. Una definizione tanto ampia da poter ricomprendere sia il singolo giudice sia l'ordine giudiziario nel suo complesso.

⁷⁶⁴ La vaghezza messa in risalto anche da C. CASONATO – B. MARCHETTI, *Prime osservazioni*, cit., p. 427.

diritto dell'Unione e nazionale⁷⁶⁵. Il bilanciamento tra questi diritti viene dunque lasciato al diritto nazionale e, come abbiamo visto, secondo la giurisprudenza amministrativa prevalgono le garanzie del giusto processo ai diritti delle imprese produttrici di algoritmi.

La seconda sembra essere legata a una *ragion pratica* che l'*AI Act* non intende risolvere, vale a dire il noto problema della *technical black box*: essendo i sistemi di IA spesso incomprensibili anche a chi li progetta, viene ovviato pragmaticamente il problema attraverso obblighi *comunicativi* di funzionamento degli stessi sotto forma di istruzioni d'uso, documentazione tecnica ecc...⁷⁶⁶. Tale flusso informativo lungo la catena di produzione e utilizzo di IA ha un punto di caduta nell'art. 86 del Regolamento "Diritto alla spiegazione dei singoli processi decisionali"⁷⁶⁷ il quale, probabilmente intercettando le critiche che erano state rivolte al GDPR dopo la sua approvazione, prevede che chiunque sia oggetto di una decisione adottata dal *deployer* sulla base dell'*output* di un sistema di IA ad alto rischio che produca effetti giuridici o incida su tale persona in un modo in cui essa ritenga avere un impatto negativo sulla sua salute, sulla sua sicurezza o sui suoi diritti fondamentali «ha il diritto di ottenere dal *deployer* (e quindi, per quanto qui interessa, dal giudice) spiegazioni chiare e significative sul ruolo del sistema di IA nella procedura decisionale e sui principali elementi della decisione adottata»⁷⁶⁸.

È evidente che se caliamo tale previsione nel contesto della giustizia algoritmica poco viene aggiunto rispetto all'onere motivazionale da parte del giudice per ogni provvedimento giurisdizionale, riconosciuto sia a livello nazionale sia europeo (si v. *supra* capitolo VII § 7). Soprattutto, non risolve i problemi che abbiamo segnalato rispetto alla trasparenza necessaria (anche per il giudice) a permettere a chi impiega lo strumento algoritmico di comprenderne il funzionamento e darne

⁷⁶⁵ Si v. *considerando n. 88 e n. 167*, art. 25 § 5, art. 53 § 1 lett. b) e art. 78 del Regolamento.

⁷⁶⁶ G. LO SAPIO, *L'Artificial Intelligence Act e la prova di resistenza per la legalità algoritmica*, cit., p. 287.

⁷⁶⁷ Si segnala che la norma non era prevista nella proposta iniziale delle Commissioni, ma è stata aggiunta a seguito del trilogico con il Consiglio e il Parlamento europei.

⁷⁶⁸ Il § 2, tuttavia, probabilmente con in mente i diritti di privacy, specifica che il diritto alla spiegazione comunque «non si applica all'uso di sistemi di IA per i quali sono previste eccezioni o limitazioni all'obbligo stabilito in tale paragrafo in virtù del diritto dell'Unione o del diritto nazionale, in linea con il diritto dell'Unione».

conto nella motivazione. Così, se gli obblighi di trasparenza *a monte*, come sembra, non garantiscono piena spiegabilità del sistema algoritmico, parrebbero frustrate sia le segnalate esigenze di tutela delle parti nel processo, sia la capacità del giudice di adempiere a pieno la sua funzione giurisdizionale.

Ancora non è possibile sapere la portata applicativa degli obblighi previsti dal Regolamento, il livello di «chiarezza e significatività» richiesto al fornitore o il livello richiesto per far sì che il sistema sia «sufficientemente trasparente». Non possiamo tuttavia esimerci dal considerare tali disposizioni, nell'ambito della giustizia algoritmica, insufficienti. Gli obblighi imposti sono previsti per tutti i sistemi considerati ad alto rischio e, se possono risultare adeguati in determinati contesti, in sede processuale la comprensibilità della decisione assunta dal giudice e degli elementi da lui considerati acquistano un valore maggiore, oltre che costituzionalmente rilevante. Si pensi per esempio a un *deployer* diverso dal giudice, quale un operatore sanitario che impiega un sistema di IA per effettuare diagnosi, con percentuali di accuratezza altissime ma di difficile comprensione tecnica: probabilmente gli obblighi informativi previsti nel Regolamento risulteranno sufficienti per ritenere legittima tale diagnosi e, soprattutto, il paziente sottoposto alla diagnosi artificiale non avrà bisogno di spiegazioni e comprensioni particolarmente pregnanti (del resto, quando richiediamo delle cure, ci *fidiamo* del medico curante e non abbiamo bisogno di comprendere le ragioni tecniche e mediche che portano alla diagnosi e alla successiva cura, ci basta una trasparenza «comunicativa»).

Diverso è il caso della giustizia algoritmica, dove le esigenze di trasparenza sono declinate, come visto nel capitolo precedente, nella conoscibilità degli elementi rilevanti per la decisione per poter esercitare i diritti di difesa e a una tutela giurisdizionale *effettiva* (come la possibilità di impugnare la decisione).

La trasparenza algoritmica è dunque essa stessa condizione di legittimità delle decisioni giudiziarie e, se oggi, nell'incertezza che permane sui possibili utilizzi dell'IA nel processo, essa può fungere da bussola per la garanzia della tutela di diritti fondamentali ad essa connessi, in un futuro ancora incerto in cui l'intelligenza artificiale avrà un utilizzo sempre più frequente e incisivo nella giurisdizione, essa

dovrà rappresentare la stella polare per verificare la rispondenza di decisioni algoritmiche con il dettato costituzionale interno ed europeo tanto per i decisori politici quanto per i giudici.

In questo senso, per chiudere il discorso, nel giusto processo europeo integrato si può individuare un fondamento costituzionale della trasparenza algoritmica, individuato da alcuni autori, con diverse declinazioni, nell'art. 24 della Costituzione⁷⁶⁹. Ampliando il discorso, potremmo dire che la trasparenza algoritmica trovi il fondamento nel diritto a una tutela giurisdizionale, la cui effettività può garantirsi solo se la giurisdizione è esercitata secondo le regole del giusto processo. Ecco che quindi la trasparenza diventa una componente essenziale della garanzia costituzionale alla tutela giurisdizionale effettiva, che trova fondamento in tutte le disposizioni costituzionali e sovranazionali che impongono il rispetto del *fair trial* (gli artt. 24, 101, 104 e 111 Cost, l'art. 6 CEDU e l'art. 47 CDFUE). E se, come affermato dalla Corte costituzionale, la tutela giurisdizionale costituisce un principio supremo del nostro ordinamento, riconoscere la trasparenza algoritmica come corollario di tale principio la porrebbe, secondo la teoria dei controlimiti, dietro a uno scudo protettivo anche di fronte a eventuali norme europee in contrasto con esso.

Per concludere, non è forse utile oggi declinare tale trasparenza in un termine preciso o in una scelta pratica definita (accesso al codice, spiegabilità dei risultati, comprensione dell'*output* ecc...: indipendentemente dalla complessità degli algoritmi che la giurisdizione utilizza o utilizzerà in procedure come quella qui esaminata, ciò che appare ineludibile per superare il vaglio di legittimità di siffatte

⁷⁶⁹ Si cfr. A. SIMONCINI, *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*, cit. in cui l'A. ricostruisce la portata estensiva dell'art. 24 secondo un principio di «lealtà digitale», secondo cui «una applicazione effettiva ed attuale del principio costituzionale di difesa dei diritti richiede che ogniqualvolta un soggetto – sia esso pubblico o privato, persona fisica o giuridica – impieghi un dispositivo di Intelligenza Artificiale per prendere decisioni suscettibili di interferire con un diritto o interesse, chiunque subisce tale decisione ha il diritto: a) di sapere se nella decisione è stato utilizzato uno strumento tecnologico (ovvero, il diritto ad essere informati se e quando stiamo interagendo con una macchina ovvero con un essere umano); b) di comprendere la decisione (ovvero, il diritto di conoscere le ragioni che sostengono tale decisione). Si cfr. anche A. FORMISANO, *L'impatto dell'intelligenza artificiale in ambito giudiziario sui diritti fondamentali*, cit., che invece declina il fondamento della trasparenza nell'art. 24 Cost nel principio di spiegabilità e *traducibilità*, in un linguaggio comprensibile, della decisione algoritmica

procedure non sarà né la disponibilità del codice sorgente e nemmeno la mera spiegazione della logica utilizzata dell'algoritmo. Ciò che invece risulta imprescindibile è che il funzionamento di tali algoritmi consenta *effettivamente* alle parti di interpretare e utilizzare l'*output* per la tutela dei propri diritti (*in primis* la tutela giurisdizionale)⁷⁷⁰. Non si richiede dunque al potere giudiziario che intenda valersi di strumenti algoritmici un'obbligazione di mezzi, quanto piuttosto un'obbligazione di risultato, a prescindere dalle modalità e dalle scelte compiute su come tale risultato (la trasparenza) debba essere conseguito.

4. Alfabetizzazione e autonomia del giudice

Come emerso nel corso del lavoro, l'opacità algoritmica non rende soltanto illusori i diritti delle parti alla conoscibilità delle prove e al diritto di difesa, ma tende a svuotare anche il ruolo del giudice come custode del processo e dei diritti delle parti. Come abbiamo visto, infatti, non solo a queste ultime verrebbe impedito di comprendere la decisione del giudice, ma anche questa sarebbe fondata su strumenti non del tutto comprensibili dallo stesso giudice.

In un passo della sentenza *Loomis* si afferma che «*[t]he circuit court and Loomis had access to the same copy of the risk assessment*», considerazione che mostra chiaramente che l'algoritmo (in quel caso era un *risk assessment tool*, ma potrebbe essere in futuro un altro algoritmo decisionale, o un sistema di giustizia predittiva o uno strumento di IA generativa) che ha contribuito a determinare diritti e obblighi della parte è lo stesso algoritmo che il giudice ha impiegato per assumere la propria decisione.

È ancora muovendo alla giurisprudenza statunitense che diventa facile, quasi automatico, fornire alcune soluzioni.

⁷⁷⁰ Come è stato scritto: «*Transparency is not enough, in itself: transparency must be meaningful*», si v. S. QUATTROCOLO - C. ANGLANO - M. CANONICO - M. GUAZZONE, *Technical Solutions for Legal Challenges: Equality of Arms in Criminal Proceedings*, in *Global Jurist*, 20, 1, 2020.

Le decisioni dei giudici americani esaminate nel presente lavoro spaziano da alcune posizioni di assoluta chiusura ad altre assai permissive rispetto all'impiego *dei risk assessment tools* nel *sentencing*.

La *ratio decidendi* delle sentenze, come visto, ruota intorno al riconoscimento o meno di un abuso di discrezionalità del giudice nel momento in cui deve determinare l'ammontare della pena anche sulla base dei risultati prodotti dal *tool* algoritmico. La riflessione ha mosso dalle argomentazioni assai radicali esposte in Indiana nei casi *Cardwell* e *Rhodes*, decisioni in cui si ostacola con forza qualunque utilizzo di algoritmi nel *sentencing*. In queste sentenze si scrive chiaramente che il ricorso a formule algoritmiche priva il giudice della possibilità di giudicare in autonomia tutti i fattori rilevanti per determinare la pena, inaridisce il suo giudizio discrezionale e lo deresponsabilizza.

