



La giustizia algoritmica, tra esperienze comparate e la nuova regolamentazione europea in materia di intelligenza artificiale

di Andrea Giubilei*

Abstract: *The paper explores the legal and constitutional issues on the use of AI in the judicial sector. It analyses the risks of algorithms affecting court proceedings – not only judgements, but also lawyer's deeds – and examines the effects on compliance with the principle of transparency and guarantees of due process. The paper also focuses on the AI Act, which provides an advanced regulatory model based on principles of transparency, human oversight and risk management, which can also be replicated in non-European legal systems. In fact, in a comparative analysis, the paper examines the decision of Colombian Constitutional Court T-323/2024, on the use of AI in judicial decisions, which reaffirms the centrality of the human factor, looking at the AI Act as model for a national law on IA.*

SOMMARIO: 1. I rischi dell'uso "amatoriale" dell'intelligenza artificiale nel settore giudiziario – 2. Il peso delle valutazioni algoritmiche sugli atti processuali in alcune vicende italiane, europee ed internazionali – 3. L'*AI Act* e la regolamentazione dei sistemi di intelligenza artificiale nell'amministrazione della giustizia – 4. Il modello europeo come punto di riferimento: il caso della Colombia

1. I rischi dell'uso "amatoriale" dell'intelligenza artificiale nel settore giudiziario

Il progresso scientifico e tecnologico, soprattutto nelle sue manifestazioni più avanzate e sofisticate, ha reso più labile – se non, in alcuni casi, inesistente – quel confine tra *naturale* e *artificiale* che per secoli ha confortato l'umanità. Com'è

* Ricercatore di Diritto costituzionale e pubblico presso il Dipartimento di Scienze politiche dell'Università degli Studi Roma Tre. Contributo sottoposto a referaggio cieco (*double blind peer review*); il testo aggiorna e rielabora alcune considerazioni già sviluppate in S. BARBARESCI, A. GIUBILEI, *Il giudice chatbot: riflessioni sull'importanza delle garanzie italiane, prendendo spunto da una recente decisione della Corte costituzionale colombiana*, in «BioLaw Journal – Rivista di BioDiritto», n. 2, 2025, pp. 169-182.



Saggi

evidente, questa condizione pone alle società contemporanee una serie di sfide cruciali che interessano, in prima battuta, la tutela di quei diritti fondamentali che rischiano di essere compromessi da alcune delle tecnologie emergenti¹.

Per questo è chiamato in causa il diritto: d'altronde, pure il diritto rappresenta uno degli espedienti estranei all'ordine naturale delle cose – dunque “artificiali” – elaborati dall'uomo per governare la vita². Il diritto può divenire la grammatica di una coesistenza tra umano e artificiale che non sacrifichi i valori fondamentali in nome del progresso, ma li riaffermi come sua condizione irrinunciabile.

Tuttavia, queste possibilità sembrano oggi messe in crisi da quelle tecnologie che impattano sul diritto stesso e che compromettono la sua funzione di assicurare tutela agli individui secondo i canoni dell'equità, della razionalità e della certezza. Guardando ai tempi più recenti, queste criticità sembrerebbero caratterizzare, tra tutti, i sistemi di intelligenza artificiale (IA). Questa tecnologia, infatti, dirompe oggi in moltissimi ambiti di rilievo per l'ordinamento giuridico, oltre che per la vita delle persone: istruzione, industria, ricerca scientifica, salute, arte, informazione, ambiente, economia, giustizia.

Ecco, volendo concentrare l'attenzione sull'ultimo dei profili indicati, l'interazione tra IA e attività giurisdizionale rappresenta senza dubbio uno degli aspetti più problematici, per quanto affascinanti, dell'impatto delle nuove tecnologie sul diritto. Questo perché l'utilizzo dell'IA in ambito giudiziario, benché promettente sul piano dell'efficienza – ad esempio, nell'organizzazione degli uffici giudiziari, nell'analisi di precedenti, nel contenere la durata dei procedimenti – se non adeguatamente governato può incidere in maniera determinante sulla tenuta

¹ Sul punto, cfr. C. COLAPIETRO, *The impact of the information and technology revolution on constitutional rights with particular attention to personal data protection issues*, in M. BELOV (a cura di), *The Information and Technology Revolution and Its Impact on State, Constitutionalism and Constitutional Law*, Bloomsbury, London 2021, p. 231 ss.

² Possono riprendersi, al riguardo, le considerazioni di S. RODOTÀ, *Il diritto di avere diritti*, Laterza, Roma-Bari 2012, p. 352, secondo cui «di fronte all'interminata creazione di un mondo nuovo, i giuristi manifestano timori che sembrano confermare la tesi che li vuole espulsi dalla dimensione dove si scrive il palinsesto della vita. Tesi che, nella sua allarmante radicalità, finisce con il congiungersi con quella che al diritto assegna l'unico compito di legittimare tutto ciò che la scienza rende praticabile: ieri *instrumentum regni*, oggi, e soprattutto domani, mero *instrumentum scientiae*».



di quell'ideale di giustizia che viene declinato nei principi sul giusto processo e, conseguentemente, anche sulla tutela dei diritti in sede giurisdizionale³.

Eppure, allo stato attuale, si registra una certa impreparazione degli ordinamenti giuridici rispetto all'ingresso dell'IA nei contesti istituzionali. Ci si riferisce a tutte quelle ipotesi in cui l'IA influenza l'esercizio di funzioni pubbliche: nei Parlamenti, come supporto ai legislatori nella produzione normativa; nella pubblica amministrazione, quando i funzionari basano l'adozione di provvedimenti su valutazioni elaborate dagli algoritmi; nei Tribunali, determinando una netta trasformazione delle funzioni riservate a giudici e avvocati. La citata condizione di impreparazione rispetto a simili novità può ancora ritenersi fisiologica, giacché il progresso tecnologico tende sempre a precedere la produzione normativa e i decisori politici faticano a stare al passo con una realtà che si evolve troppo rapidamente⁴. Pertanto, anche l'utilizzo delle tecnologie algoritmiche da parte di chi amministra la cosa pubblica sembrerebbe ancora affidato, più che altro, al buon senso dei singoli. Il riferimento, qui, non è ad eventuali sistemi di IA regolarmente in dotazione degli uffici, bensì all'utilizzo che, al pari dei comuni cittadini, i singoli dipendenti pubblici possono fare dell'IA per svolgere le funzioni a cui sono deputati. In dottrina, è stato definito come un uso "amatoriale"⁵ o, comunque, non professionale dell'IA che, nell'ombra, sottotraccia, più o meno consapevolmente, condiziona istruttorie amministrative, provvedimenti o anche decisionali giudiziali.

Tuttavia, l'introduzione, anche inconsapevole, di strumenti di IA utilizzati in maniera amatoriale all'interno di contesti istituzionali comporta delle evidenti criticità. L'integrazione di soluzioni algoritmiche non validate, non supervisionate e prive di un'adeguata verifica tecnico-scientifica all'interno di processi

³ Sulla "tenuta" del principio costituzionale di tutela dei diritti a fronte di decisioni algoritmiche, cfr. A. SIMONCINI, *Il linguaggio dell'Intelligenza artificiale e la tutela costituzionale dei diritti*, in «Rivista AIC», n. 2, 2023, pp. 34 ss.

⁴ Sul punto si veda, fra tutti, A. IANNUZZI, *Il diritto capovolto. Regolazione a contenuto tecnico-scientifico e Costituzione*, Editoriale Scientifica, Napoli 2018.

⁵ In tal senso, v. E. LONGO, *La "giustizia digitale" tra nuove forme di efficienza e garanzia dei diritti fondamentali*, in «La nuova giuridica», n. 2, 2023, p. 187 ss., spec. 191 ss.



Saggi

decisionali istituzionali – ma anche più semplicemente professionali – può compromettere gravemente l'affidabilità degli esiti, minare i principi di legalità e trasparenza e determinare lesioni anche gravi a diritti soggettivi tutelati dall'ordinamento, quali il diritto alla riservatezza, alla non discriminazione, o – per quel che rileva ai fini della presente analisi – anche le garanzie del giusto processo. Le cause di simili criticità sono molteplici e possono rinvenirsi, principalmente, nell'opacità dei modelli utilizzati e nella riproduzione di *bias* e risultati distorti, generati da errori e pregiudizi (umani) che possono alterare i dati di addestramento dell'IA.

Il punto è che, in assenza di formazione e competenze specifiche, oltre che di un quadro regolatorio chiaro e condiviso, si rischia una proliferazione incontrollata di pratiche tecnologiche informali, spesso inidonee a rispettare i criteri minimi di sicurezza, equità e *accountability*.

Eppure, per rimanere al settore giudiziario, si registrano casi sempre più numerosi in cui gli stessi giudici fondano – più o meno dichiaratamente – le proprie decisioni sulle risposte dell'IA e, nello specifico, di *chatbot*⁶. Come precisato, quindi, non si tratta di *software* messi a disposizione degli uffici giudiziari, bensì di giudici che, al pari del *quisque de populo*, pongono domande ai più comuni sistemi di IA per trovare conferme in merito alla soluzione dei casi sottoposti al proprio giudizio o, comunque, per avere un supporto nella stesura della pronuncia.

Com'è noto, infatti, con «*chatbot*» si intende un *software* progettato per simulare una conversazione umana attraverso interfacce di chat, prevalentemente testuali, ma ormai in grado di esprimersi anche attraverso simulazioni vocali sempre più accurate. I *chatbot* possono essere più semplici o più complessi, a seconda

⁶ Nel nostro ordinamento, già da qualche anno si discute, a livello dottrinario, sulla legittimità o comunque sull'opportunità di decisioni giudiziarie robotiche. Fra tutti, si vedano, senza pretesa di esaustività, F. DONATI, *Intelligenza artificiale e giustizia*, in «Rivista AIC», n. 1, 2020, p. 414 ss.; M. FASAN, *Intelligenza artificiale e costituzionalismo contemporaneo. Principi, diritti e modelli in prospettiva comparata*, Editoriale Scientifica, Napoli 2024, p. 237 ss.; E. LONGO, *Giustizia digitale e Costituzione. Riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, FrancoAngeli, Milano 2023; M. LUCIANI, *La decisione giudiziaria robotica*, in «Rivista AIC», n. 3, 2018, p. 872 ss.