In una prospettiva quasi del tutto opposta si posizionano invece le decisioni assunte nei casi *Malenchik* e *J.S.*, nelle quali si argomenta che proprio l'impiego di algoritmi di previsione del rischio di recidiva, che rappresentano uno dei tanti elementi da considerare nel procedimento di *sentencing*, permetterebbe di rinvigorire e potenziare (*enhance*) la valutazione autonoma del giudice attraverso la considerazione di un elemento aggiuntivo in grado di indirizzare la decisione verso il trattamento più adeguato per l'imputato (si cfr. *supra* capitolo VIII § 2).

Le motivazioni spese da ultimo dai giudici dell'Indiana diventano presto il perno su cui ruoterà la giurisprudenza successiva di altri stati e che inizierà a introdurre nel processo anche le considerazioni riguardanti il rispetto del *due process of law*.

Sintetizzando, l'abuso di discrezionalità e le violazioni del giusto processo non si configurano in quanto – ancorché esistano dubbi sull'affidabilità degli strumenti utilizzati, sulla loro provenienza privata e sulla loro conseguente inaccessibilità – essi non catturano né assorbono la decisione ma costituiscono solo uno dei tanti elementi che il giudice può prendere in considerazione per assumere la decisione sulla pena. In questo senso non vi è violazione delle regole processuali né quando il giudice trascura l'algoritmo (si cfr. *Samsa*) né quando la pena è determinata assegnando un peso rilevante al risultato algoritmico (si cfr. *Loomis* e *People v. Younglove*, *supra* capitolo VIII § 3).

Sulla base di siffatte decisioni alcuni giudici hanno poi segnalato come la distinzione tra *reliance* e *consideration* appare in realtà ingannevole e nasconde la propensione mostrata dai giudici ad assegnare un peso decisivo all'algoritmo predittivo (si cfr. *supra*, capitolo VIII § 3, spec. *Gordon* e la *concurring opinion* della giudice Roggensack in *Loomis*).

La tendenza ad affidare la decisione al risultato automatizzato comporta quindi che una parte della decisione non rimane affidata al giudice ma viene esternalizzata a soggetti privati che progettano e immettono nel mercato gli algoritmi (si cfr. ancora, *People v. Younglove*).

Tale fenomeno si aggrava ulteriormente se si considerano le caratteristiche degli algoritmi più evoluti, il cui funzionamento, come segnalato in altre sentenze (si cfr. *Gordon*, *Guise* e la *concurring opinion* della Giudice Cavanagh in *People v. Canedo*), molto spesso sfugge a chi è chiamato a utilizzarli.

Tutto ciò ci ha concesso di sollevare dubbi sul rispetto di alcune garanzie istituzionali e procedurali che regolano la giurisdizione, dall'indipendenza e imparzialità del giudice (capitolo VIII § 6), alla *fairness* del procedimento giurisdizionale (capitolo VIII § 7) fino al diritto a una tutela giurisdizionale *effettiva*.

Ora, ciò che in effetti sembra accomunare tutte le decisioni statunitensi è l'assenza di una cornice che possa racchiudere il quadro complesso in cui si è sviluppata la giurisprudenza americana. Tale cornice, a nostro avviso, deve essere ricercata in una progressiva e imprescindibile formazione del giudice sugli algoritmi e sull'intelligenza artificiale e la conseguente capacità di comprendere come l'algoritmo conduce propri calcoli, come da determinati dati viene prodotto l'*output*, come funziona il *machine learning*. Ciò in quanto non è possibile convincersi della utilità di un continuo sviluppo della giustizia algoritmica se coloro i quali sono chiamati a farne applicazione non ne hanno il pieno controllo.

Sul punto, tanto la giurisprudenza quanto la dottrina statunitensi sembrano concordi nel ritenere fondamentale una nuova formazione tecnologica del giudice.

In *Loomis*, l'imputato nel procedimento dinanzi alla *trial court* aveva chiamato a testimoniare un esperto criminologo, il quale aveva opportunamente riferito che

le «[s]entencing courts have very little information about how a COMPAS assessment analyzes the risk», specificando che «[t]he Court does not know how the COMPAS compares that individual's history with the population that it's comparing them with. The Court doesn't even know whether that population is a Wisconsin population, a New York population, a California population. . . . There's all kinds of information that the court doesn't have, and what we're doing is we're mis-informing the court when we put these graphs in front of them and let them use it for sentence».

A differenza della Corte distrettuale, che non accoglie gli argomenti del perito, la Corte suprema pare invece condividere le perplessità sollevate rispetto al *deficit* di preparazione del giudice e, citando nella motivazione le parole tratte da uno studio specialistico in cui si sosteneva che «*unless criminal justice system actors are made fully aware of the limits of the tools they are being asked to implement, they are likely to misuse them*», la Corte afferma di prestare attenzione al suddetto monito, imponendo alcune cautele per l'utilizzo dei *risk assessment tools* (come visto *supra* capitolo VIII § 3).

aAncora, in una relazione prodotta in giudizio nel caso *Samsa* l'imputato notava opportunamente che: «*[w]ithout training on the COMPAS assessment, it would be easy to misinterpret the results, which are often displayed in bar chart form*».

Proprio questa mancanza di comprensione del fenomeno algoritmico si lega a doppio filo con quello che abbiamo definito *automation bias* (si v. *supra* capitolo VI § 3) e che tende in fondo a catturare la stessa decisione del giudice. In questo senso la dottrina statunitense ha fatto notare che «*individuals who receive such training are more likely to scrutinize an automated system's suggestions*» e che gli studi hanno dimostrato che i corsi di *training* ai giudici sulla valutazione dei pareri di esperti tecnici sia stata molto efficace per evitare l'*automation bias*⁷⁷¹, l'intensità del quale è sì «direttamente proporzionale alla diffusione dell'innovazione tecnologica»⁷⁷² ma anche alla scarsa consapevolezza degli strumenti utilizzati.

⁷⁷¹ D.K. CITRON, *Technological Due Process*, cit., p. 1306. S. BREWER, *Scientific Expert Testimony and Intellectual Due Process*, in *Yale Law Journal*, 107(6), 1998, pp. 1531-1681.

⁷⁷² G. LO SAPIO, *L'Artificial Intelligence Act e la prova di resistenza per la legalità algoritmica*, in *Federalismi*, cit., p. 281.

Specificamente rispetto all'impiego di COMPAS, si è affermato che «*in order to use COMPAS during sentencing, judges would have to attend trainings that would assist them in understanding the benefits as well as the limitations of COMPAS*»⁷⁷³. Ancora più nettamente, e con uno sguardo generale sul giudiziario statunitense, si è affermato che «*we live in an age of breakneck technological change that will thrust many difficult technical and scientific issues on judges, for which very few of them (of us, I should say) are prepared, because of the excessively rhetorical emphasis of legal education and the weak scientific background of most law students*»⁷⁷⁴.

Dunque, senza un'adeguata preparazione e formazione sul funzionamento degli algoritmi e dell'intelligenza artificiale non sembra possibile raggiungere uno *standard* minimo per poter garantire un esercizio corretto della funzione giurisdizionale e, in ultimo, garantire il rispetto del giusto processo.

Del resto, la preparazione e la volontà dei giudici di comprendere il fenomeno algoritmico potrebbe comportare una maggiore ricerca di trasparenza da parte di coloro che progettano gli algoritmi: come è stato scritto «*Going forward, technologists need to be far more transparent about the nature of the choices they make when designing new technology. They must create detailed documentation of the design choices made and the rationales for them, courts need to be prepared to ask more specific questions about technology design choices and require explanations of how these choices might have led to the outcome being adjudicated, and what alternative design strategies might have been suitable. This will have feedback effects as well. If courts demonstrate a willingness to deconstruct design, designers will have a greater incentive to design transparently to reduce litigation costs and will ultimately design more thoughtfully to incorporate not just efficiency and scaling considerations, but potential harms that they might have to litigate later*»⁷⁷⁵.

⁷⁷³ K. FREEMAN, *Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis*, cit., p. 104.

⁷⁷⁴ R. A. POSNER, *The Role of The Judge on the Twenty-First Century*, in 86 *Boston University Law Review*, pp. 1049- 1068, spec. 1049.

⁷⁷⁵ A. SELBST - S. VENKATASUBRAMANIAN - I. E. KUMAR, *Deconstructing design decisions: why courts must interrogate machine learning and other technologies*, cit., p. 60.

La necessità di consapevolezza e conoscenza dei nuovi paradigmi portati da intelligenza artificiale e decisioni automatizzate è intercettata anche dall'*AI Act*, dove anzi le esigenze di comprensione del fenomeno dell'IA sembrano assumere un peso centrale.

Il Regolamento, infatti, sembra considerare decisiva la “alfabetizzazione in materia di IA” (*AI literacy* nel testo inglese), definita come «le competenze, le conoscenze e la comprensione che consentono ai fornitori, ai deployer e alle persone interessate, tenendo conto dei loro rispettivi diritti e obblighi nel contesto del presente regolamento, di procedere a una diffusione informata dei sistemi di IA, nonché di acquisire consapevolezza in merito alle opportunità e ai rischi dell'IA e ai possibili danni che essa può causare» (art. 3, n. 56). La centralità dell'alfabetizzazione per il legislatore europeo⁷⁷⁶ è manifestata anche dal collocamento dell'articolo dedicato, inserito appena dopo aver definito oggetto, ambito di applicazione e definizioni (artt. 1-3). L'art. 4 dell'*AI Act* prevede infatti che «[i] fornitori e i deployer dei sistemi di IA adottano misure per garantire nella misura del possibile un livello sufficiente di alfabetizzazione in materia di IA del loro personale nonché di qualsiasi altra persona che si occupa del funzionamento e dell'utilizzo dei sistemi di IA per loro conto, prendendo in considerazione le loro conoscenze tecniche, la loro esperienza, istruzione e formazione, nonché il contesto in cui i sistemi di IA devono essere utilizzati, e tenendo conto delle persone o dei gruppi di persone su cui i sistemi di IA devono essere utilizzati». Compito della Commissione, tramite l'Ufficio per l'IA, e degli Stati Membri, anche tramite il Consiglio per l'IA (si v. il capitolo V § 3 per la struttura di governance prevista nel Regolamento), è quella di «promozione dell'alfabetizzazione in materia di IA, della sensibilizzazione del pubblico e della comprensione dei benefici, dei rischi, delle garanzie e dei diritti e degli obblighi in relazione all'uso dei sistemi di IA» (art. 66 § 1 lett. f); v. anche art. 95 § 2 lett. c)).

⁷⁷⁶ Obiettivo simile, come visto (capitolo V § 6) è stato perseguito dal dalla Convenzione sull'IA che all'art. 20 obbliga gli Stati a «*encourage and promote adequate digital literacy and digital skills for all segments of the population, including specific expert skills for those responsible for the identification, assessment, prevention and mitigation of risks posed by artificial intelligence systems*».