Saggi

delle tecnologie utilizzate. Quelli più avanzati, ma che si stanno anche diffondendo rapidamente, sono quelli basati sull'IA. Più nello specifico, i *chatbot* che pongono più criticità in merito alla loro applicabilità all'interno di sedi istituzionali sono quelli basati su *Large Language Models (LLM)*, perché, grazie al *machine learning*, possono autonomamente migliorare le loro prestazioni nel tempo e fornire risposte sempre più accurate e personalizzate. L'esempio più noto e diffuso sul mercato è *ChatGPT*, sviluppato dalla società statunitense *OpenAI*, che è anche al centro di alcune vicende che verranno analizzate nel presente contributo.

Si tratta di un modello di IA progettato per comprendere, generare e interagire con il linguaggio umano, che si basa su enormi quantità di dati per elaborare e produrre testo in base agli *input* ricevuti. Grazie alla loro struttura, simili strumenti sono capaci di eseguire una vasta gamma di compiti linguistici, come traduzione, completamento del testo, generazione creativa e risposte a domande complesse.

È legittimo chiedersi, allora, se tali strumenti siano funzionali per rispondere, in maniera più efficace dell'essere umano, anche a domande di giustizia. Non vi è dubbio, infatti, che in ambito giudiziario le determinazioni umane potrebbero risultare astrattamente «viziate dagli stati d'animo, dalle emozioni, dalle amicizie, dagli interessi particolari di ognuno», escludendo la possibilità di «esprimere alcuna scelta realmente razionale poiché i (falsi) motivi logici per raggiungerla sarebbero creati artificiosamente in un secondo tempo, dopo che, per ragioni in realtà poco o nulla sensate, avremmo già deciso il senso da dare alla nostra volontà»⁷.

Non vi è dubbio che, a prima vista, il processo logico che caratterizza la risposta dei *chatbot* e quella dei giudici sembrerebbe sovrapponibile: a) l'intelligenza artificiale, a fronte di una domanda posta dall'utente, sceglie con cura fra gli elementi a sua disposizione (i *dataset* inseriti in fase di addestramento) e fornisce una risposta secondo criteri dati dall'algoritmo; b) il giudice, a fronte di una domanda di

⁷ Cfr. C. CASONATO, *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in «Diritto pubblico comparato ed europeo», n. speciale, 2019, p. 119, che, sul punto, rinvia a J. NATHAN MATHIAS, *Bias and Noise: Daniel Kahneman on Errors in Decision-Making*, in *Medium*, 17 ottobre 2017 (<https://natematias.medium.com/bias-and-noise-daniel-kahneman-on-errors-in-decision-making-6bc844ff5194>). Su simili criticità, v. anche A. SANTOSUOSSO, *Intelligenza artificiale e diritto. Perché le tecnologie di IA sono una grande opportunità per il diritto*, Mondadori università, Milano 2020, p. 93 ss.



giustizia del cittadino, sceglie con cura fra gli elementi a sua disposizione (l'apparato normativo e, entro certi limiti, giurisprudenziale) e fornisce una risposta secondo criteri dati dal diritto, in base al principio di soggezione del giudice soltanto alla legge.

Tuttavia, tali considerazioni necessitano di un approccio costituzionalmente orientato in merito all'opportunità che la decisione umana venga sostituita da quella artificiale⁸. Difatti, come in parte già anticipato, sono molteplici le frizioni con alcuni principi fondamentali sul giusto processo che possono venire in rilievo.

Una prima criticità concerne l'opacità dei processi decisionali elaborati dall'IA (la c.d. *black box*). I moderni modelli di *machine learning* elaborano previsioni a partire da correlazioni statistiche spesso non intelleggibili nemmeno per i loro stessi sviluppatori. Trasposta in sede processuale, tale opacità genera il rischio che decisioni potenzialmente incidenti sulla libertà o, comunque, sulla vita dei soggetti coinvolti, vengano pronunciate senza un controllo effettivo sui criteri logici e fattuali che hanno guidato l'algoritmo. Come si vedrà anche più avanti, ne deriva una tensione con il principio di motivazione e con il diritto alla comprensione delle ragioni del provvedimento.

Strettamente connesso è il problema dei *bias* discriminatori. Gli algoritmi "apprendono" dai dati storici, che possono riflettere disuguaglianze pregresse (ad esempio, di natura razziale, di genere o socio-economica). In assenza di un'accurata procedura di *de-biasing*, tali pregiudizi vengono perpetuati e potenziati dall'automazione⁹, con conseguente violazione del principio di uguaglianza.

⁸ Cfr. S. PENASA, *Intelligenza artificiale e giustizia: il delicato equilibrio tra affidabilità tecnologica e sostenibilità costituzionale in prospettiva comparata*, in «DPCE Online», n. 1, 2022, p. 298, secondo cui, rispetto all'interazione tra strumenti tecnologici e funzioni pubbliche, «si pone l'esigenza di assicurarne la sostenibilità costituzionale, la quale si compone tanto di una dimensione assiologica, relativa all'impatto sui diritti e sulle libertà delle persone coinvolte, quanto di una dimensione propriamente istituzionale, la quale richiede una valutazione degli effetti – fisiologici o potenzialmente distorsivi – dell'impiego di sistemi autonomi sulla natura della funzione giurisdizionale e sul ruolo del giudice all'interno di uno stato costituzionale di diritto».

⁹ Sul punto, A. IANNUZZI, *La definizione di Intelligenza Artificiale fra Regolamento europeo e norma tecnica internazionale. L'uso umano degli esseri umani e l'uso tecnico delle macchine*, in F. PIZZETTI (a



La questione della responsabilità costituisce un ulteriore snodo problematico. Nell'ecosistema dell'IA, si moltiplicano i livelli di intervento (sviluppatori, fornitori di dati, magistrati utilizzatori), rendendo ardua l'attribuzione univoca della decisione. Una scelta eccessivamente piegata sulla soluzione offerta dall'IA rischia, così, di alimentare una deresponsabilizzazione diffusa, incompatibile con il principio personalistico della responsabilità che permea le tradizioni giuridiche occidentali.

Infine, l'autonomia del giudice viene messa in crisi da modelli che tendono a standardizzare l'esito delle controversie. L'utilizzo di strumenti che uniformano o suggeriscono soluzioni mediane può intaccare l'essenza stessa del giudizio come atto umano, responsabile e argomentato, trasformandolo in una procedura meccanica governata dal dato statistico, ove si disperdono le specificità fattuali e le circostanze soggettive che caratterizzano il caso concreto.

Alla luce di tali rilievi, l'adozione dell'IA in sede giudiziaria impone un approccio precauzionale: occorrono robusti meccanismi di verificabilità e controllo degli algoritmi, protocolli di validazione indipendente, procedure di *human-in-the-loop* che preservino il vaglio critico del magistrato e un chiaro regime di *accountability*.

Ed è in questa direzione che, come si vedrà, sembrerebbero collocarsi alcune recenti decisioni adottate sul piano nazionale, europeo ed internazionale, in cui, anche laddove manchi una disciplina sulle condizioni di legittimità per l'uso dell'IA da parte dei giudici, le Corti superiori tentano di fare ordine.

cura di), *La regolazione europea dell'Intelligenza artificiale nella società digitale*, Giappichelli, Torino, 2025, p. 28 ss., osserva, infatti, che «la discriminazione prodotta dalla macchina non è altro che la pericolosa riproduzione delle discriminazioni umane. E ciò è evidente, con tutta la sua carica di amarezza, se si pone mente al fatto che i soggetti che sono stati già discriminati dall'AI o che potenzialmente sono esposti al rischio di discriminazioni appartengono alle medesime categorie che subiscono già oggi trattamenti ingiusti. È l'uomo che discrimina mentre il problema della discriminazione algoritmica è determinato dalla sua capacità di replicare e di amplificare le ingiustizie e le nefandezze che si verificano nel mondo e che si trasferiscono alla macchina attraverso i dati».



2. Il peso delle valutazioni algoritmiche sugli atti processuali in alcune vicende italiane, europee ed internazionali

L'IA è entrata da tempo nelle aule giudiziarie, ancor prima di trovare una compiuta regolazione da parte del legislatore. Ma, al momento, vi ha fatto ingresso principalmente come *oggetto* della decisione giudiziale, cioè come sistema dal cui utilizzo è scaturita la controversia.

Ciò è avvenuto anche in Italia, come dimostra la nota vicenda relativa all'utilizzo di procedure automatizzate nei concorsi per la scuola, che ha consentito al giudice amministrativo di affermare alcuni fondamentali principi sull'assunzione di decisioni algoritmiche da parte della pubblica amministrazione¹⁰. Pur non trattandosi ancora di IA in senso stretto, senza dubbio questo filone giurisprudenziale ha gettato delle solide basi per l'individuazione dei limiti sull'utilizzo delle nuove tecnologie nell'assunzione di decisioni pubbliche, con un possibile impatto sui diritti e, più in generale, sulla vita dei cittadini.

In quell'occasione, la vicenda era sorta dall'iniziativa assunta dal Ministero dell'istruzione, dell'Università e della ricerca (MIUR) di adottare un sistema informatico automatizzato – un algoritmo – come strumento per la gestione e la riorganizzazione delle assegnazioni del personale docente sul territorio nazionale¹¹. La gestione informatizzata della procedura ha dato adito a significative

¹⁰ Sul punto, v. F. LAVIOLA, *Algoritmico, troppo algoritmico: decisioni amministrative automatizzate, protezione dei dati personali e tutela delle libertà dei cittadini alla luce della più recente giurisprudenza amministrativa*, in «BioLaw Journal», n. 3, 2020, p. 389 ss. e L. MUSELLI, *La decisione amministrativa nell'età degli algoritmi: primi spunti*, in «MediaLaws», n. 1, 2020, p. 18 ss., spec. 25 ss.

¹¹ Tale decisione si collocava nel più ampio contesto della riforma introdotta dalla legge 13 luglio 2015, n. 107 – conosciuta anche come “Buona scuola” – recante disposizioni in materia di riorganizzazione del sistema nazionale di istruzione e formazione e volta a rafforzare l'efficienza del servizio scolastico attraverso un piano straordinario di immissione in ruolo di personale docente. La normativa in questione aveva previsto, oltre a una fase ordinaria di reclutamento, un secondo momento attuativo avente carattere straordinario, finalizzato, da un lato, alla copertura dei posti comuni e di sostegno rimasti vacanti all'esito della fase ordinaria, e dall'altro, all'assegnazione di ulteriori cattedre destinate sia al potenziamento dell'offerta formativa sia alla riduzione del ricorso a supplenze temporanee, soprattutto nelle scuole di primo e secondo grado. Tale ulteriore fase procedurale si sarebbe



Saggi

contestazioni da parte di alcuni insegnanti, che lamentavano l'assegnazione a posizioni professionali non corrispondenti al proprio percorso di esperienza e specializzazione. Al centro delle contestazioni, vi era proprio la gestione automatizzata delle assegnazioni, che non consentiva di conoscere in modo trasparente – per l'inaccessibilità al funzionamento interno dell'algoritmo – i criteri applicati per garantire un'adeguata considerazione delle preferenze manifestate dai candidati.

Come anticipato, la vicenda è giunta dinnanzi al giudice amministrativo, che ha avuto modo di pronunciarsi sulla questione in plurime occasioni¹². Le considerazioni sviluppate nelle motivazioni di tali decisioni, pur riferendosi più strettamente al procedimento amministrativo, possono rilevare trasversalmente per ogni momento applicativo del diritto in cui sia coinvolto un algoritmo e, dunque, anche per la sua applicazione giudiziale. Si può considerare, in tal senso, la sentenza del Consiglio di Stato, sez. VI, 13 dicembre 2019, n. 8472, con la quale i giudici di Palazzo Spada hanno messo in luce l'indubbio apporto che alcune innovazioni tecnologiche sofisticate possono fornire in termini di efficienza, evidenziando come «in molti campi gli algoritmi promettono di diventare lo strumento attraverso il quale correggere le storture e le imperfezioni che caratterizzano tipicamente i processi cognitivi e le scelte compiute dagli esseri umani»¹³.

Allo stesso tempo, però, il Consiglio di Stato ha chiarito come l'impiego di strumenti algoritmici comporti, in realtà, «una serie di scelte e di assunzioni tutt'altro che neutre: l'adozione di modelli predittivi e di criteri in base ai quali i dati sono raccolti, selezionati, sistematizzati, ordinati e messi insieme, la loro interpretazione e la conseguente formulazione di giudizi sono tutte operazioni frutto di precise scelte e di valori, consapevoli o inconsapevoli; da ciò ne consegue

svolta secondo modalità derogatorie rispetto ai tradizionali meccanismi di assegnazione, prevedendo – per l'appunto – l'impiego di un sistema algoritmico per la dislocazione dei docenti a livello nazionale.

¹² Possono considerarsi, in tal senso, due coppie di pronunce di primo e secondo grado, rispettivamente: TAR Lazio, sez. III-*bis*, 1° dicembre 2016, n. 12026 e Consiglio di Stato, sez. VI, 8 aprile 2019, n. 2270; TAR Lazio, sez. III-*bis*, 12 settembre 2018, n. 9230 e Consiglio di Stato, sez. VI, 13 dicembre 2019, n. 8472.

¹³ Consiglio di Stato, sez. VI, 13 dicembre 2019, n. 8472, par. 7.1.



che tali strumenti sono chiamati ad operano una serie di scelte, le quali dipendono in gran parte dai criteri utilizzati e dai dati di riferimento utilizzati, in merito ai quali è apparso spesso difficile ottenere la necessaria trasparenza»¹⁴.

Quest'ultimo, d'altronde, sembrerebbe il profilo più critico. Perché, anche in questo caso, come accade per ogni decisione adottata da chi detiene il potere di incidere sulla sfera giuridica dei cittadini, si pone l'esigenza di verificare «che i criteri, i presupposti e gli esiti del procedimento robotizzato siano conformi alle prescrizioni e alle finalità stabilite dalla legge o dalla stessa amministrazione a monte di tale procedimento e affinché siano chiare – e conseguentemente sindacabili – le modalità e le regole in base alle quali esso è stato impostato»¹⁵.

Sulla base di tali premesse, il Consiglio di Stato individua due elementi «di minima garanzia per ogni ipotesi di utilizzo di algoritmi in sede decisoria pubblica: a) la piena conoscibilità a monte del modulo utilizzato e dei criteri applicati; b) l'imputabilità della decisione all'organo titolare del potere, il quale deve poter svolgere la necessaria verifica di logicità e legittimità della scelta e degli esiti affidati all'algoritmo»¹⁶.

Notare che, pur occupandosi di un episodio avente ad oggetto un procedimento amministrativo, i giudici di Palazzo Spada sembrerebbe riferirsi generalmente all'uso di algoritmi per *ogni* decisione assunta nell'esercizio di una funzione pubblica, tra cui può rientrare, a ben vedere, anche quella di natura

¹⁴ Consiglio di Stato, sez. VI, 13 dicembre 2019, n. 8472, par. 7.2.

¹⁵ Ivi, par. 13.1. Tali considerazioni sembrerebbero riferibili all'esistenza di un vero e proprio *right to explanation*, elaborato dalla dottrina anzitutto a partire dalle garanzie conoscitive previste dalla normativa in materia di privacy e approntate, in particolare, dal GDPR per quanto attiene il divieto di processi decisionali automatizzati (art. 22). Sul punto, cfr. C. COLAPIETRO, *Circolazione dei dati, automatizzazione e regolazione*, in «Osservatorio sulle fonti», n. 2, 2021, p. 831 ss.

¹⁶ Consiglio di Stato, sez. VI, 13 dicembre 2019, n. 8472, par. 12. Sul punto, v. C. BENETAZZO, *IA, giustizia e P.A.: la sfida etica ed antropologica*, in «Federalismi.it», n. 8, 2025, p. 17, secondo cui, pur nella consapevolezza dell'apporto che gli sviluppi tecnologici possono offrire al buon andamento della pubblica amministrazione (art. 97 Cost.), deve considerarsi che «sul piano della vita, della dignità umana e del diritto di eguaglianza l'individualismo normativo richiede che l'intervento dell'uomo nel compimento della scelta amministrativa non sia del tutto obliterato o marginalizzato, proprio per rispettare la gerarchia assiologica voluta dalla Costituzione e alzare un argine alla "disumanizzazione" dell'amministrazione, riportando la macchina a semplice strumento dell'azione amministrativa».



giurisdizionale. Infatti, come si vedrà meglio in seguito, i due elementi sopra indicati rappresentano i principali nodi critici delle prime sentenze in cui viene dato conto dell'utilizzo dell'IA da parte dei giudici. In particolare, al fine di misurare la legittimità delle decisioni *latu sensu* "automatizzate", il fattore davvero determinante sembrerebbe essere quello individuato dal Consiglio di Stato al punto *sub b)*, cioè l'imputabilità della decisione all'organo – in questo caso – giudicante e non alla macchina.

La vicenda sopra analizzata non è che l'inevitabile conseguenza dell'avvento inatteso – quantomeno in termini così dirompenti – degli algoritmi nella vita delle persone. Perché è evidente che la loro diffusione capillare nei rapporti sociali, ancora in assenza di una sistematica regolazione da parte del legislatore, pone in capo ai giudici il gravoso compito di individuare presupposti e condizioni per un utilizzo delle nuove tecnologie rispettoso dei diritti e delle libertà fondamentali coinvolte. Anche all'interno di altri ordinamenti iniziano a registrarsi vicende analoghe, ove l'adozione di alcune decisioni automatizzate o basate su sistemi di IA è giunta all'interno delle aule giudiziarie, anche se ancora come mero oggetto della controversia e, dunque, della decisione giudiziale.

Quel che è interessante, notare, però, è come gradualmente l'IA si stia spostando dall'essere la causa del contendere all'essere la soluzione della contesa. In altri termini, a leggere alcune sentenze adottate in Italia e nel resto del mondo, si scorge sempre più frequentemente il peso dell'IA generativa e delle valutazioni algoritmiche sul contenuto degli atti processuali.

Tra i primi casi può considerarsi quello avvenuto nel luglio del 2024 nei Paesi Bassi, ove il *Rechtbank Gelderland* – uno dei tribunali di primo grado del paese – ha dichiaratamente fatto uso di *ChatGPT* nell'ambito di una controversia civile inerente a rapporti di vicinato e all'utilizzo dei pannelli solari¹⁷. A ben vedere, l'apporto del *chatbot*, in questo caso, sembrerebbe essere paragonabile a quello di

¹⁷ Trattasi della decisione ECLI:NL:RBGEL:2024:3636, pubblicata il 26 luglio 2024, e consultabile sul sito istituzionale del *Rechtbank Gelderland* al seguente collegamento:

<https://uitspraken.rechtspraak.nl/details?id=ECLI:NL:RBGEL:2024:3636&showbutton=true&keyword=ECLI%253aNL%253aRBGEL%253a2024%253a3636&idx=1>.



Saggi

un “consulente tecnico” del giudice¹⁸. Quest’ultimo, infatti, si è servito del noto sistema di IA per avere informazioni particolarmente specialistiche legate all’utilizzo dei pannelli solari. In particolare, la controversia era sorta perché uno dei due vicini coinvolti nella vicenda lamentava un calo nella produzione di energia dei propri pannelli solari, imputabile – a suo dire – alla costruzione realizzata sul tetto del convenuto, che ne avrebbe causato l’ombreggiamento.

La questione centrale era quindi la stima del danno economico derivante dalla presunta perdita di rendimento fotovoltaico. A questo fine, il giudice olandese ha deciso di consultare direttamente *ChatGPT* per ottenere alcune informazioni ausiliarie sulla vita utile media dei pannelli solari e sul prezzo medio dell’elettricità nel periodo considerato (euro per kWh). Il giudice, infatti, chiarisce di aver individuato i dati necessari per assumere la decisione «mede met behulp van *ChatGPT*»¹⁹, cioè “anche con l’aiuto di *ChatGPT*”; frase che ritorna – in maniera pressoché identica – in un altro passaggio della sentenza, concernente la quantificazione del risarcimento.

A ben vedere, quindi, le risposte fornite dall’IA sono state incorporate nella motivazione della sentenza per stimare il danno, pur con la cautela (data da quel «anche con l’aiuto») di precisare che le informazioni utilizzate non sono state decisive per l’assunzione della decisione finale sulla controversia.

Ciò nonostante, vi sono vari aspetti che lasciano perplessi. Anzitutto, quello di una fiducia ingiustificatamente cieca, da parte del giudice, sull’affidabilità della soluzione individuata: com’è noto, i modelli generativi come *ChatGPT* non si basano su dati necessariamente aggiornati o verificati in tempo reale, ma su informazioni pregresse di natura probabilistica. La possibilità che l’*output* generato sia impreciso, manipolato o addirittura “allucinato” – cioè completamente

¹⁸ Eppure, è bene ricordare, come evidenzia E. LONGO, *La “giustizia digitale” tra nuove forme di efficienza e garanzia dei diritti fondamentali*, cit., p. 192, che «laddove venga utilizzato per decidere o per selezionare casi, l’IA non può essere equiparata a strumenti convenzionali, non essendo assimilabile né a un perito né a un consulente tecnico»; ragione per cui è necessario che «l’uso delle tecnologie nelle attività processuali sia accompagnato da una seria riflessione sulle implicazioni che le tecnologie digitali hanno per i diritti fondamentali tutelati da e attraverso i processi».