Nel *considerando n. 20* vengono infine delineati metodi e obiettivi dell'alfabetizzazione per consentire il raggiungimento di «massimi benefici dai sistemi di IA proteggendo nel contempo i diritti fondamentali, la salute e la sicurezza e di consentire il controllo democratico»: la comprensione della corretta applicazione degli elementi tecnici durante la fase di sviluppo del sistema di IA, le modalità adeguate per interpretare l'*output* del sistema di IA, le conoscenze necessarie per comprendere in che modo le decisioni adottate con l'assistenza dell'IA incideranno sulle persone. Tutto ciò per «sostenere il consolidamento e il percorso di innovazione di un'IA affidabile nell'Unione».

È dunque una catena che dovrebbe partire dalla struttura di governance prevista nel regolamento, passare dai fornitori e dai *deployer* e giungere sino alla sensibilizzazione del pubblico⁷⁷⁷.

Sono previsioni, come si intuisce, molto generiche e difficilmente in grado di creare veri e propri obblighi ai destinatari – la valutazione di tali norme andrà valutata nel tempo rispetto alle misure di attuazione del Regolamento e l'accettazione più o meno propositiva e condivisa degli attori in campo – ma non v'è dubbio che segnano il percorso verso una direzione che sembra imprescindibile per poter consentire un ingresso graduale degli algoritmi nel processo.

Le previsioni sull'alfabetizzazione, inoltre, sembrano prodromiche a rendere effettivamente realizzabili gli obblighi previsti per i *deployer* per quanto concerne la sorveglianza umana (*human oversight*) che «mira a prevenire o ridurre al minimo i rischi per la salute, la sicurezza o i diritti fondamentali che possono emergere quando un sistema di IA ad alto rischio è utilizzato conformemente alla sua finalità prevista o in condizioni di uso improprio ragionevolmente prevedibile, in particolare qualora tali rischi persistano nonostante l'applicazione di altri requisiti» (art. 14 § 2 del Regolamento). La progettazione dovrebbe consentire al *deployer*, e quindi sia al potere giudiziario nel suo complesso sia, di conseguenza, al singolo giudice, di comprendere le capacità e i limiti del sistema di IA ed essere in grado di

⁷⁷⁷ Gli ultimi dati mostrano che l'Italia ha un livello ampiamente sotto la media europea di *Digital literacy*. Si cfr. i dati pubblicati da Eurostat al seguente indirizzo: <https://data.europa.eu/en/publications/datastories/digital-literacy-eu-overview>.

monitorarne il funzionamento; di «restare consapevole della possibile tendenza a fare automaticamente affidamento o a fare eccessivo affidamento sull'output prodotto da un sistema di IA ad alto rischio («distorsione dell'automazione»), in particolare in relazione ai sistemi di IA ad alto rischio utilizzati per fornire informazioni o raccomandazioni per le decisioni che devono essere prese da persone fisiche» (abbiamo già spiegato che la *distorsione dell'automazione* è la formulazione italiana che il Regolamento impiega per l'*automation bias*, su cui si v. *supra*, capitolo VI § 3); di interpretare correttamente l'*output* del sistema di IA ad alto rischio, tenendo conto ad esempio degli strumenti e dei metodi di interpretazione disponibili; di ignorare, annullare o ribaltare l'*output* del sistema di IA ad alto rischio (art. 14 § 4).

Tali obblighi, uniti a quelli previsti dall'art. 26 e 27 del Regolamento⁷⁷⁸ per i *deployer* di sistemi ad alto rischio come quelli utilizzati dall'autorità giudiziaria, paiono particolarmente rigorosi e assai penetranti; viene da chiedersi, quindi, in che modo il sistema giurisdizionale, in Italia ma anche in tanti paesi europei, per come oggi lo conosciamo – quanto a alfabetizzazione digitale, capacità di adattamento, sviluppo tecnologico – possa essere in grado di adempiere a tutti i compiti previsti nell'*AI Act*. Gli obiettivi di alfabetizzazione potranno senz'altro aiutare il processo, ma forse per pensare a una giustizia algoritmica che si fa spazio nelle aule giudiziarie in conformità con le nuove regole europee sembra necessaria una notevole riorganizzazione dell'amministrazione della giustizia oltre a una rivoluzione culturale in materia di intelligenza artificiale che deve includere ogni attore in campo, dalla società civile alle istituzioni, dai progettisti ai giudici. Solo in questo modo, per usare le parole del Regolamento, il processo decisionale finale rimarrà «un'attività a guida umana»⁷⁷⁹, l'unica oggi coerente con le garanzie costituzionali.

⁷⁷⁸ Sui quali si cfr. M. GAUR, *An In-Depth Exploration of Predictive Justice with AI*, in *AI and Emerging Technologies, Automated Decision-Making, and Ethical Considerations*, 2024.

⁷⁷⁹ *Considerando n. 61* del Regolamento. Si v., in tema, A. Z. HUQ, *A Right to a Human Decision*, cit., nonché E. LONGO, *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit., pp. 335-337 e S. GABORIAU, *Libertà e umanità del giudice: due valori fondamentali della giustizia. La giustizia digitale può garantire nel tempo questi valori?*, in *Questione giustizia*, trimestrale n. 4, 2018. A questo riguardo, l'allora Ministra Marta Cartabia, in

Un tentativo (parziale ma ambizioso) di conclusione: la giustizia algoritmica come opportunità

Giunti sino a qui – dopo un percorso che dall’analisi dello stato dell’arte della giustizia algoritmica, passando per la regolazione di tale fenomeno, ci ha condotto ad analizzare in una chiave costituzionalistica i *vulnera* di un inevitabile cambiamento (in alcuni casi ancora *in potenza*, in altri già *in atto*), e a delineare soluzioni spesso non risolutive – la risposta più semplice, più immediata, sembra essere una sola: negare *sic et simpliciter* l’impiego di algoritmi nel processo, mantenere una giustizia esclusivamente umana ed escludere l’intelligenza artificiale da un’attività così delicata come quella del giudice.

Se non ci si vuole accontentare di una siffatta soluzione (che sarebbe probabilmente antistorica, vista l’irrefrenabile evoluzione tecnologica), scavando sul fondo delle parole dei giudici americani e di quelli europei, della letteratura specialistica e dei primi dati normativi – consci ormai, arrivati sino a qui, delle complessità che attengono all’impatto degli algoritmi sulla giurisdizione – si intravedono alcuni spazi di riflessione, seppure ancora parziale e incompleta.

Le riserve e le paure generalizzate su un impiego strutturale di algoritmi nel processo è, a nostro avviso, il riflesso di uno sguardo prospettico catastrofista che spesso sembra circondare la tecnologia in generale. Tale sguardo osserva con grande attenzione i rischi dell’uso delle tecnologie mentre guarda solo di traverso i vantaggi e le opportunità che un’evoluzione tecnologica controllata potrebbe portare alla società. Lo stesso sguardo peraltro ha influenzato anche il presente lavoro, probabilmente più concentrato sui problemi che sulle soluzioni della giustizia algoritmica. E si badi: tale angolo visivo è del tutto comprensibile. I

occasione del *Digital Justice Ministerial Forum* del 12 ottobre 2021 affermò la necessità del pieno rispetto dell’indipendenza della magistratura e, inoltre, specificò che «nell’attività giudiziaria si possono distinguere le fasi del conoscere e quelle del decidere e nella prima, che richiede lo studio dei fatti e delle norme e della giurisprudenza applicabili alla controversia, l’intelligenza artificiale può portare un grande valore aggiunto per la sua immensa capacità di raccogliere ed elaborare dati, con una potenza che sfugge alle umane capacità. Viceversa, il momento del decidere deve rimanere sempre nelle mani del giudice in carne ed ossa, capace di cogliere tutte le sfumature, le peculiarità e le irripetibili specificità di ogni singolo caso».

problemi che stanno nascendo con l'impiego dell'intelligenza artificiale sono innumerevoli e le esperienze fino ad oggi di giustizia algoritmica non di rado sono parse impermeabili a tentativi di canalizzare il fenomeno su binari pienamente legittimi.

Proprio per questa ragione, tuttavia, in queste ultime righe vi è il tentativo di cambiare l'angolo prospettico e muovere da una visione di giustizia algoritmica come opportunità.

Le soluzioni che abbiamo analizzato in questo ultimo capitolo sono ancora da verificare nella pratica, anche perché ancora non si conosce l'interpretazione e l'applicazione che verrà data del Regolamento sull'intelligenza artificiale. Ciò nonostante, sembra potersi dire che se in futuro si concepirà la giustizia algoritmica come decisiva per rafforzare le garanzie costituzionali, specialmente quelle riferite al giusto processo, sarà più facile concepire soluzioni e guarentigie che siano finalizzate a tale ambizioso obiettivo.

Controllo pubblico di dati e algoritmi, trasparenza del procedimento algoritmico e preparazione del giudice sembrano pilastri a cui non sembra potersi rinunciare, ma la loro declinazione pratica potrà assumere le forme più disparate, purché coerenti con l'obiettivo appena indicato.

In questo senso, la preventiva e profonda consapevolezza del giudice sul funzionamento degli algoritmi getterebbe, infatti, nuova luce sulle sentenze dei giudici americani qui prese in esame e, di conseguenza, sulle garanzie del giusto processo: in questo senso le decisioni di assoluta chiusura all'impiego di algoritmi, apparirebbero poco condivisibili, in quanto l'impiego di strumenti algoritmici in grado di produrre ed elaborare dei dati difficilmente acquisibili dall'essere umano, purché mantenuti nel circuito pubblico e controllati da un giudice in grado di comprenderne le peculiarità e le eventuali storture, fornirebbero allo stesso decidente un elemento di prova utile per una decisione più accurata e razionale.

Di contro, le sentenze che qualificano gli algoritmi come strumenti in grado di migliorare la decisione discrezionale del giudice acquisirebbero un significato diverso. Il giudice con la giusta preparazione, infatti, di fronte a un algoritmo trasparente e pubblico, ben potrebbe motivare sulla rilevanza nella decisione

dell'algoritmo stesso e sulle ragioni per le quali in un caso ha deciso di discostarsi dal risultato prodotto, valorizzando altri elementi della causa, in un altro decidendo di tenerlo, invece, in maggiore considerazione.

Il pieno controllo degli strumenti algoritmici, oltre alle altre cautele afferenti alla trasparenza degli stessi e alla loro progettazione pubblica, rallenterebbe il fenomeno della esternalizzazione e riporterebbe il giudizio discrezionale nelle mani del giudice, il quale si assumerebbe la piena responsabilità della decisione finale ma con l'ausilio di strumenti artificiali che ampliano la sua base di conoscenza.

In ultimo, un uso accorto e consapevole di strumenti tecnologici e algoritmici da parte dei giudici contribuirebbe a creare un clima di maggior fiducia nei confronti della magistratura, requisito irrinunciabile stante il ruolo fondamentale che svolge in uno Stato di diritto⁷⁸⁰.

Ora, se le soluzioni rispetto al controllo pubblico sugli algoritmi e all'opacità algoritmica paiono poter essere raggiunte con adeguate cautele sia tecniche sia normative, più complicato sembra produrre con gli strumenti del diritto un cambiamento del ruolo che deve avere il giudice di fronte all'avvento della giustizia algoritmica. La rivoluzione della giustizia algoritmica, in altre parole, se vuole insinuarsi tra le pieghe del processo, non passa solo da soluzioni tecniche o normative, ma da una necessaria rivoluzione culturale dei giudici, non dissimile da quella che hanno dovuto compiere con l'avvento del costituzionalismo⁷⁸¹ e del processo di integrazione europea⁷⁸². Ciò su cui si dovrà, nei prossimi tempi,

⁷⁸⁰ Corte Edu, *De Haes and Gijssels c. Belgio*, 1997, § 37. Sul tema si cfr., ancora, B. RANDAZZO, *Imparzialità del giudice e fiducia nella magistratura nella prospettiva costituzionale europea*, cit., p. 233.