¹⁹ *Rechtbank Gelderland*, ECLI:NL:RBGEL:2024:3636, par. 5.7.



Saggi

inventato – rende problematico un suo uso diretto come fonte conoscitiva in ambito processuale, soprattutto in assenza di validazione umana e confronto tra le parti. Inoltre, anche tale vicenda solleva evidenti dubbi in termini di trasparenza del processo decisionale. Nella sentenza, infatti, non viene pubblicato il *prompt* utilizzato dal giudice, né vengono date spiegazioni sul controllo che ha esercitato sull'*output*. Il risultato è quello di un utilizzo autonomo ed amatoriale del *chatbot*, senza che le parti del processo ne fossero informate preventivamente, in modo da consentire un contraddittorio sulle valutazioni sviluppate dall'IA.

Il caso olandese appare dunque emblematico rispetto ad una tendenza che si va diffondendo in diversi ordinamenti: il ricorso a strumenti di IA generativa non come supporto alla macchina della giustizia, ma come elemento che, più o meno consapevolmente, si rivela determinante nella costruzione dell'argomentazione giudiziale. Si tratta di una dinamica che impone riflessioni profonde sulla natura del provvedimento giurisdizionale, sulla sua intelligibilità e sulla responsabilità del giudice rispetto alle fonti da cui attinge.

E, da quest'ultimo punto di vista, deve prendersi in considerazione una tendenza, apparentemente meno problematica, ma che condivide alcuni degli elementi critici considerati fino ad ora: quella dell'utilizzo sconsiderato dell'IA generativa come supporto difensivo da parte degli avvocati.

Già nel 2023, negli Stati Uniti ha fatto particolarmente discutere un caso sottoposto alla Corte federale del *Southern District of New York* che, pur dovendosi inizialmente occupare di tutt'altra vicenda²⁰, è finita per pronunciarsi sull'uso maldestro di *ChatGPT* ad opera degli avvocati di una delle parti in causa. In particolare, gli avvocati Steven A. Schwartz e Peter LoDuca hanno depositato presso la Corte una memoria difensiva contenente – come di consueto – il riferimento a svariati precedenti giurisprudenziali²¹. Tuttavia, come emerso, poi, durante processo, nel tentativo di velocizzare l'attività di ricerca giurisprudenziale,

²⁰ Trattasi dell'azione civile 22-cv-01461-PKC, intentata nel febbraio del 2022 contro la compagnia aerea colombiana *Avianca* da uno dei passeggeri che, durante un volo di linea, era stato ferito alla gamba da un carrello di servizio.

²¹ Il documento è consultabile al seguente collegamento:
<https://storage.courtlistener.com/recap/gov.uscourts.nysd.575368/gov.uscourts.nysd.575368.21.0.pdf>.



gli avvocati hanno fatto ricorso a *ChatGPT* per redigere l'atto difensivo. Il risultato è stato quello di includere, a supporto delle proprie argomentazioni, sei precedenti giudiziari completamente inventati, frutto di un'allucinazione dell'IA utilizzata, senza una verifica, anche minima, sull'esistenza o sulla fondatezza delle informazioni riportate²².

E, infatti, a seguito delle sollecitazioni della controparte e delle richieste dello stesso giudice, è emersa l'effettiva inesistenza di alcune delle decisioni citate, poi confermata dagli stessi legali interessati. In particolare, nella dichiarazione formalmente resa alla Corte federale dall'avvocato Steven A. Schwartz, emerge come egli abbia utilizzato *ChatGPT* «to assist with legal research», effettuando la ricerca «in the same general manner I searched any legal research database»²³.

Questa specifica questione, che ha interessato incidentalmente il processo, si è conclusa con una sanzione di 5.000 dollari nei confronti degli avvocati coinvolti e del loro studio, a cui è stato anche ordinato di inviare lettere di rettifica a tutti i giudici le cui decisioni erano state falsamente attribuite.

Interessante notare le considerazioni sviluppate dal giudice distrettuale nel proprio ordine sanzionatorio, secondo cui «technological advances are

²² Il dialogo tra il *chatbot* di *OpneAI* e l'avvocato Steven A. Schwartz, ove è stato dato conto della (errata) esistenza dei precedenti giurisprudenziali in questione, è stato depositato come prova documentale ed è consultabile al seguente collegamento: <https://storage.courtlistener.com/recap/gov.uscourts.nysd.575368/gov.uscourts.nysd.575368.46.1.pdf>. In particolare, i precedenti poi rivelatisi inesistenti sarebbero: *Varghese v. China Southern Airlines Co Ltd*; *Shaboon v. Egyptair*; *Petersen v. Iran Air*; *Martinez v. Delta Airlines, Inc.*; *Estate of Durden v. KLM Royal Dutch Airlines*; *Miller v. United Airlines, Inc.*

²³ Tali affermazioni sono riportate nell'atto depositato il 6 giugno 2023 e contenute, per l'appunto, le dichiarazioni formalmente rese alla Corte, consultabile al seguente collegamento: <https://storage.courtlistener.com/recap/gov.uscourts.nysd.575368/gov.uscourts.nysd.575368.46.0.pdf>. La confusione di cui è stato vittima l'avvocato coinvolto, che ha visto in un *chatbot* basato su LLM le medesime funzionalità che poteva individuare su un database di natura legale, è un ulteriore sintomo dei rischi che si celano nel calare un uso ciecamente fiducioso dell'IA nell'ambito di un contesto professionale. D'altronde, sono ormai numerosi i sistemi di IA appositamente elaborati per supportare giudici e avvocati nella loro attività, che potrebbero risultare senz'altro più mirati nel fornire supporto agli avvocati. Sul punto, anche per un'esauritiva individuazione dei sistemi in questione, cfr. L. S'AVACK, *Intelligenza artificiale e diritto: problematiche etiche e giuridiche*, in «Dir. fam. pers.», n. 4, 2023, p. 1714 ss.



Saggi

commonplace and there is nothing inherently improper about using a reliable artificial intelligence tool for assistance. But existing rules impose a gatekeeping role on attorneys to ensure the accuracy of their filings»²⁴. A ben vedere, quindi, la sanzione non derivava dall'uso dell'IA in sé, ma dalla negligenza professionale e dalla mancanza di un'adeguata verifica umana, che ha dato luogo ad una violazione dei doveri deontologici di diligenza e di correttezza verso la Corte.

Il caso Schwartz evidenzia, in modo emblematico, i limiti attuali dell'IA generativa utilizzata in maniera amatoriale nell'ambito di un contesto professionale, specie in quello giuridico. Tali strumenti, senza un adeguato – anche se, verrebbe da dire, “minimo” – livello di preparazione e competenza sul loro utilizzo, non sono ancora in grado di sostituire l'essere umano anche nello svolgimento di attività semplici e ripetitive come la selezione delle fonti nell'ambito di una ricerca giurisprudenziale. Il tutto lasciando ancora senza risposta i consueti interrogativi sulla responsabilità per gli errori generati dall'algoritmo, sulla trasparenza nell'uso delle nuove tecnologie in ambito professionale e sull'opportunità di adottare un apparato regolatorio e presidi etici – anche che si tratti solo di linee guida uniformi in ogni ordinamento giudiziario – sulle condizioni di utilizzo dell'IA negli atti processuali.

Eppure, nonostante le evidenti criticità sollevate dal caso Schwartz, una vicenda analoga non ha tardato a giungere anche nel nostro ordinamento. E, infatti, un'interessante – e più recente – occasione per riflettere sulle implicazioni dell'impiego dell'IA negli atti processuali è offerta dall'ordinanza del Tribunale di Firenze del 14 marzo 2025, che si è pronunciato in merito all'utilizzo di un sistema di IA generativa (nella specie, ancora una volta *ChatGPT*) da parte di un difensore nel corso di un giudizio civile.

²⁴ L'ordine sanzionatorio adottato dal giudice distrettuale degli Stati Uniti P. Kevin Castel, in data 22 giugno 2023, è consultabile al seguente collegamento: https://storage.courtlistener.com/recap/gov.uscourts.nysd.575368/gov.uscourts.nysd.575368.54.0_8.pdf. Nella decisione emerge come il giudice abbia motivato la sanzione sulla base del fatto che i legali «abandoned their responsibilities when they submitted non-existent judicial opinions with fake quotes and citations created by the artificial intelligence tool ChatGPT, then continued to stand by the fake opinions after judicial orders called their existence into question».



Il caso ha assunto rilievo per ragioni analoghe alla vicenda statunitense sopra richiamata e, cioè, l'indicazione, all'interno della comparsa di costituzione della parte resistente, di alcune pronunce della Corte di cassazione che, all'esito di verifiche sollecitate dalla controparte, si sono rivelate del tutto inesistenti o non pertinenti rispetto al contenuto attribuito.

La questione è emersa in sede di reclamo avverso un'ordinanza cautelare, allorché la parte reclamante ha invocato l'applicazione dell'art. 96 c.p.c. nei confronti dell'avvocato di controparte, ipotizzando una responsabilità aggravata per lite temeraria²⁵. Il Tribunale, nel motivare la reiezione della domanda, ha ricostruito in modo accurato la dinamica dell'accaduto: il difensore della resistente ha chiarito che i riferimenti giurisprudenziali contestati erano frutto dell'attività di una collaboratrice di studio, la quale aveva effettuato una ricerca tramite *ChatGPT*, ottenendo risultati apparentemente plausibili ma in realtà del tutto infondati. Si è trattato, secondo quanto ricostruito nell'ordinanza, di un classico «fenomeno delle cc.dd. allucinazioni di intelligenza artificiale, che si verifica allorché l'IA inventi risultati inesistenti ma che, anche a seguito di una seconda interrogazione, vengono confermati come veritieri»²⁶.