⁷⁸¹ Si cfr., sul nuovo ruolo del giudice con l'avvento del costituzionalismo, tra i tanti, P. GROSSI, *Ritorno al diritto*, cit.

⁷⁸² Sono note le difficoltà che ha dovuto affrontare il giudice nell'interpretazione e applicazione della legge quando ha dovuto iniziare a fare i conti sia con il sistema del Consiglio d'Europa (e quindi con la giurisprudenza della Corte EDU), sia con le fonti dell'Unione europea (e la giurisprudenza della Corte di Giustizia). Su tali tematiche, soprattutto con riguardo alle problematiche afferenti il rapporto tra giudici domestici e CEDU, si cfr., B. RANDAZZO, *Interpretazione delle sentenze della Corte europea dei diritti ai fini dell'esecuzione (giudiziaria) e interpretazione della sua giurisprudenza ai fini dell'applicazione della CEDU*, in *Rivista AIC*, 2/2015. Già prima si cfr. B. RANDAZZO, *Giudici comuni e Corte europea dei diritti*, in *Atti del Seminario del "Gruppo di Pisa", La Corte costituzionale e le Corti d'Europa*, (Catanzaro, 31 maggio – 1° giugno 2002), Giappichelli, Torino 2003, pp. 217-288. Sul mutamento ruolo del giudice

lavorare non è tanto la ricerca di un'intelligenza artificiale più coerente con i principi che reggono lo stato democratico costituzionale – che rimane imprescindibile – quanto piuttosto la formazione dei giudici che con questa intelligenza artificiale dovranno fare i conti e stanno già facendo i conti⁷⁸³. Se ora i giudici appaiono ancora riluttanti a “decodificare” l'intelligenza artificiale⁷⁸⁴, con una formazione adeguata avranno più strumenti per mantenere il controllo della decisione e – in fondo – per riappropriarsi del potere decisionale di cui stanno venendo spogliati a causa delle applicazioni algoritmiche. Occorre dunque che i giuristi, e in particolari i magistrati, «arricchiscano il loro patrimonio culturale con l'innesto di conoscenze tecnico-informatiche ... anche per acquisire una piena consapevolezza dei problemi, senza acritici entusiasmi né eccessivi timori che ostacolino l'evoluzione tecnologica»⁷⁸⁵.

A quest'ultimo riguardo, nel corso del lavoro abbiamo tentato di ricostruire la tenuta delle garanzie costituzionali del giusto processo di fronte al nuovo paradigma della giustizia algoritmica. Forti dell'interpretazioni delle Corti abbiamo cercato di dimostrare che le decisioni assunte (anche) tramite algoritmi rientrano nell'estensione applicativa dei principi costituzionali presi in considerazione, senza necessità di formulare nuove garanzie *specificamente* applicabili al nuovo contesto tecnologico.

in generale si cfr., *ex multis*, R. ROMBOLI, *Il ruolo del giudice in rapporto all'evoluzione del sistema delle fonti ed alla disciplina dell'ordinamento giudiziario*, cit.

⁷⁸³ Nella prospettiva della tutela *by education*, proposta da A. SIMONCINI in *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, cit., 88. In realtà l'A. utilizza tale espressione per sostenere la necessità di una formazione giuridica per i tecnici che progettano l'algoritmo. Invece, nella prospettiva qui proposta, viene impiegata per indicare la necessità di una formazione tecnica dei giudici.

⁷⁸⁴ In dottrina si è fatto notare che «*the court was disturbed at the lack of evidence it received as to exactly how this calculation had been done but, interestingly, did not seem interested to find out more but rather to exclude it from influence. It seems quite likely that courts will be reluctant to become activists about disclosures of source code, let alone algorithmic training sets and models, until they feel more confident of their ability to comprehend and use such evidence - which may take some time*» cfr. L. EDWARDS - M. VEALE, *Enslaving the Algorithm: From a “Right to an Explanation” to a “Right to Better Decisions”?*, in *IEEE Security & Privacy*, 16(3), 2018, p. 12. Si cfr. sul tema anche A. SELBST - S. VENKATASUBRAMANIAN - I. E. KUMAR, *Deconstructing design decisions: why courts must interrogate machine learning and other technologies*, cit.

⁷⁸⁵ G. CARLOTTI, *La giustizia predittiva e le fragole con la panna*, cit.

Ciò detto, forse la formulazione di un unico nuovo principio per guidare la rivoluzione formativa, organizzativa e culturale dei giudici può essere utile. Tale principio dovrà fondare l'eventuale decisione giudiziale "promiscua" (assunta cioè da un giudice con l'ausilio dell'algoritmo) e che si può mutuare, con qualche inevitabile adattamento, dall'aforismo *iura novit curia*. Come si presuppone che il giudice sia a conoscenza delle norme e abbia l'obbligo di farne applicazione anche in mancanza di una corretta individuazione delle parti, così questi dovrà conoscere l'algoritmo per poter individuare le storture e per poter eventualmente decidere di discostarsene: *algorithmia novit curia*.

Per concludere, solo assumendo come stella polare questo nuovo principio sarà immaginabile un futuro in cui la giustizia, razionalizzata e migliorata dall'impiego dell'intelligenza artificiale e dagli algoritmi, rimarrà appannaggio di un giudice umano, non riducibile a una macchina artificiale. In questo senso, anche seguendo lo spirito comparativo che ha avuto il presente lavoro, riportiamo due citazioni di due intellettuali che in tempi non sospetti hanno valorizzato la componente umana del giudice. Calamandrei scriveva che il giudice «non è un meccanismo, non è una macchina calcolatrice. È un uomo vivo: e quella funzione di applicare la legge è in realtà un'operazione di sintesi che si compie a caldo, misteriosamente, nel crogiuolo sigillato dello spirito, ove la saldatura tra la legge astratta e il fatto concreto ha bisogno, per compiersi, della intuizione e del sentimento acceso in una coscienza operosa». E ancora: «Ridurre la funzione del giudice a un puro sillogizzare vuol dire impoverirla, inaridirla, disseccarla. La giustizia è qualcosa di meglio: è creazione che sgorga da una coscienza viva, sensibile, vigilante, umana. È proprio questo calore vitale, questo senso di continua conquista, di vigile responsabilità che bisogna pregiare e sviluppare nel giudice»⁷⁸⁶. Similmente, il Giudice statunitense Frankfurter, in una sentenza risalente, affermava che «[t]o believe that this judicial exercise of judgment could be avoided by freezing "due process of law" [...] is to suggest that the most important aspect of constitutional

⁷⁸⁶ P. CALAMANDREI, *Processo e Democrazia*, in M. CAPPELLETTI (a cura di), *Opere giuridiche*, vol. I, Morano, 1965, p. 647.

adjudication is a function for inanimate machines and not for judges [...]. Even cybernetics has not yet made that haughty claim»⁷⁸⁷.

Mantenere vive queste parole e, insieme, dar loro un rinnovato significato attraverso l'introduzione graduale di algoritmi nel processo per rafforzare le garanzie costituzionali e migliorare la funzione giurisdizionale: questa è la sfida, a un tempo difficile ed entusiasmante, della giustizia algoritmica.

Siamo ancora agli albori di tale sfida epocale ma, come scrisse Antonio Machado in una bellissima poesia, «Viandante, sono le tue orme /il sentiero e niente più;/ viandante, non esiste il sentiero/, il sentiero si fa camminando»⁷⁸⁸.

⁷⁸⁷ U.S. Supreme Court, *Rochinv. California*, 342 U.S. 165, 171 (1952).

⁷⁸⁸ «Caminante, son tus huellas/el camino, y nada más;/ caminante, no hay camino:/se hace camino al andar», A. MACHADO, *Caminante*, 1998.

BIBLIOGRAFIA

ABOUFOUL M., *Despite Their Feats, Large Language Models Still Haven't Contributed to Linguistics. A review of Chomsky's views on linguistics and LLMs*, in *Towards Data Science*, 5 dicembre 2022.

ALETRAS N. - TSARAPATSANIS D. - PREOȚIUC-PIETRO D. - V. LAMPOS, *Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective*, in *PeerJ Computer Science*, 2, 2016.

ALZUBI J. - NAYYAR A. - KUMAR A., *Machine Learning from Theory to Algorithms: An Overview*, in *J. Phys.: Conf. Ser.*, 2018, pp. 1142 ss.

ANDRONIO A., *Art. 111*, in BIFULCO R. - CELOTTO A. - OLIVETTI M. (a cura di), *Commentario alla Costituzione*, Utet giuridica, Torino, 2006, pp. 2111 ss.

ANGWIN J.- LARSON J.- MATTU S.- KIRCHNER L., *Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks*, Propublica, 23 maggio 2016.

ARAUJO T. *et al.*, *In AI we trust? Perceptions about automated decision-making by artificial intelligence*, in *AI and Society*, 35, 2020, 611-623.

ARTOSI A. - BONGIOVANNI G. - VIDA S. (a cura di), *Problemi della produzione e dell'attuazione normativa*, III, *Analisi del linguaggio giuridico, legistica e legimatica*, Bologna, 2001.

ASHLEY K. D. - BRÜNINGHAUS S., *Automatically classifying case texts and predicting outcomes*, in *Artificial Intelligence and Law*, 17 (2), 2009, pp. 125-165.

ASHLEY K. D., *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*, in *Cambridge University Press*, 2017.

BALKIN M., *The three laws of robotics in the age of big data*, in *Ohio State Law Journal*, 78, 2017.

BARBARO C., *Usò dell'intelligenza artificiale nei sistemi giudiziari: verso la definizione di principi etici condivisi a livello europeo?*, in *Quest. giust.*, 4/2018, pp. 189 ss.

BASILE F., *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Dir. pen. e uomo*, 29 settembre 2019.

BASSU C. - BETZU M. - CLEMENTI F. - COINU G., *Diritto costituzionale degli Stati Uniti d'America. Una introduzione*, Giappichelli, 2022.

BEAU M., *Décret «DataJust»: stupeur et mécontentements*, in *Le Monde du Droit*, 6 aprile 2020.

BERK R. *et al.*, *Fairness in Criminal Justice Risk Assessments: The State of the Art*, in *Sociological Methods & Research*, 50 (1), pp. 3-44.

BERMUDEZ J. - NYRUP R. - DETERDING S. - MOUGENOT C. - MORADBAKHTI L. - YOU F. - CALVO R., *What is a subliminal technique?*, in *Ieee Ethics-2023 Conference Proceedings*, 2023.

BERTI R. - ZUMERLE F., *Prodigit: l'intelligenza artificiale che migliora la giustizia tributaria*, in *Agenda Digitale*, 24 luglio 2023.

BETZU M., *I poteri privati nella società digitale: oligopoli e antitrust*, in *Rivista di diritto pubblico*, 3, 2021, pp. 739-760.