L'ordinanza, pur stigmatizzando l'omessa verifica della veridicità dei riferimenti inseriti in atti ufficiali, ha ritenuto insussistenti i presupposti per una condanna ai sensi dell'art. 96 c.p.c. In particolare, si è sottolineato come l'indicazione delle pronunce fittizie non fosse stata volta a manipolare il giudizio, né ad introdurre una nuova linea difensiva, bensì a rafforzare argomentazioni già espresse e già valutate favorevolmente dai giudici di primo grado. In tal senso, il Tribunale ha osservato che l'indicazione di tali riferimenti giurisprudenziali è stata posta a fondamento della tesi sostenuta sin dal principio dalla parte resistente, «proposta quindi a supporto di una struttura difensiva rimasta immutata sin dal primo grado del giudizio ed oggettivamente non finalizzata ad influenzare il collegio,

²⁵ Al riguardo, si ricorda che l'art. 96 c.p.c. stabilisce, al primo comma, che «se risulta che la parte soccombente ha agito o resistito in giudizio con mala fede o colpa grave, il giudice, su istanza dell'altra parte, la condanna, oltre che alle spese, al risarcimento dei danni, che liquida, anche d'ufficio, nella sentenza».

²⁶ Così Trib. Firenze, sez. imprese, (ord. 14 marzo 2025, punto n. 3).



appuntandosi piuttosto su quanto già indicato, in senso analogo, anche nelle decisioni di prime cure»²⁷.

Sulla base di tali considerazioni, il Tribunale ha escluso l'applicabilità dell'art. 96 c.p.c., perché, «fermo restando il disvalore relativo all'omessa verifica dell'effettiva esistenza delle sentenze risultanti dall'interrogazione dell'IA, [...] l'indicazione di estremi di legittimità nel giudizio di reclamo ad ulteriore conferma della linea difensiva già esposta [...] si può quindi considerare diretta a rafforzare un apparato difensivo già noto e non invece finalizzata a resistere in giudizio in mala fede, conseguendone la non applicabilità delle disposizioni di cui all'art. 96 c.p.c.»²⁸.

Anche in questo caso, si può ritenere che l'ordinanza del Tribunale di Firenze rappresenti un precedente significativo non solo per l'approccio sostanzialmente garantista con cui è stata trattata la questione, ma anche per la chiarezza con cui viene delineata la necessità, per gli operatori del diritto, di maneggiare con cautela strumenti tecnologici avanzati, le cui potenzialità generative devono essere accompagnate da un adeguato controllo umano, palesando la necessità, ormai non più differibile, di una riflessione sistematica su *standard* minimi di correttezza e verificabilità nell'uso delle tecnologie nei procedimenti giudiziari.

3. L'*AI Act* e la regolamentazione dei sistemi di intelligenza artificiale nell'amministrazione della giustizia

Dall'analisi delle vicende sopra riportate, si evince come gli ordinamenti e, nello specifico, i settori giudiziari, si siano trovati alquanto impreparati rispetto all'avvento dell'IA. Il punto è che si tratta di una tecnologia che presenta degli innegabili vantaggi astratti per l'esercizio della funzione giurisdizionale. Può considerarsi, in primo luogo, il buon andamento dell'amministrazione della giustizia (art. 97 Cost.), in termini di maggiore efficienza, economicità e rapidità

²⁷ *Ibidem*.

²⁸ *Ibidem*.



Saggi

nell'adozione delle decisioni, componenti che rilevano sia come interesse collettivo al funzionamento dell'apparato giudiziario, sia come diritto del soggetto coinvolto nella singola decisione (può considerarsi, tra tutti, il diritto alla durata ragionevole del processo, di cui all'art. 111, comma 2 Cost.). Inoltre – ma, lo ripetiamo, solo sul piano astratto – un secondo ordine di vantaggi potrebbe individuarsi nella maggiore imparzialità nell'adozione della decisione (art. 111, comma 2 Cost.) e nella garanzia della certezza giuridica. Si pensi, in particolare, alla possibilità che l'IA faciliti il compito del giudice circoscrivendo, in maniera meno fallibile di un essere umano, l'ambito applicativo della normativa rilevante per il caso di specie, ad esempio individuando la normativa vigente al netto di un'avvenuta abrogazione o superando altre forme di antinomia.

Non solo. Com'è evidente, l'uso di *chatbot* sempre più sofisticati possiede anche un'evidente aura attrattiva per i singoli operatori del diritto – giudici o avvocati – che, al pari di chiunque, vedono nell'IA una pratica soluzione a portata di mano cui potersi affidare per lo svolgimento dei propri compiti istituzionali²⁹, fiduciosi (forse troppo ciecamente) nella qualità e nella correttezza degli *output*. È quella che è stata descritta come «la travolgente forza pratica dell'algoritmo», per cui «una volta introdotto un sistema automatico di decisione all'interno di un processo decisionale umano, il sistema automatico tende, nel tempo, a catturare la decisione stessa»³⁰.

Tuttavia, i casi sopra analizzati, dimostrano come l'IA in ambito giudiziario debba essere maneggiata con cautela e che, senza dubbio, presenta degli innegabili fattori di *rischio* per i diritti dei soggetti coinvolti.

²⁹ Cfr. C. CASONATO, *Costituzione e intelligenza artificiale: un'agenda per il prossimo futuro*, in AA.VV., *Liber amicorum per Pasquale Costanzo*, vol. I, Consulta online, Genova 2020, p. 386, secondo cui, rispetto a tali attività, «l'ipotesi più convincente [...] riguarda non la completa sostituzione dell'uomo con la AI, ma un'attività di assistenza: una collaborazione in termini di sussidiarietà, in cui ogni parte svolge i compiti che maggiormente le si addicono. Per quanto riguarda l'amministrazione della Giustizia, così, si sono sottolineati i possibili vantaggi del supporto algoritmico in termini, ad esempio, di rapidità e ampiezza dell'istruttoria; di completezza degli elementi su cui costruire la decisione; di certezza (o esattezza) del diritto».

³⁰ Così A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in «BioLaw Journal», n. 1, 2019, p. 63 ss., spec. 81.



Ebbene, com'è noto, è proprio attorno all'elemento del "rischio" che ruota la nuova disciplina eurolunitaria sull'IA e che caratterizza, nello specifico, le previsioni sull'applicazione di tali tecnologie in ambito giudiziario. Questo approccio ha accompagnato il cammino della regolazione europea dell'IA sin da tempi meno sospetti³¹ e, da ultimo, è stato fatto proprio dal regolamento (UE) 2024/1689, «che stabilisce regole armonizzate sull'intelligenza artificiale» (c.d. *AI Act*). Si tratta di un atto normativo che segna una svolta epocale nel governo delle tecnologie digitali, anche perché è la prima normativa al mondo che affronta in modo sistematico e giuridicamente vincolante l'uso dell'IA nei diversi ambiti della vita sociale, economica e pubblica³². Al riguardo, sin dal principio, il nuovo regolamento chiarisce esplicitamente che il proprio obiettivo è quello di «promuovere la diffusione di un'intelligenza artificiale (IA) antropocentrica e affidabile, garantendo nel contempo un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'Unione europea, compresi la democrazia, lo Stato di diritto e la protezione

³¹ Per una ricostruzione dell'impegno delle Istituzioni europee nella regolazione dell'IA, v. A. SANTOSUOSSO, M. TOMASI, *Diritto, scienza, nuove tecnologie*, Wolters Kluwer, 2021, p. 314 ss., ove gli Autori sottolineano come le proposte della Commissione europea siano sempre state guidate dal concetto di "rischio", seguendo due criteri cumulativi: da una parte, guardando all'insorgenza del rischio sulla base del settore specifico in cui l'IA è impiegata (come potrebbe essere quello sanitario o quello bancario); dall'altra parte, tenendo in considerazione le specifiche modalità con cui l'IA viene utilizzata, anche se il settore non rientra fra quelli considerati rischiosi (pensiamo, tra tutti, all'uso di sistemi di riconoscimento facciale).

³² Il nuovo regolamento europeo sull'IA sembrerebbe rientrare in quel percorso che, per quanto riguarda il governo della società digitale, si caratterizza per «uno spostamento del baricentro della produzione legislativa, in questi ambiti, dall'autoregolazione privata al regolamento europeo». In tal senso, A. IANNUZZI, *Le fonti del diritto dell'Unione europea per la disciplina della società digitale*, in F. PIZZETTI (a cura di), *La regolazione europea della società digitale*, Giappichelli, Torino, 2024, p. 13 ss., secondo cui «la strada imboccata dall'Europa con questa decisa operazione di recupero della centralità delle fonti del diritto è quella giusta perché può consentire di affrontare correttamente il tema della sovranità e del potere nell'era digitale che rappresenta uno dei problemi più rilevanti del costituzionalismo del XXI secolo». Sul punto, per quanto riguarda più nello specifico i temi trattati nel presente contributo, v. anche M. FASAN, *L'intelligenza artificiale nel settore della giustizia. Prime riflessioni alla luce della proposta di Regolamento (UE) in materia di IA*, in «Queste istituzioni», n. 4, 2022.



dell'ambiente, proteggere contro gli effetti nocivi dei sistemi di IA nell'Unione, nonché promuovere l'innovazione» (art. 1)³³.

Per perseguire simili obiettivi, come già anticipato, la nuova disciplina intende «stabilire un meccanismo autorizzativo circa l'immissione nel mercato dei “prodotti” che incorporano sistemi di IA [...] seguendo un approccio basato sul rischio, secondo la regola: più alto è il rischio, più severe sono le regole»³⁴. A ben vedere, quindi, il legislatore europeo non tenta di circoscrivere i sistemi di IA sulla base delle loro caratteristiche tecniche, ma si concentra, invece, su *come* essi vengono utilizzati e su quali rischi possono generare in concreto per i diritti degli utenti³⁵.

³³ Sul punto, cfr. M. OROFINO, *Obiettivi, ambito di applicazione e principi fondamentali dell'AI Act*, in F. PIZZETTI (a cura di), *La regolazione europea dell'Intelligenza artificiale nella società digitale*, cit., p. 35 ss., che si concentra, in particolare, sulla necessità che l'IA sia antropocentrica ed affidabile, individuando: nell'“antropocentrismo” la condizione secondo cui l'IA «dovrebbe essere progettata e sviluppata tenendo in considerazione l'essere umano come punto di riferimento. Non però come entità monodimensionale, ma come persona collocata in un ambiente naturale, sociale e culturale»; nell'“affidabilità”, «la capacità di un sistema, dispositivo o processo di svolgere le sue funzioni in modo accurato e coerente». Distingue tra prospettiva “antropomorfa” e prospettiva “antropocentrica” C. PINELLI, *L'AI Act: gestione del rischio e tutela dei diritti*, in «Giur. it.», n. 2, 2025, p. 453, secondo cui «la prima proietta la ragione e più in generale l'identità umana su macchine pensanti, l'altra mira a far sì che le innovazioni scientifiche e tecniche avvengano nel rispetto delle garanzie dei diritti fondamentali. Il timore di una dipendenza da qualcosa di incontrollabile può infatti nascere più da un senso di spossessamento dell'identità che dalla preoccupazione per la perdita dei diritti. E tale timore si pone nel momento in cui l'identità ha comunque già conosciuto una trasformazione decisiva con un'acquisizione sempre più massiccia di dati personali da parte degli algoritmi e comunque di fonti esterne».

³⁴ In tal senso, E. LONGO, *Le pratiche di IA vietate e i sistemi di IA ad alto rischio: metodi strumenti per la società del “rischio digitale”*, in F. PIZZETTI (a cura di), *La regolazione europea dell'Intelligenza artificiale nella società digitale*, cit., p. 64.

³⁵ D'altronde, come osserva A. IANNUZZI, *La definizione di Intelligenza Artificiale fra Regolamento europeo e norma tecnica internazionale. L'uso umano degli esseri umani e l'uso tecnico delle macchine*, cit., p. 26 ss., «diventa difficile qualificare l'Intelligenza artificiale perché già di per sé è difficile definire l'intelligenza umana, che è evidente se solo si pensa al fatto che c'è chi ha autorevolmente negato che l'intelligenza umana sia riconducibile ad un concetto unitario. Di conseguenza, l'accostamento di una tecnologia dirompente ad un concetto liquido, anzi gassoso, come l'intelligenza umana non può che produrre un esito definitorio (quasi) sempre vago». Per tale ragione, prosegue l'Autore, «la scelta del legislatore europeo è andata nel senso di definire, e quindi circoscrivere, il fenomeno dell'Intelligenza



Per questo, il Regolamento classifica i sistemi di IA in quattro categorie sulla base dell'intensità del rischio generato dal loro utilizzo, per ciascuna delle quali è previsto un diverso livello di regolamentazione: 1) rischio inaccettabile³⁶; 2) alto rischio³⁷; 3) rischio limitato³⁸; 4) rischio minimo o nullo³⁹.

In tale quadro, l'utilizzo dell'IA per l'amministrazione della giustizia viene espressamente qualificato come attività ad "alto rischio". Più in particolare, il punto n. 8 dell'Allegato III del regolamento⁴⁰ include in tale ambito «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». Si tratta di una formulazione (volutamente) ampia, idonea a ricomprendere qualsiasi sistema capace di influenzare, anche in modo non vincolante, l'elaborazione della decisione giudiziale.

Artificiale intorno alle sue potenziali finalità ed applicazioni, costituendo questi i principali aspetti in cui può rinvenirsi un effettivo pregiudizio per i diritti fondamentali degli individui».

³⁶ Qui rientrano tutti quei sistemi che l'UE considera incompatibili con i valori fondamentali: pensiamo, ad esempio, a tecnologie che manipolano il comportamento delle persone sfruttandone vulnerabilità psicologiche, a sistemi di riconoscimento biometrico in tempo reale usati in spazi pubblici, o ai meccanismi di *social scoring*, in cui il comportamento dei cittadini viene valutato per assegnare diritti o comminare sanzioni. Questi sistemi sono vietati in modo assoluto.

³⁷ Si tratta di tecnologie che non sono vietate, ma che possono avere un impatto molto significativo sui diritti delle persone. Per esempio: *software* usati nella selezione del personale, algoritmi per il credito bancario, sistemi di IA utilizzati nella sanità o – come si vedrà – nel settore della giustizia.

³⁸ Si tratta di tecnologie che non pongono rischi particolarmente gravi, ma che devono comunque rispettare alcune regole di trasparenza. È il caso, per esempio, dei *chatbot* o degli assistenti virtuali: l'utente deve essere chiaramente informato che sta interagendo con una macchina e non con un essere umano. Lo stesso vale per i sistemi che generano contenuti (come testi o immagini), rispetto ai quali andrebbe indicato in modo visibile che si tratta di *output* prodotti artificialmente.

³⁹ Vi rientrano sistemi quali quelli usati nei videogiochi, nei filtri *antispam* delle *e-mail*, o nei sistemi di raccomandazione personalizzata per lo *shopping*. Per questi casi, l'uso è libero e non soggetto a particolari vincoli, salvo il rispetto delle normative generali (come potrebbe essere, ad esempio, la disciplina sulla protezione dei dati personali).

⁴⁰ L'Allegato III è dedicato ai sistemi di IA ad alto rischio e viene in rilievo sulla base di un espresso rinvio disposto dall'art. 6, par. 2, del regolamento.



Le ragioni dell'inclusione tra i sistemi ad alto rischio vengono meglio esplicitate al *considerando* n. 61, in base al quale «alcuni sistemi di IA destinati all'amministrazione della giustizia e ai processi democratici dovrebbero essere classificati come sistemi 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»⁴¹. Per tali ragioni, l'utilizzo di sistemi di IA nell'amministrazione della giustizia implica l'imposizione di una serie particolarmente stringente di obblighi in capo ai soggetti che sviluppano, forniscono o impiegano tali sistemi.

Anzitutto, si stabilisce – in via generale – che i sistemi di IA ad alto rischio rispettino i requisiti disposti dal regolamento stesso, «tenendo conto delle loro previste finalità nonché dello stato dell'arte generalmente riconosciuto in materia di IA e di tecnologie correlate all'IA» (art. 8). Tale disposizione può ritenersi integrata da quanto previsto dal *considerando* n. 46 dello stesso regolamento, secondo cui tali requisiti dovrebbero garantire che i sistemi di IA «non presentino rischi inaccettabili per interessi pubblici importanti dell'Unione, come riconosciuti e tutelati dal diritto dell'Unione».

⁴¹ A integrazione di tali previsioni, può considerarsi anche quanto disposto dal *considerando* n. 59, in base al quale «tenuto conto del loro ruolo e della loro responsabilità, le azioni delle autorità di contrasto che prevedono determinati usi dei sistemi di IA sono caratterizzate da un livello significativo di squilibrio di potere e possono portare alla sorveglianza, all'arresto o alla privazione della libertà di una persona fisica, come pure avere altri impatti negativi sui diritti fondamentali garantiti nella Carta. In particolare, il sistema di IA, se non è addestrato con dati di elevata qualità, se non soddisfa requisiti adeguati in termini di prestazione, accuratezza o robustezza, o se non è adeguatamente progettato e sottoposto a prova prima di essere immesso sul mercato o altrimenti messo in servizio, può individuare le persone in modo discriminatorio o altrimenti errato o ingiusto. 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. È pertanto opportuno classificare come ad alto rischio, nella misura in cui il loro uso è consentito dal pertinente diritto dell'Unione e nazionale, una serie di sistemi di IA destinati a essere utilizzati nel contesto delle attività di contrasto, in cui l'accuratezza, l'affidabilità e la trasparenza risultano particolarmente importanti per evitare impatti negativi, mantenere la fiducia dei cittadini e garantire la responsabilità e mezzi di ricorso efficaci».



Inoltre, nel quadro delle garanzie previste per l'utilizzo di IA nel settore della giustizia, un aspetto centrale riguarda la qualità e la *governance* dei dati impiegati per lo sviluppo di tali sistemi⁴². Il legislatore europeo, consapevole delle implicazioni che un uso non controllato dell'IA può determinare in settori sensibili (come, appunto, in ambito giudiziario), ha previsto un'articolata disciplina finalizzata ad evitare che *bias* strutturali, errori sistematici o difetti informativi possano riverberarsi negativamente sui diritti delle persone o sull'imparzialità del processo decisionale.

In particolare, l'art. 10 del regolamento stabilisce che i sistemi ad alto rischio che utilizzano tecniche che prevedono l'uso di dati per l'addestramento di modelli di IA sono sviluppati sulla base di *set* di dati di addestramento, convalida e prova che rispettino elevati standard qualitativi. Tali *set* devono essere sottoposti ad adeguate pratiche di *governance* e gestione dei dati, che includano, tra gli altri: la possibilità di tracciare i dati all'origine; lo svolgimento di operazioni di aggiornamento e aggregazione; l'adozione di misure specifiche per valutare «le possibili distorsioni suscettibili di incidere sulla salute e sulla sicurezza delle persone, di avere un impatto negativo sui diritti fondamentali o di comportare discriminazioni vietate dal diritto dell'Unione, specie laddove gli *output* di dati influenzano gli *input* per operazioni future» (art. 10, par. 2, lett. *f*). Il regolamento precisa, inoltre, che i dati utilizzati debbano essere «pertinenti, sufficientemente rappresentativi e, nella misura del possibile, esenti da errori e completi», nonché capaci di riflettere con accuratezza il contesto specifico – geografico, comportamentale, funzionale – in cui il sistema sarà applicato (art. 10, parr. 3 e 4).

Di rilevante interesse è, infine, la possibilità – prevista in via eccezionale – di trattare categorie particolari di dati personali, al solo fine di rilevare e correggere distorsioni nei modelli di IA. Tale trattamento è ammesso solo ove non possa essere realizzato efficacemente mediante l'utilizzo di altri dati, compresi i dati sintetici o anonimizzati ed è subordinato al rispetto di una serie di garanzie rafforzate (misure tecniche di pseudonimizzazione e sicurezza, accesso ristretto ai

⁴² Sull'importanza della qualità dei dati impiegati nell'ambito delle decisioni algoritmiche, cfr. C. COLAPIETRO, *Circolazione dei dati, automatizzazione e regolazione*, cit., p. 839 ss.



soli soggetti autorizzati vincolati da obblighi di riservatezza, divieto assoluto di trasferimento a terzi e obbligo di cancellazione al termine della finalità perseguita).

Nel *corpus* normativo dell'*AI Act*, un ulteriore passaggio di cruciale rilevanza per l'impiego dell'IA in ambito giudiziario è senza dubbio rappresentato dalle disposizioni concernenti la trasparenza dei sistemi (art. 13), che si pone come condizione imprescindibile per garantire un uso consapevole, responsabile e giuridicamente controllabile delle tecnologie algoritmiche nei contesti in cui sono in gioco diritti fondamentali o decisioni di natura autoritativa, come accade in ambito processuale.