BETZU M., *poteri pubblici e poteri privati nel mondo digitale*, in *Rivista Gruppo di Pisa*, 2, 2021, pp. 166-191.

BEX F. - PRAKKEN H., *Can predictive justice improve the predictability and consistency of judicial decision-making?*, in SCHWEIGHOF E. (a cura di), *Legal knowledge and Information Systems*, Amsterdam, 2021, pp. 207-213.

BICHI R., *Intelligenza digitale, giurmetria, giustizia predittiva e algoritmo decisorio. Machina sapiens e il controllo sulla giurisdizione*, in RUFFOLO U. (a cura di), *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Milano, 2020, pp. 423-447.

BILANCIA P., PIZZETTI F.G., *Aspetti e problemi del costituzionalismo multilivello*, Milano, Giuffrè, 2004.

BLACK J., *Risk-based regulation: choices, practices and lessons being learnt. Risk and Regulatory Policy*, in *Improving the Governance of Risk*, OECD, Paris, 2010.

BONDERO B., *Justice prédictive: la fin de l'aléa judiciaire?*, in *Recueil Dalloz*, n. 10, 2017, pp. 532 ss.

BOSTROM N., *Superintelligence: path, dangers, strategies*, Oxford University Press, New York, 2014.

BRADFORD A., *Brussels Effect. How the European Union rules the world*, Oxford University Press, 2020.

BREMMER I. - SULEYMAN M., *The AI Power Paradox: Can States Learn to Govern Artificial Intelligence—Before It's Too Late?*, in *Foreign Affairs*, 2023.

BRENNER M. *et al*, *Constitutional Dimensions of Predictive Algorithms in Criminal Justice*, in *Harvard Civil Rights*, 55, 2020.

BREWER S., *Scientific Expert Testimony and Intellectual Due Process*, in *Yale Law Journal*, 107(6), 1998, pp. 1531-1681.

BROWN T. - MANN B. - RYDER N. *et al.*, *Language models are few-shot learners*, in *Advances in neural information processing systems*, 2020.

BUCHANAN G.- HEADRICK THOMAS T.E., *Some Speculation About Artificial Intelligence and Legal Reasoning*, in *Harvard Law Review*, 23, 1970, 40 ss.

BURCHARD C., *L'intelligenza artificiale come fine del diritto penale? Sulla trasformazione algoritmica della società*, in *Riv. it. Dir. Proc. Pen.*, vol. 4, 2019.

CALAMANDREI P., *Opere giuridiche. Volume I. Problemi generali del diritto e del processo*, RomaTre-Press, 2019.

CALAMANDREI P., *Processo e Democrazia*, in CAPPELLETTI M. (a cura di), *Opere giuridiche*, vol. I, Morano, 1965.

CALO R.-. CITRON D.K., *The Automated Administrative State: A Crisis of Legitimacy*, in *Emory Law Journal*, 70, 4, 2021, pp. 799-845.

CARLEO A. (a cura di), *Calcolabilità giuridica*, Il Mulino, Bologna, 2017.

CARLEO A. (a cura di), *Decisione robotica*, Il Mulino, Bologna, 2019.

CARLEO A. (a cura di), *Il vincolo giudiziale del passato*, Bologna 2018.

CARLOTTI G., *La giustizia predittiva e le fragole con la panna*, in *Giustizia amministrativa*, 2023.

CARRATTA A., *Decisione robotica e valori del processo*, in *Riv. dir. proc.*, 2/2020.

CARULLO G., *Decisione amministrativa e intelligenza artificiale*, in *Diritto dell'informazione e dell'informatica*, n. 3, 2021.

CASONATO C. - MARCHETTI B., *Prime osservazioni sulla proposta di regolamento dell'Unione Europea in materia di intelligenza artificiale*, in *BioLaw Journal – Rivista di BioDiritto*, n. 3/2021, pp. 415-437.

CASONATO C., *Giustizia e intelligenza artificiale: considerazioni introduttive*, in *BioLaw Journal – Rivista di BioDiritto*, 2, 2021.

CASONATO C., *Intelligenza artificiale e giustizia: potenzialità e rischi*, in *DPCE online*, 2020/3, p. 3369-3388.

CASONATO C., *L'intelligenza artificiale e il diritto pubblico comparato ed europeo*, in *DPCE Online*, 1/2022.

CASTELLI C. - PIANA D., *Giustizia predittiva. La qualità della giustizia in due tempi*, in *Questione giustizia*, trimestrale n. 4, 2018.

CASTELLI C., *Giustizia Predittiva: i progetti in corso in Italia*, in *Agenda Digitale*, 2 agosto 2023.

CASTELLI C., *La giustizia predittiva*, in *Lamagistratura.it*, 11 novembre 2021.

CASTETS-RENARD C., *Human Rights and Algorithmic Impact Assessment for Predictive Policing*, in MICKLITZ H. - POLLICINO O. - REICHMAN A. - SIMONCINI A. - SARTOR G. - DE GREGORIO G. (a cura di), *Constitutional Challenges in the Algorithmic Society*, Cambridge University Press, 2022.

CATH C. et al., *Artificial Intelligence and the 'Good Society', the US, EU, and UK approach*, in *Sci Eng Ethics*, 2018.

CAVAGGION G. - OROFINO M., *Lingua e costituzione: l'irrompere dei linguaggi algoritmici*, in *Rivista AIC*, n. 4/2023, pp. 155-184.

CECCHETTI M., *La riforma dell'art. 111 Cost.: tra fonti preesistenti, modifica della Costituzione, diritto intertemporale ed esigenze di adeguamento della legislazione ordinaria*, in DE SIERVO U. (a cura di), *Osservatorio sulle fonti*, Torino, Giappichelli, 2001, pp. 35-66.

CHANG Y. - WANG X. - WANG J. *et al.*, *A survey on evaluation of large language models*, in *ACM Transactions on Intelligent Systems and Technology*, 2018.

CITRON D. K. - PASQUALE F., *The scored society: Due process for automated predictions*, in *Washington Law Review*, 89, 1, 2014, pp. 1-33.

CITRON D.K., *Technological Due Process*, in 85 *Wash. U.L. Rev.*, 2008, p. 1261.

CLARK B.B. - ROBERT C. - HAMPTON S. A., *The Technology Effect: How Perceptions of Technology Drive Excessive Optimism*, in 31 *Journal of Business and Psychology*, 2016, pp. 87-102.

COGLIANESE C.- LEHR D., *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, in *University of Pennsylvania Carey Law School*, All Faculty Scholarship, vol. 105, pp. 1147 ss.

COGLIANESE C., *AI, Due Process, and Trade Secrets*, in *The Regulatory Review*, 2023.

COLLINGRIDGE D., *The Social Control of Technology*, New York, 1980.

CONTI G., *Il sistema di tutela multilivello e l'interazione tra ordinamento interno e fonti sovranazionali*, in *Questione Giustizia*, n. 4/2016, pp. 89-114.

CONTISSA G. - LASAGNI G. - SARTOR G., *Quando a decidere in materia penale sono (anche) algoritmi e IA: alla ricerca di un rimedio effettivo*, in *Diritto di internet*, n. 4/2019, pp. 619-634.

COTINO HUESO L., *The Council of Europe's Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*, in *CERIDAP*, 9 settembre 2024.

COWGER, JR. A. R., *The Threats of Algorithms and AI to Civil Rights, Legal Remedies, and American Jurisprudence: One Nation Under Algorithms*, London, 2020, p. 14.

CUSENZA G. G., *Litigating Governmental Use Of AI*, in *Ohio State Technology Law Journal*, 2024.

D'ALOIA A., *Il diritto verso il "mondo nuovo". Le sfide dell'intelligenza artificiale*, in *Biolaw Journal – Rivista di Biodiritto*, n. 1/2019.

D'AMATO A., *Can/Should Computers Replace Judges?*, in *Georgia Law Review*, 11, 1977, pp. 1277-1301.

DE GREGORIO G., *Digital Constitutionalism in Europe. Reframing Rights and Powers in the Algorithmic Society*, Cambridge, 2022.

DE MINICO G., *Giustizia e intelligenza artificiale: un equilibrio mutevole*, in *Rivista AIC*, n. 2/2024.

DEEKS A., *The judicial demand for explainable artificial intelligence*, in *Columbia Law Review*, Vol. 119, No. 7, pp. 1829- 1850.

DEMICHELE M. - COMFORT M. - BARRICK K. - BAUMGARTNER P., *The Intuitive-Override Model: Nudging Judges Toward Pretrial Risk Assessment Instruments*, in *Federal Probation Journal*, vol. 85, n. 2, 2021.

DI GIOVINE O., *Il judge bot e le sequenze giuridiche in materia penale (intelligenza artificiale e stabilizzazione giurisprudenziale)*, in *Cass. pen.*, 2020, pp. 951 ss.

DI PORTO A., *Calcolo giuridico secondo la legge nell'età della giurisdizione, il ritorno del testo normativo*, in CARLEO A. (a cura di), *Calcolabilità giuridica*, Il Mulino, Bologna, 2017.

DI PORTO A., *Calcolo giuridico secondo la legge nell'età della giurisdizione, il ritorno del testo normativo*, in CARLEO A. (a cura di), *Calcolabilità giuridica*, Il Mulino, Bologna, 2017.

DONATI F., *Diritti fondamentali e algoritmi nella proposta di regolamento sull'intelligenza artificiale*, in PAJNO A. - DONATI F. - PERRUCCI A., *Intelligenza artificiale e diritto: una rivoluzione?*, volume I, 2022, pp. 111-126.

DONATI F., *Intelligenza artificiale e giustizia*, in *Rivista AIC*, 1/2020, pp. 415-436.

DONATI F., *Trasparenza della giustizia e anonimizzazione dei provvedimenti giudiziari*, in PAJNO A. - DONATI F. - PERRUCCI A. (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, volume 2, 2022, pp 433-452.

DUFOUR O. - CORVIAUX A., *Datajust: Plutôt que de faire de la justice prédictive, il faut engager une démarche d'indexation et de tri des décisions*, in *Actu-Juridique*, 27 gennaio 2022.

DUMOULIN L., *De quoi la 'justice prédictive' est-elle le nom? Algorithmes, décision et jugement*, Science politique, Université Paris-Saclay, 2022.

EBERS M. - TUPAY P. K., *Due Process, Fair Trial, Transparency, and Explainability*, in *ID, Artificial Intelligence and Machine Learning Powered Public Service Delivery in Estonia: Opportunities and Legal Challenges*, 2023.

EBERS M., *Standardizing AI - The Case of the European Commission's Proposal for an Artificial Intelligence Act*, in DI MATTEO L.A. - CANNARSA M. - PONCIBÒ C. (a cura di), *The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics*, Cambridge, 2022.

EDWARDS L. - VEALE M., *Enslaving the Algorithm: From a "Right to an Explanation" to a "Right to Better Decisions"?*, in *IEEE Security & Privacy*, 16 (3), 2019, pp. 46-54.

ELARDO D.J., *A Case for Computers in Law Practice* in *Cleveland State Law Rev.*, 17(3), 1968, pp. 552-566.

FALLETTI E., *L'Artificial Intelligence Act Proposal e la regolamentazione degli algoritmi predittivi: luci e ombre*, in *CERIDAP*, 4/2023, pp. 173-217.

FEIGENSON N., *Brain Imaging and Courtroom Evidence: On the Admissibility and Persuasiveness of fMRI*, in *International Journal of Law in Context*, n. 3, 2006.

FELDSTEIN S., *Evaluating Europe's push to enact AI regulations: how will this influence global norms?*, in *Democratization*, 31, 5, 2024.

FERRANTI I., *Prodigit, come funziona il progetto per la giustizia tributaria digitale*, in *Agenda Digitale*, 10 Marzo 2023.

FERRI G., *Autonomia e indipendenza della magistratura tra 'vecchio' e 'nuovo' ordinamento giudiziario*, in *Rivista AIC*, n. 4, 2017, pp. 1-52.

FINOCCHIARO G., *La regolazione dell'intelligenza artificiale*, in *Riv. Trim. Dir. Pubbl.*, 4, 2022.

FLORIDI L., *Establishing the Rules for Building Trustworthy AI*, in L. FLORIDI (a cura di) *Ethics, Governance, and Policies in Artificial Intelligence. Philosophical Studies Series*, vol 144. Springer, 2021.

FLORIDI L., *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Raffaello Cortina Editore, Milano, 2017.

FLORIDI L., *The European Legislation on AI: A Brief Analysis of its Philosophical Approach*, in *Philos. Technol*, 34, 2021, pp. 215-222.

FORMISANO A., *L'impatto dell'intelligenza artificiale in ambito giudiziario sui diritti fondamentali*, in *Federalismi*, n. 22/2024, pp. 112-149.

FORREST K. B., *When machines can be judge, jury, and executioner: justice in the age of artificial intelligence*, World Scientific Publishing Co. Pte. Ltd., Singapore, 2021.

FRANKEL M., *Criminal Sentences: Law Without Order*, Hill and Wang, New York, 1973.

FREEMAN K., *Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis*, in *North Carolina Journal of Law & Technology*, 18, 2016.

FRIDMAN L. - DING L. - JENIK B. - REIMER B., *Arguing Machines: Human Supervision of Black Box AI Systems That Make Life-Critical Decisions*, 2019.

FRIEDLER S. - HAVEN J. - J. CHEN, *How the AI Executive Order and OMB memo introduce accountability for artificial intelligence*, in *Brookings.edu*, 16 novembre 2023.

FRIEDLER S. A. - SCHEIDEGGER C.- VENKATASUBRAMANIAN S., *On the (im)possibility of fairness*, in *ArXiv*, 23 settembre 2016.

FRIENDLY, H.J. *Some Kind of Hearing*, in *123 University of Pennsylvania Law Review*, 1975, pp. 1267 ss.

FRONZA E.- CARUSO C. (a cura di), *Ti faresti giudicare da un algoritmo. Intervista ad Antoine Garapon*, in *Questione Giustizia*, trimestrale, n. 4, 2018, pp. 189-199.

G'SELL F., *Les progrès à petits pas de la «justice prédictive» en France*, in *Era Forum*, 2020, pp. 299 ss.

GABORIAU S., *Libertà e umanità del giudice: due valori fondamentali della giustizia. La giustizia digitale può garantire nel tempo questi valori?*, in *Questione giustizia*, trimestrale n. 4, 2018.

GALETTA D.U., *Algoritmi, procedimento amministrativo e garanzie: brevi riflessioni, anche alla luce degli ultimi arresti giurisprudenziali in materia*, in *Rivista Italiana di Diritto Pubblico Comunitario*, 3, 2020, pp. 501 ss.

GALLI F. - SARTOR G., *AI Approaches to Predictive Justice: A Critical Assessment*, in *Humanities And Rights Global Network Journal*, 5, 2023.

GARAPON A. - LASSÈGUE J., *Justice digitale*, PUF, Paris, 2018.

GARAPON A., *La despazializzazione della giustizia*, Milano-Udine, 2021.

GARAPON A., *Les enjeux de la justice predictive*, in *La semaine juridique*, 1, 2017, pp. 47-52.

GARRETT B. L. - MONAHAN J., *Judging Risk*, in *California Law Review*, vol. 108, 2020.

GARRIDO MARTIN E. - HERRERO ALONSO C., *Influence of the prosecutor's Plea on the Judge's Sentencing in Sexual Crimes: Hypothesis of the Theory of Anchoring by Tversky and Kahneman*, 1997.

GAUDEMET Y., *La justice à l'heure des algorithmes. À propos de justice prédictive*, in *Revue du droit public et de la science politique en France et à l'étranger*, n. 3, 2018, pp. 651 ss.

GAUR M., *An In-Depth Exploration of Predictive Justice with AI*, in *AI and Emerging Technologies, Automated Decision-Making, and Ethical Considerations*, 2024.

GIALUZ M., *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti ed Europa*, in *Diritto penale contemporaneo*, 29 maggio 2019.

GIALUZ M., *Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra Stati Uniti e Europa*, in *Dir. Pen. Cont.*, 29 maggio 2019.

GRAVETT W., *Jailed by a "black box": the impact of opaque algorithms on the right to a fair trial in the Unites State of America*, in 84 *THRHR*, 2021, pp. 299 ss.

GREEN A., *Humans May Put Too Much Trust in Robots, Study Finds*, in *Mental Floss*, marzo 2016.

GROSSI P., *Ritorno al diritto*, Editori Laterza, Roma-Bari, 2015.

GROSSI P., *Storicità versus prevedibilità: sui caratteri di un diritto pos-moderno*, in *Questione giustizia*, trimestrale n. 4, 2018, pp. 17-23.

GUNNINGHAM N. - REES J., *Industry Self-Regulation: An Institutional Perspective*, in *Law & Policy*, 19, 1997.

GUTIÉRREZ J.D., *ChatGPT in Colombian Courts: Why we need to have a conversation about the digital literacy of the judiciary*, in *VerfBlog*, 23 febbraio 2023.

GUTIÉRREZ J.D., *Judges and Magistrates in Peru and Mexico Have ChatGPT Fever*, in *Tech Policy Press*, 19 aprile 2023.

HAATAJA M. - BRYSON J., *The European Parliament's AI regulation: should we call it progress?*, in *Amicus Curiae*, Series 2, Vol 4, No 3, 2023, pp. 707-718.

HAMILTON M., *Risk and Needs Assessment: Constitutional and Ethical Challenges*, in *American Criminal Law Review*, vol. 52, 2015.

HELBERGER N. - DIAKOPOULOS N., *ChatGPT and the AI Act*, in *Internet Policy Review*, 12 (1), 2023.

HICKMAN E. - PETRIN M., *Trustworthy AI and Corporate Governance – The EU's Ethics Guidelines For Trustworthy Artificial Intelligence from a Company Law Perspective*, in *European Business Organization Law Review*, 2021, pp. 593-625.

HILDEBRANDT M., *Technology and the end of law*, in B. KEIRSBILCK - W. DEVROE - E. CLAES (a cura di), *Facing the limits of the law*, Springer, 2009, pp. 443-464.

HINE E. - FLORIDI L., *Artificial intelligence with American values and Chinese characteristics: a comparative analysis of American and Chinese governmental AI policies*, in *AI and Society*, 39, 2022.

HINE E. - FLORIDI L., *The Blueprint for an AI Bill of Rights: In Search of Enaction, at Risk of Inaction*, in *Minds and Machines*, 30 novembre 2022.

HSU T. - THOMPSON S.A., *Disinformation Researchers Raise Alarms About A.I. Chatbots*, in *New York Times*, 20 giugno 2023.

HUQ A. Z., *Racial equity in algorithmic criminal justice*, in *Duke Law Journal*, vol. 68 (6), 2019.

HUQ A. Z., *A Right to a Human Decision*, in *105 Virginia Law Review*, 2020.

HUQ A. Z., *Constitutional Rights in the Machine Learning State*, in *Cornell Law Review*, Vol. 105, 2020.

IANNOTTI DELLA VALLE A., *Le regole di internet tra poteri pubblici e privati. Tutela dei diritti e ruolo dell'antitrust in una prospettiva costituzionale*, Editoriale Scientifica, Napoli, 2023.

IRTI N., *L'età della decodificazione*, Giuffrè, Milano, 1999.

IRTI N., *La crisi della fattispecie*, in ID., *Un diritto incalcolabile*, Torino, Giappichelli, 2016.

J. SAKS M. - KIDD R. F., *Human Information Processing and Adjudication: Trial by Heuristics*, in *Law & Society Review*, Vol. 15, No. 1, 1980, pp. 123-160.

JESSUP T., *'I think therefore I AM'? - Artificial intelligence and the Chinese Social Credit System: A Human Rights Critical Analysis*, 17 maggio 2021.

JOHNSON D. R. - POST D., *Law and Borders: The Rise of Law in Cyberspace*, in *Stanford Law Review*, 48, 5, 1996.

KAHNEMAN D. - SIBONY O. -. SUNSTEIN C. R., *Noise: A Flaw in Human Judgement*, trad. it. *Rumore: un difetto del ragionamento umano*, Milano, 2021.

KAPLAN J., *Generative Artificial Intelligence: What Everyone Needs to Know*, Oxford, 2024.

KATZ D. M. - BOMMARITO M. J. - J. BLACKMAN, *A general approach for predicting the behavior of the Supreme Court of the United States*, in *Plos One*, n. 4/2017.

KATZ D.M., *Quantitative Legal Prediction--or--How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry*, in *Emory L.J.*, n. 4, 2012, pp. 906-966.

KATZ Y., *Artificial Whiteness: Politics and Ideology in Artificial Intelligence*, in *Columbia University Press*, New York, 2020.

KEHL D. - GUO P. - KESSLER S., *Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing*, in *Responsive Communities Initiative, Berkman Klein Center for Internet & Society*, Harvard Law School, 2017.

KERIKMÄE T. - HAMUĽÁK O. - GÁBRIŠ T., *Frontiers in AI Judiciary: A Contribution to Legal Futurology*, in *Acta Baltica Historiae et Philosophiae Scientiarum*, 11, 2, 2023.

KHDER M. A., *Web Scraping or Web Crawling: State of Art, Techniques, Approaches and Application*, in *International Journal of Advances in Soft Computing & Its Applications*, 13 (3), 2021.

KIESOW E. - MASLEY N., *Adjudication of artificial intelligence and automated decision-making cases in Europe and the USA*, in *European Journal of Risk Regulation*, 2023, pp. 1-19.

KLEINBERG J.- LUDWIG J. - MULLAINATHANY S. - SUNSTEIN C. R., *Discrimination in the age of algorithms*, in *Journal of Legal Analysis*, vol. 10, 2018.

KOSTORIS E., *Intelligenza artificiale, strumenti predittivi e processo penale*, in *DisCrimen*, 5 marzo 2024.

KRIEBEL D. et al., *The Precautionary Principle in Environmental Science*, in *Environmental Health Perspectives* 109 (9), 2001, pp. 871-876.

LALLI A., *I non radicalmente diversi approcci alla regolazione dell'intelligenza artificiale di Europa, Stati Uniti e Cina*, in LALLI A. (a cura di), *La regolazione pubblica delle tecnologie digitali e dell'intelligenza artificiale*, Giappichelli, 2024, pp. 115-140.

LAWLOR R.C., *What Computers Can Do: Analysis and Prediction of Judicial Decisions*, in *Am Bar Assoc J*, 49 (4), 1963, pp. 337-344.

LEE N. T. - CHIJOKE O., *Why states and localities are acting on AI*, in *Brookings.edu*, 15 dicembre 2023.