Il legislatore europeo impone, anzitutto, che i sistemi di IA ad alto rischio siano progettati e sviluppati «in modo tale da garantire che il loro funzionamento sia sufficientemente trasparente da consentire ai *deployer* di interpretare l'*output* del sistema e utilizzarlo adeguatamente». La finalità dichiarata è quella di garantire «un tipo e un livello di trasparenza adeguati», funzionali non solo alla piena comprensibilità dell'*output*, ma anche al rispetto degli obblighi gravanti sia sui fornitori che sugli stessi *deployer* (termine che indica gli utilizzatori professionali del sistema, come potrebbero essere – per quanto ci interessa – giudici o ausiliari del giudice).

In tale ottica, il regolamento stabilisce che ogni sistema di IA ad alto rischio debba essere accompagnato da istruzioni d'uso «concise, complete, corrette e chiare», accessibili e comprensibili da parte degli utilizzatori finali. A ben vedere, queste istruzioni non potranno ridursi a meri supporti tecnici oscuri e indecifrabili, ma dovranno rappresentare veri e propri strumenti di garanzia, contenenti informazioni essenziali per esercitare – come si vedrà – un controllo umano sull'IA, così da scongiurare il rischio di affidamento cieco e inconsapevole nei confronti della tecnologia.

Tra gli elementi informativi obbligatori, particolare rilievo assumono: la descrizione della finalità prevista del sistema; le indicazioni relative al livello atteso di accuratezza, robustezza e cbersicurezza; l'esplicitazione delle circostanze note o prevedibili che potrebbero compromettere la sicurezza o i diritti fondamentali delle persone; le capacità esplicative del sistema rispetto ai propri *output*; le



Saggi

eventuali specificità prestazionali in relazione a determinati gruppi di persone. Ulteriori obblighi informativi riguardano le specifiche tecniche sui dati di *input* e sui set utilizzati per l'addestramento, convalida e test e la descrizione delle misure di sorveglianza umana, comprensive di soluzioni tecniche atte a facilitare l'interpretazione degli *output* dei sistemi di IA da parte degli utilizzatori.

Nel settore della giustizia, tali previsioni risultano particolarmente rilevanti. La trasparenza e l'esplicabilità degli strumenti algoritmici, infatti, si pongono come condizione necessaria per assicurare il rispetto dell'obbligo di motivazione delle decisioni giudiziarie e, più in generale, la comprensibilità di ogni atto di natura giurisdizionale di cui si è destinatari, anche al fine di esercitare i diritti garantiti in sede processuale (tra tutti quello del contraddittorio tra le parti)⁴³. Difatti, un *output* algoritmico generato secondo processi non decifrabili dall'operatore umano non consente, a valle, al destinatario del provvedimento di verificare il percorso logico-argomentativo che ha condotto il giudice all'adozione della decisione; il che rappresenterebbe un'evidente violazione di tutti i presidi costituzionali a garanzia del giusto processo.

Da ultimo, tra le previsioni di maggior rilievo per quanto riguarda l'impiego di tecnologie algoritmiche in ambito giudiziario, vi è senza dubbio l'articolato regime dedicato alla sorveglianza umana dei sistemi di IA. L'art. 14 del regolamento individua, in proposito, una serie di obblighi stringenti che si pongono come garanzia essenziale per assicurare il rispetto dei diritti fondamentali, l'affidabilità funzionale dei sistemi e, più in generale, la conformità dell'impiego dell'IA ai principi dello Stato di diritto.

⁴³ Cfr. G. MOBILIO, *I giudici e le nuove tecnologie per giudicare: una occasione per riscoprire i caratteri fondanti della funzione giurisdizionale*, in V. MESSERINI, R. ROMBOLI, E. ROSSI, A. SPERTI, R. TARCHI (a cura di), *Ricordando Alessandro Pizzorusso. L'ordinamento giudiziario*, Pisa University Press, Pisa 2021, p. 483 ss., secondo cui la tendenza all'opacità degli algoritmi «può spiegarsi con una pluralità di motivi, come la tutela della sicurezza del sistema stesso, la garanzia di segreti commerciali, ma anche per rendere indecifrabili le decisioni algoritmiche e l'uso dei dati. A ogni modo, tale circostanza non pare affatto compatibile con i citati riferimenti normativi che informano il contributo decisorio del giudice, il quale non potrebbe fondare la propria decisione sul responso imperscrutabile di un algoritmo, ovvero una sequenza di istruzioni di cui non si conosce la struttura, il peso accordato alle variabili considerate e i dati processati».



In primo luogo, la norma stabilisce che i sistemi di IA ad alto rischio debbano essere progettati e sviluppati in modo da garantire, anche attraverso interfacce uomo-macchina adeguate, una supervisione efficace da parte di persone fisiche durante l'intero periodo di utilizzo. Non si tratta di un requisito meramente formale, ma di una esigenza sostanziale: la sorveglianza umana deve essere tale da prevenire o ridurre al minimo i rischi per la salute, la sicurezza o i diritti fondamentali che possano emergere dall'utilizzo ordinario o dall'uso improprio – ma ragionevolmente prevedibile – del sistema⁴⁴. In questo senso, la supervisione si configura come un presidio residuale di legalità, chiamato a intervenire laddove le misure tecniche di sicurezza e conformità *ex ante* risultino insufficienti o non esaustive.

Le modalità della sorveglianza umana devono essere proporzionate al rischio connesso al sistema, al grado di autonomia di quest'ultimo e al contesto di utilizzo. A tal fine, l'art. 14, par. 3, del regolamento stabilisce che le misure di sorveglianza siano garantite mediante almeno uno dei seguenti strumenti: «a) misure individuate e integrate nel sistema di IA ad alto rischio dal fornitore prima della sua immissione sul mercato o messa in servizio, ove tecnicamente possibile; b) misure individuate dal fornitore prima dell'immissione sul mercato o della messa in servizio del sistema di IA ad alto rischio, adatte ad essere attuate dal *deployer*».

Ciò implica che i sistemi di IA ad alto rischio debbano essere forniti assieme strumenti tecnici e documentali tali da consentire al soggetto umano deputato alla sorveglianza – ad esempio, nel contesto giudiziario, il magistrato o un

⁴⁴ Sul punto, cfr. A. PUNZI, *La decisione giudiziaria nell'AI Act*, in «Giur. it.», n. 2, 2025, p. 450, che evidenzia come «il principio già affermato nell'art. 22 del GDPR trova ora un'opportuna specificazione: non è solo vietata “una decisione basata unicamente sul trattamento automatizzato”, bensì prescritta un'efficace supervisione nel corso dell'attività svolta dal sistema di AI, dunque un controllo su tutti i passaggi rilevanti ai fini della decisione finale, che deve sempre essere assunta dall'uomo. Tale previsione esprime una visione antropocentrica aperta e, anche grazie al riferimento ad un'interfaccia adeguata uomo-macchina, suggerisce l'idea che il sistema di AI non fornisca solo un supporto in sede di ricognizione dei contributi dottrinali e giurisprudenziali rilevanti sul *thema decidendum*, bensì interagisca attivamente nel processo decisionale del giudicante, sempre rimanendo da questo sorvegliato».



Saggi

ausiliario del giudice – di comprendere le capacità e i limiti del sistema, monitorarne il funzionamento e intervenire in caso di anomalie, disfunzioni o risultati inattesi.

Di particolare rilievo è l'esplicito riferimento alla necessità che i *deployer* – da intendersi, quindi come gli esseri umani incaricati della sorveglianza – restino consapevoli del rischio di *automation bias*, ovvero della tendenza a fare automaticamente affidamento o eccessivo affidamento sull'*output* fornito dal sistema. Si tratta dell'importante riconoscimento normativo di un fenomeno psicologico e cognitivo e che – se non adeguatamente contenuto – rischia di trasformare la supervisione umana in una sorveglianza meramente formale, se non illusoria.

Il regolamento prevede, inoltre, che l'operatore umano debba essere posto nelle condizioni di: interpretare correttamente l'*output* algoritmico; decidere di non utilizzare il sistema o, altrimenti di ignorare, annullare o ribaltare l'*output* generato; interrompere in sicurezza il funzionamento del sistema, mediante un apposito comando di arresto.

Nel complesso, l'art. 14 dell'*AI Act* si propone come il cardine di una strategia di responsabilizzazione del fattore umano nel governo dell'IA. La sorveglianza umana, infatti, viene elevata a prerequisito strutturale di legittimità per l'uso dei sistemi ad alto rischio, specialmente nei settori – come quello giudiziario – in cui l'algoritmizzazione della decisione incide direttamente sull'integrità di un processo decisionale pubblico, su posizioni giuridiche soggettive, sulla libertà personale e sulla tutela dei diritti fondamentali⁴⁵.

⁴⁵ In questo senso, l'*AI Act* sembrerebbe fare un passo in avanti rispetto alla garanzia del principio di non esclusività della decisione algoritmica che, in passato, trovava una (blanda) tutela nella previsione del già citato art. 22 del regolamento generale sulla protezione dei dati (Reg. UE 2016/679), in base al quale «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». Sul punto, cfr. A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, cit., p. 69, che evidenziava – prima dell'entrata in vigore dell'*AI Act* – come il principio in questione, pur apparentemente risolutivo, auspicando quella collaborazione “intelligenza naturale-intelligenza artificiale”, fosse intrinsecamente debole, dal momento che, in alcuni casi, una decisione algoritmica poteva legittimamente incidere sui diritti della persona anche



Quanto analizzato fino ad ora consente di inquadrare la via europea all'IA come una sintesi tra innovazione, responsabilizzazione e tutela della dignità umana. L'ottica antropocentrica, che vuole un'IA quale «strumento per le persone, con il fine ultimo di migliorare il benessere degli esseri umani» (*considerando* n. 6), intende creare le condizioni affinché le logiche dell'efficienza tecnologica non prevalgano sui diritti fondamentali delle persone. Questo vale in particolar modo per l'uso di sistemi algoritmici nell'amministrazione della giustizia, ove l'impianto del regolamento (UE) 2024/1689 consente di immaginare un'IA al servizio del giudice e non in sua sostituzione. Ciò dà la conferma di quanto l'UE tenda alla regolazione del progresso, del fenomeno tecnologico e del digitale, con un'ispirazione di fondo senz'altro positiva, cioè quella di predisporre gli strumenti per assicurare che l'innovazione tecnologica manifesti le sue potenzialità in termini di crescita, progresso e miglioramento della condizione umana e, al contempo, evitare che ciò determini il consolidamento di posizioni di privilegio, diseguaglianze e asimmetrie di potere all'interno della società⁴⁶.