LEE N. T. - MALAMUD J., *Opportunities and blind spots in the White House's blueprint for an AI Bill of Rights*, in *Brookings.edu*, 19 dicembre 2022.

LEE N.T. - CHIJOKE O., *Why states and localities are acting on AI*, in *Brookings.edu*, 15 dicembre 2023.

LEE N.T., *How the White House Executive Order on AI ensures an effective governance regime*, in *Brookings.edu*, 28 marzo 2024.

LIBERTINI M. - ROSARIA MAUGERI M. - VINCENTI E., *Intelligenza artificiale e giurisdizione ordinaria. una ricognizione delle esperienze in corso*, in PAJNO A. -

DONATI F. - PERRUCCI A. (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, volume 2, 2022, pp. 483-514.

LICOPPE C. - DUMOULIN L., *Le Travail des juges et les algorithmes de traitement de la jurisprudence. Premières analyses d'une expérimentation de "justice prédictive" en France*, in *Droit et Société*, n. 103, 2019, p. 535-554.

LIU H. - LIN C. - CHEN Y., *Beyond State v Loomis: artificial intelligence, government algorithmization and accountability*, in *International Journal of Law and Information Technology*, 2019, vol. 27, pp. 122-141.

LO SAPIO G., *L'Artificial Intelligence Act e la prova di resistenza per la legalità algoritmica*, in *Federalismi*, n. 16/2024, pp. 265-290.

LONGO E., *Giustizia digitale e Costituzione: riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, Milano, 2023.

LONGO E., *La "giustizia digitale" tra nuove forme di efficienza e garanzia dei diritti costituzionali*, in *La Nuova Giuridica - Florence Law Review*, 2023, 2, pp. 187-209.

LONGO E., *Predizione algoritmica e giustizia: le prospettive costituzionali*, in PANZERA C. - RAUTI A. (a cura di), *Attualità di diritto pubblico*, II, Editoriale Scientifica, 2022, pp. 87-125, spec. pp. 112-115.

LOUVEL B., *Communication à l'occasion du colloque pour le bicentenaire de l'ordre des avocats au Conseil d'État et à la Cour de Cassation*, in *Ordre des avocats au Conseil d'État et à la Cour de cassation, La Justice prédictive*, Paris, 2018.

LUCIANI M., *La decisione giudiziaria robotica*, in *Rivista AIC*, n. 3/2018.

MAGLIONE V., *Pisa allena l'algoritmo che «prevede» le sentenze*, in *Norme e Tributi Plus*, 2021.

MALDONATO L., *Algoritmi predittivi e discrezionalità del giudice: una nuova sfida per la giustizia penale* in *Diritto penale contemporaneo*, 2/2019, p. 410.

MANES V., *L'oracolo algoritmico e la giustizia penale*, in *DisCrimen*, 15 maggio 2020.

MARCHETTI B., *La garanzia dello human in the loop alla prova della decisione amministrativa algoritmica*, in *BioLaw Journal – Rivista di BioDiritto*, n. 2/2021.

MARION O. *et al.*, *Algorithmic risk assessment policing models: lessons from the Durham HART model and ‘Experimental’ proportionality*, in *Information & Communications Technology Law*, 2/2018, 223-250.

MARJANOVIC O. - CECEZ-KECMANOVIC D. - VIDGEN R., *Theorizing algorithmic justice*, in *European Journal of Information Systems*, 31(3), 269-287.

MASELLA G., “*Giurisprudenza predittiva*”. *Risultati operativi e prospettive future*, in *L’esperienza Uni4Justice e le prospettive future Le ricerche del team di Ca’ Foscari*, 2023, pp. 279-292.

MATTEUCCI S.C., “*«Umano troppo umano»*. *Decisioni amministrative automatizzate e principio di legalità*, in *Riv. Dir. Pubbl.*, 1/2019.

MCCARTHY J. - MINSKY M. - ROCHESTER N. - SHANNON C., *A proposal for the Dartmouth summer research project on artificial intelligence*, 1955.

MEDVEDEV M. - VOLS M. - M. WIELING, *Using Machine Learning to Predict Decisions of the European Court of Human Rights*, in *Artificial intelligence and Law*, vol. 28, 2020, pp. 237-266.

MINGARDO L., *Online Dispute Resolution. Involuzioni ed evoluzioni di telematica giuridica*, in MORO P. - SARRA C. (a cura di), *Tecnodiritto: temi e problemi di informatica e robotica giuridica*, Milano 2017, pp. 121 ss.

MOBILIO G., *L’intelligenza artificiale e i rischi di una “disruption” della regolamentazione giuridica*, in *BioLaw Journal – Rivista di BioDiritto*, n. 2/2020, pp. 401-424.

MOOBS D., *The Implications of Brain Imaging Studies for the Law*, presentato alla conferenza *Law, Mind and Brain interdisciplinary colloquium*, University College London, febbraio 2006.

MORISON J. - HARKENS A., *Re-engineering justice? Robot judges, computerised courts and (semi) automated legal decision-making*, in *Legal Studies*, vol. 39 (4), 2019.

MORTATI C., *Note sul potere discrezionale*, 1936, ora in *Problemi di diritto pubblico nell'attuale esperienza costituzionale repubblicana*, Raccolta di scritti, vol. III, 1972.

MUSSELLI L., *La decisione amministrativa nell'età degli algoritmi: primi spunti*, in *MediaLaws-Rivista di diritto dei media*, 1/2020.

NIEVA-FENOLL J., *Intelligenza artificiale e processo*, trad.it di P. COMOGLIO, Giappichelli, Torino, 2019, pp. 38 ss.

NOURI Z.- SALAH W. B. -. OMRANE N. A., *Artificial Intelligence and Administrative Justice: An Analysis of Predictive Justice in France*, in *Hasanuddin Law Review*, 10 no. 2, 2024.

ORWAT C., *Risks of Discrimination through the Use of Algorithms*, in *Federal Anti-Discrimination Agency*, 2020.

OTRANTO P., *Riflessioni in tema di decisione amministrativa, intelligenza artificiale e legalità*, in *Federalismi*, n. 7/2021.

PACK H., *Regulation of Artificial Intelligence in the United States*, in *Interactive Robotics: Legal, Ethical, Social and Economic Aspects*, Springer International Publishing, 2018.

PAJNO A. et al., *AI: profili giuridici: Intelligenza Artificiale: criticità emergenti e sfide per il giurista*, in *BioLaw Journal – Rivista di BioDiritto*, n. 3/2019.

PAJNO A., *Intelligenza artificiale e autonomia del giudice. il ruolo del fatto e il valore del precedente*, in PAJNO A. - DONATI F. - PERRUCCI A. (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, volume 2, 2022, pp. 389-432.

PARDOLESI R. - DAVOLA A., *Algorithmic legal decision making: la fine del mondo (del diritto) o il paese delle meraviglie?*, in *Questione Giustizia*, trimestrale n. 1, 2020 pp. 104 ss.

PARK E., *The AI Bill of Rights: A Step in the Right Direction*, in *Orange County Lawyer Magazine*, Vol. 65, 2, 2023.

F. PASQUALE, *The Black Box Society*, Harvard University Press, 2015.

PASQUALE F.- CASHWELL G., *Prediction, persuasion, and the jurisprudence of behaviourism*, in *Univ. of Toronto L.J.*, vol. 68 (suppl. 1), 2018.

PASQUALE F., *Secret Algorithms Threaten the Rule of Law*, in *MIT Technology Review*, 1 giugno 2017.

PATRONE J., *Il prosecutor negli Stati Uniti. Un esempio da seguire?*, in *Questione Giustizia*, 11 marzo 2021.

PATRONI GRIFFI F., *Decisione robotica e giudice amministrativo*, in A. CARLEO (a cura di), *Decisione robotica*, Il Mulino, Bologna, 2019.

PENASA S., *Intelligenza artificiale e giustizia: il delicato equilibrio tra affidabilità tecnologica e sostenibilità costituzionale in prospettiva comparata*, in *DPCE Online*, 1/2022.

PERES E., *Che cosa sono gli algoritmi*, Salani, Firenze, 2020.

PERONA R., *ChatGPT e decisione giudiziale: per un primo commento alla recente sentenza del Juzgado Primero Laboral di Cartagena de Indias (Colombia)*, in *Diritti Comparati*, 21 febbraio 2023.

PISANO G. et al., *La sommarizzazione delle sentenze tributarie di merito nel Progetto PRODIGIT*, in *Rivista di Scienze Giuridiche, Scienze Cognitive ed Intelligenza Artificiale*, vol. 17 (1), 2024.

PIZZETTI F., *La proposta di Regolamento sull'IA della Commissione Europea presentata il 21.4.2021 (COM (2021) 206 final) tra Mercato Unico e competizione digitale globale*, in *Diritto di Internet* n. 4/2021., pp. 591 ss.

POLICE A., *Art. 24*, in BIFULCO R. - CELOTTO A. - OLIVETTI M. (a cura di), *Commentario alla Costituzione*, Utet giuridica, Torino, 2006, pp. 502 ss.

POLLICINO O. - G. DE GREGORIO, *Constitutional Law in the Algorithmic Society*, in MICKLITZ H. - POLLICINO O. - REICHMAN A. - SIMONCINI A. - SARTOR G. - DE GREGORIO G. (a cura di), *Constitutional Challenges in the Algorithmic Society*, Cambridge University Press, 2022.

POSNER R. A., *The Role of The Judge on the Twenty-First Century*, in 86 *Boston University Law Review*, pp. 1049- 1068.

POUND R., *Interpretations of legal history*, The Macmillan Company, New York, 1923.

POUND R., *Mechanical Jurisprudence*, in *Columbia Law Review*, 8 605-623, 1908.

PUNZI A., *Judge in the machine. E se fossero le macchine a restituirci l'umanità nel giudicare?* In A. CARLEO (a cura di) *Decisione robotica*, Il Mulino, Bologna, 2019.

QUATTROCOLO S., *Artificial Intelligence, Computational Modelling and Criminal Proceedings. A Framework for A European Legal Discussion*, 2020.

QUATTROCOLO S., *Equità del processo penale e automated evidence alla luce della convenzione europea dei diritti dell'uomo*, in *Revista Ítalo-Española de Derecho Procesal*, 2019, pp. 107-123.

QUATTROCOLO S., *Equo processo penale e sfide della società algoritmica*, in *BioLaw Journal – Rivista di BioDiritto*, n. 1/2019, pp. 135-144.

QUATTROCOLO S., *Intelligenza artificiale e giustizia: nella cornice della Carta etica europea gli spunti per un'urgente discussione tra scienze penali e informatiche*, in www.lalegislazionepenale.it, 18 dicembre 2018.

QUATTROCOLO S., *Quesiti nuovi e soluzioni antiche? Consolidati paradigmi normativi vs rischi e paure della giustizia digitale "predittiva"*, in *Cassazione penale*, 4, 2019, 1748 ss.

RANDAZZO B., *Access to Justice in a Multilevel Constitutional System, Protecting Human Rights*, Giappichelli, 2024.

RANDAZZO B., *Accesso alla giustizia e diritti fondamentali nella dimensione costituzionale nazionale e sovranazionale*, in AMALFITANO C. - GALETTA D.U. - VIOLINI L., "Law, Justice and sustainable development". *L'accesso alla giustizia nel quadro del Sd goal 16*, 2022 pp. 117-138.

RANDAZZO B., *Giudici comuni e Corte europea dei diritti*, in *Atti del Seminario del “Gruppo di Pisa”, La Corte costituzionale e le Corti d’Europa*, (Catanzaro, 31 maggio – 1° giugno 2002), Giappichelli, Torino 2003, pp. 217-288.

RANDAZZO B., *Il ‘riaccentramento’ del giudizio costituzionale nella prospettiva di un sistema integrato di giustizia costituzionale*, in *Federalismi*, 3, 2021, pp. 144-160.

RANDAZZO B., *Imparzialità del giudice e fiducia nella magistratura nella prospettiva costituzionale europea*, in *Questione Giustizia*, 1/2/2024, pp. 231-241.

RANDAZZO B., *Interpretazione delle sentenze della Corte europea dei diritti ai fini dell’esecuzione (giudiziaria) e interpretazione della sua giurisprudenza ai fini dell’applicazione della CEDU*, in *Rivista AIC*, 2/2015.

REILING D., *Courts and Artificial Intelligence*, in *International Journal for Court Administration*, 8, 2020.

REILING D., *Technology for Justice. How Information Technology Can Support Judicial Reform*, Leiden, 2009.

RESTA G., *Governare l’innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza*, in *Politica del diritto*, 2019.

RICHARDS B. - AGÜERA B. - ARCAS Y.- G. LAJOIE - D. SRIDHAR, *The illusion of AI’s existential risk*, in *Noema Magazine*, 18 luglio 2023.

RIVOLLIER V., *Datajust. Histoire d’un échec. Séminaire Nouvelles technologies et justice*, Centre internet et société, marzo 2023.

ROBERTS H. - COWLS J. - HINE E., *et al*, *Achieving a ‘Good AI Society’: Comparing the Aims and Progress of the EU and the US*, in *Sci Eng Ethics* 27, 68, 2021.

ROMBOLI R., *Il ruolo del giudice in rapporto all’evoluzione del sistema delle fonti ed alla disciplina dell’ordinamento giudiziario*, in *Associazione per gli studi e le ricerche parlamentari*, Quaderno n. 16, Torino, 2006, pp. 63-87.

RUFFOLO U., *Giustizia predittiva, decisione algoritmica e ruolo ancillare della machina come ausiliare necessario del giudice umano*, in PAJNO A. - DONATI F. -

PERRUCCI A. (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, Bologna, Il Mulino, 2022, pp. 453-481.

SALAZAR C., *Judex ex machina? Note su giustizia, giudici e intelligenza artificiale*, in *ConsultaOnline*, 18 novembre 2021.

SAMOILI S. - LÓPEZ COBO M. - GÓMEZ E. - DE PRATO G. - MARTÍNEZ-PLUMED F. - DELIPETREV B., *AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence*, Publications Office of the European Union, Lussemburgo, 2020.

SANTOSUOSSO A. - SARTOR G., *La giustizia predittiva: una visione realistica*, in *Giur. It.*, 7, 2022, pp. 1759 ss.

SARTOR G., *Intelligenza artificiale e diritto. Un'introduzione*, Giuffrè, Milano, 1996.

SCHEPISI C., *Le "dimensioni" della regolazione dell'intelligenza artificiale nella proposta di regolamento della Commissione*, in *I Post di AISDUE*, IV, 16, 28 marzo 2022.

SCHERER M., *Artificial Intelligence and legal Decision-Making: the wide open?*, in *Journal of International Arbitration*, 36, 5, 2019, pp. 563 ss.

SCHUILENBURG M. - PEETERS R. (a cura di), *The Algorithmic Society. Technology, Power, and Knowledge*, Abingdon-New York, 2021.

A. SELBST, *Meaningful information and the right to explanation*, in *7(4) International Data Privacy Law*, 2017.

SELBST A. - BOROCAS S., *The Intuitive Appeal of Explainable Machines*, in *Fordham Law Review*, 87 (3), 2018, pp. 1089 ss.

SELBST A. - VENKATASUBRAMANIAN S. - I. KUMAR E., *Deconstructing design decisions: why courts must interrogate machine learning and other technologies*, in *85 Ohio State Law Journal*, 2024.

SILVESTRI G., *Consiglio superiore della magistratura e sistema costituzionale*, in *Questione Giustizia*, trimestrale n. 4 del 2017, pp. 19-29.

SIMONCINI A.- SUWEIS S., *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, in *Rivista di filosofia del diritto*, 1/2019.

SIMONCINI A., *Amministrazione digitale algoritmica. Il quadro costituzionale*, in CAVALLO PERIN R. - GALETTA D.U. (a cura di), *Il diritto dell'amministrazione pubblica digitale*, Torino 2020.

SIMONCINI A., *Il linguaggio dell'Intelligenza Artificiale e la tutela costituzionale dei diritti*, in *Rivista AIC*, 2023, II, 1-39.

SIMONCINI A., *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *Biolaw Journal – Rivista di BioDiritto*, 1/2019, pp. 63-89.

SIMONCINI A., *La dimensione costituzionale della giustizia predittiva. Riflessioni su intelligenza artificiale e processo*, in *Rivista di diritto processuale*, 2, 2024.

SIMONCINI A., *Profili costituzionali della amministrazione algoritmica*, in *Riv. Trim. Dir. Pubbl.*, 4, 2019, pp. 1149 ss.

SIMONCINI A., *Verso la regolamentazione della Intelligenza Artificiale. Dimensioni e governo*, in *BioLaw Journal – Rivista di BioDiritto*, n. 2/2021.

SLOBOGIN C., *Principles of Risk Assessment: Sentencing and Policing*, in *Ohio State Journal of Criminal Law*, vol. 15, 2018, pp. 583- 596.

SOURDIN T., *Judge v Robot? Artificial intelligence and judicial decision-making*, in *UNSW Law Journal*, 41, 2018, pp. 1128 ss.

SOURDIN T., *Judges, Technology and Artificial Intelligence. The Artificial Judge*, Cheltenham (UK), 2021.

SOURDIN T., *What if Judges Were Replaced by AI?*, in *Turkish Policy Quarterly*, 2022.

STRICKLAND E., *6 Reactions to the White House's AI Bill of Rights*, in *IEEE Spectrum*, 14 ottobre 2022.

SUSSKIND R.E., *Online Courts and the Future of Justice*, Oxford, 2019.

TARUFFO M., *Judicial Decisions and Artificial Intelligence*, in *Artificial Intelligence and Law*, 1998, pp. 311-324.

THORNE MC CARTY L., *Reflection on 'Taxman': An Experiment in Artificial Intelligence and Legal Reasoning*, in *Harvard Law Review*, 90, 5, 1977, 837-893.

TOOHEY L. - MOORE M. - DART K. - TOOHEY D. J., *Meeting the access to civil justice challenges: digital inclusion, algorithmic justice and Human-Centred Design*, in *19 Macquarie Law Journal*, 133, 2019.

TURING A.M., *Computing machinery and intelligence*, in *Mind*, 59, 1950, pp. 433-460.

TUZET G., *L'algoritmo come pastore del giudice? Diritto, tecnologie, prova scientifica*, in *Medialaws – Rivista dir. media*, 1/2020, pp. 45 ss.

TVERSKY A. - KAHNEMAN D., *Judgment under Uncertainty: Heuristics and Biases*, in *Science, New Series*, Vol. 185, No. 4157. 1974, pp. 1124-1131.

UBERTIS G., *Intelligenza artificiale, giustizia penale, controllo umano significativo*, in *Sistema penale*, 11 novembre 2020.

ULENAERS J., *The Impact of Artificial Intelligence on the Right to a Fair Trial: Towards a Robot Judge?*, in *Asian J Law Econ*, 11 (2) 2020.

VALLOR S., *The danger of Superhuman AI is not what you think*, in *Noema Magazine*, 23 maggio 2024.

VAN DIJCK G., *Predicting Recidivism Risk Meets AI Act*, in *European Journal on Criminal Policy and Research*, 28, 2022, pp. 407-423.

VAN DIJCK J., *Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology*, in *Surveillance & society*, XII, 2, 2014, pp.197-208.

VERNILE S., *Verso la decisione amministrativa algoritmica*, in *MediaLaws-Rivista di diritto dei media*, 2/2020.

VILLASENOR - V. FOGGO J., *Artificial intelligence, due process, and criminal sentencing*, in *Mich. St. L. Rev.*, 2020, pp. 295- 354.

VINCENTI E., Il «problema» del giudice-robot, in CARLEO A., *Decisione robotica*, Il Mulino, Bologna, 2019, p. 123.

WARWICK K., *Artificial intelligence - The basics*, trad. it. di BARATTIERI DI SAN PIETRO C. - MAUGERI G., Bologna, 2015.

WEAVER J.F., *Regulation of artificial intelligence in the United States*, in BARFIELD W. - PAGALLO U. (a cura di), *Research Handbook on the Law of Artificial Intelligence*, 2018, pp. 162-178.

WINNER L., *Do artifacts have politics?* *Daedalus*, 109 (1), 1980, pp. 121-136.

WON D., *The Missing Algorithm: Safeguarding Brady Against the Rise of Trade Secrecy in Policing*, in *Michigan Law Review*, 2021.

WÖRSDÖRFER M., *Biden's Executive Order on AI and the E.U.'s AI Act: A Comparative Computer-Ethical Analysis*, in *Philosophy & Technology*, 37(3), 2024.

YALCIN G. - THEMELI E. - STAMHUIS E. - PHILIPSEN S. - PUNTONI S., *Perceptions of Justice By Algorithms*, in *Artificial Intelligence and Law*, 31 (2), 2023, pp. 269-292.

YANG C.S., *Have Inter-Judge Sentencing Disparities Increased in an Advisory Guidelines Regime? Evidence from Booker*, in *SSRN Electronic Journal*, 2013.

ZALNIERUTE M. - BELL F., *Technology and the Judicial role*, in APPLEBY G. - LYNCH A., *The Judge, the Judiciary and the Court: Individual, Collegial and Institutional Judicial Dynamics in Australia*, Cambridge University Press, 2021, pp. 116-142.

ZANON N. - BIONDI F., *Il sistema costituzionale della magistratura*, Torino, 2024.

ZANON N. - PANZERI L., *Art 101*, in BIFULCO R. - CELOTTO A. - OLIVETTI M. (a cura di), *Commentario alla Costituzione*, Utet giuridica, Torino, 2006.

ZARSKY T. Z., *An Analytic Challenge: Discrimination Theory in the Age of Predictive Analytics*, in *A Journal of Law and Policy for the Information Society*, vol. 14 (1), 2017.

ZAVRŠNIK A., *Algorithmic justice: Algorithms and big data in criminal justice settings*, in *European Journal of Criminology*, 2021, Vol. 18 (5), pp. 623–642.

ZILLER J., *The Council of Europe Framework Convention on Artificial Intelligence vs. the EU Regulation: two quite different legal instruments*, in *CERIDAP*, 2/2024.

ZOLEA S., *The european Courts Faced with the unknowns of predictive Justice*, in DUMISTRESCU A. (a cura di), *Justice, Community and Freedom: IVR World Congress Proceedings*, 2022 pp. 436-445.

ZUBOFF S., *The Age of Surveillance Capitalism, The Fight for a Human Future at the New Frontier of Power* London, 2019.