4. Il modello europeo come punto di riferimento: il caso della Colombia

Nel contesto globale, caratterizzato da una crescente diffusione e pervasività dei sistemi di IA, l'adozione dell'*AI Act* rappresenta senz'altro un punto di svolta normativo destinato a suscitare riflessioni e potenziali emulazioni anche al di là dei confini dell'Unione europea. Difatti, come visto nel paragrafo precedente, tale regolamento si distingue per un impianto regolatorio ispirato ai principi di precauzione, proporzionalità e tutela dei diritti fondamentali, fondando l'intero

senza intervento umano (ad esempio, col consenso esplicito dell'interessato, o quando necessaria per concludere o eseguire un contratto tra interessato e titolare del trattamento, oppure se autorizzata dal diritto dell'Unione o dello stato membro). A ben vedere, con il nuovo regolamento europeo sull'IA, il principio di non esclusività della decisione algoritmica sembrerebbe aver trovato ulteriori forme di garanzia, dal momento che, allo stato, la decisione di un giudice che si avvalga di sistemi di IA sembrerebbe oggi ammissibile solo a condizione che vi sia una costante supervisione umana.

⁴⁶ Cfr. G. RESTA, *Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di eguaglianza*, in «Pol. dir.», n. 2, 2019, p. 199 ss.



assetto su un approccio basato sul rischio e sull'*accountability* dei soggetti coinvolti nel ciclo di vita dell'IA.

Questa architettura normativa si configura come una risposta pionieristica alle sfide etiche, giuridiche e sociali sollevate dalle tecnologie algoritmiche, che potrà senz'altro rappresentare un modello normativo, suscettibile di ispirare la regolazione dell'IA anche in ordinamenti extra-europei. Sono, in particolare, la centralità del controllo umano nei processi decisionali automatizzati e l'attenzione per le disegualianze algoritmiche a costituire quei tratti distintivi che potrebbero guidare la riflessione di altri Paesi nella costruzione di strumenti normativi di analoga portata.

Da questo punto di vista, in tempi relativamente recenti, vi è stata una decisione adottata da un giudice colombiano che può essere presa come fondamento per sviluppare alcune considerazioni conclusive. Da un lato, tale sentenza rientra tra quelle prime ipotesi in cui l'IA diviene *soggetto* – non più mero *oggetto* – della decisione giudiziale, dal momento che l'utilizzo del *chatbot* da parte del giudice si è posto sul labile confine tra l'acquisizione di informazioni e dati su taluni aspetti del giudizio (con funzione di *supporto* alla decisione) e la ricerca di una opportuna soluzione della controversia (con funzione *sostitutiva* nell'adozione della decisione). Dall'altro lato, tale vicenda ha mostrato la sensibilità che altri ordinamenti possiedono rispetto alla disciplina europea in tema di IA, individuata quale punto di riferimento.

Come anticipato, tale caso emblematico si è registrato in Colombia, nel gennaio del 2023, quando un giudice è stato chiamato a pronunciarsi in secondo grado su una vicenda in tema di diritti sociali, che riguardava la copertura delle spese mediche e la fornitura del servizio di trasporto per un bambino con diagnosi di disturbo dello spettro autistico, in favore del quale è stata poi riconosciuta l'esenzione per il trasporto finalizzato alle terapie. Per giungere a tale conclusione, il giudice ha posto a *ChatGPT* una serie di domande relative agli aspetti più critici della propria controversia, le cui risposte sono poi confluite espressamente nel testo della sentenza⁴⁷.

⁴⁷ Trattasi della decisione del *Juzgado 1º laboral del Circuito Cartagena*, sent. 30 gennaio 2023, n. 32. Per la versione originale della sentenza, v. <https://forogpp.com/wp-content/uploads/2023/01/sentencia-tutela-segunda-instancia-rad.-13001410500420220045901.pdf>.



Va chiarito da subito che il giudice non sembrerebbe aver interpellato il *chatbot* per giungere alla soluzione del quesito giuridico sottoposto al suo scrutinio, bensì a conferma di conclusioni e considerazioni già svolte dal giudice. Nel caso di specie dopo aver ricostruito il fatto, richiamato la normativa di riferimento e la giurisprudenza rilevante in materia, nonché espresso la propria volontà di confermare l'esito del giudizio di primo grado, l'estensore del provvedimento ha riportato la sua "conversazione" con il *chatbot*, affermando espressamente di averla utilizzata a supporto della propria decisione finale⁴⁸.

In un certo senso, quindi, le richieste a *ChatGPT* sono state formulate per ottenere una conferma della bontà della decisione assunta, attribuendogli più la veste di "consulente-robot" che di vero e proprio "giudice-robot" e, dunque, non con un intento *sostitutivo* nell'assunzione della decisione.

Difatti, come è stato evidenziato dai primi commentatori, «le risposte del software di IA, in realtà molto generiche, aderiscono alle argomentazioni sviluppate dal magistrato nella parte precedente della sentenza, la cui articolazione corrisponde alla sua visione circa l'utilità dell'impiego di tali tecnologie nel processo decisionale, che è quella di "aggiungere argomentazioni" una volta presa autonomamente la decisione al fine di "agevolare la risoluzioni dei casi di tutela"»⁴⁹.

⁴⁸ Tra le domande, figurano alcune che sembrerebbero proprio mirate a trovare la soluzione del quesito giuridico posto al giudice come, ad esempio: «¿Menor autista esta exonerado de pagar cuotas moderadoras en sus terapias?», a cui il *chatbot* ha risposto nel seguente modo: «sí, es correcto. De acuerdo con la normativa en Colombia, los menores con diagnóstico de autismo están exonerados de pagar cuotas moderadoras en sus terapias. Este beneficuento está regulado por la Ley 1753 de 2015 y se aplica a todas las instituciones prestadoras de servicios de salud públicas o privadas que reciban recursos del sistema general de seguridad social en salud»; o ancora «¿Exigir en estos casos la cuota moderadora es una barrera de acceso al servicio de salud?», a cui il *chatbot* ha risposto «sí, en algunos casos, la exigencia de pagar cuotas moderadoras puede ser una barrera de acceso al servicio de salud para personas con diagnóstico de autismo, especialmente si tienen dificultades para obtener ingresos o si su situación económica es precaria. La exigencia de pagar cuotas moderadoras puede impedir o dificultar el acceso a tratamientos y terapias esenciales para el bienestar y desarrollo de los niños con autismo. Por esta razón, la normativa en Colombia exime a los menores con diagnóstico de autismo de pagar cuotas moderadoras en sus terapias, con el objetivo de garantizar el acceso a estos servicios de salud esenciales sin obstáculos financieros».

⁴⁹ Così E. LONGO, *Giustizia digitale e Costituzione. Riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, cit., p. 15.



Nonostante le migliori intenzioni, il caso colombiano sopra richiamato, forse frutto di una diffusa euforia mondiale per le IA, sembra nascondere la mancanza di una ponderata e corretta analisi circa le forme di impiego delle tecnologie nel processo e del loro rapporto con le garanzie costituzionali che presiedono allo svolgimento delle funzioni giurisdizionali.

In questo senso, risulta particolarmente significativa la successiva decisione della *Corte Constitucional de Colombia* (*sentencia* T-323, adottata il 2 agosto 2024)⁵⁰, arrivata proprio a valle della vicenda giudiziaria sopra indicata e che la Corte ha ritenuto rilevante riesaminare per un'eventuale violazione del diritto fondamentale al giusto processo nell'emissione della decisione⁵¹. Con tale decisione, i giudici costituzionali colombiani hanno messo in luce le criticità illustrate sino ad ora, stabilendo un interessante catalogo dei limiti all'uso dell'IA in ambito giudiziario.

Preliminarmente, deve osservarsi quanto si è anticipato in apertura del presente paragrafo e cioè che, nella propria decisione, la Corte colombiana ha mostrato una particolare attenzione ai principi del diritto europeo e, più nello specifico, quelli enucleati nell'ambito dell'*AI Act*. Infatti, ai parr. nn. 182 e 183 della propria decisione, la Corte costituzionale colombiana evidenzia proprio come il quadro normativo dell'IA sia ancora in fase di elaborazione in America latina, per cui al netto di alcune eccezioni – come quella peruviana⁵² – nessuno dei paesi dispone di una legge speciale in vigore che regoli l'uso e lo sviluppo dell'IA. E, per l'appunto, la Corte aggiunge che, in questa fase di elaborazione normativa, diversi esperti hanno sottolineato che «la mayoría de los países latinoamericanos propone regulaciones inspiradas o con similitudes importantes a la propuesta de Ley de Inteligencia Artificial de la Unión Europea (AI Act)».

⁵⁰ Il testo della decisione è reperibile sul sito internet istituzionale della *Corte constitucional de Colombia*, al seguente collegamento: <https://www.corteconstitucional.gov.co/relatoria/2024/T-323-24.htm>.

⁵¹ La pronuncia, infatti, è stata resa nell'ambito di un'*acción de tutela*, ipotesi riconducibile all'art. 241 della Costituzione colombiana, che annovera tra le competenze della Corte costituzionale la revisione – per l'appunto – delle azioni con cui i cittadini possono chiedere ai giudici ordinari la tutela immediata dei propri diritti costituzionali.

⁵² Il riferimento è alla Ley del Perú n. 31814 del 5 luglio 2023, che «promueve el uso de la inteligencia artificial en favor del desarrollo económico y social del país».



Difatti, è proprio alle previsioni eurounitarie che sembra ispirarsi la Corte colombiana quando individua un elenco di dodici “criteri guida” per l’uso corretto dell’intelligenza artificiale da parte degli uffici giudiziari del Paese. In particolare, avendo come punto di riferimento il diritto al giusto processo, si chiarisce che l’utilizzo dell’IA si deve fondare su: 1. trasparenza; 2. responsabilità; 3. *privacy*; 4. non sostituzione della razionalità umana; 5. serietà e verifica delle fonti; 6. prevenzione del rischio; 7. eguaglianza ed equità; 8. controllo umano; 9. regolamentazione etica; 10. adeguamento a buone pratiche e *standard* collettivi; 11. monitoraggio continuo e adattamento ai progressi giuridici, sociologici e tecnologici; 12. idoneità a facilitare e rendere più agile l’accesso alla giustizia.

Ognuno di questi punti meriterebbe un ampio approfondimento. Tuttavia, per quel che rileva maggiormente in questa sede, è interessante notare la previsione del criterio della “non sostituzione della razionalità umana”, intesa come espressione dell’impossibilità etica e giuridica di sostituire l’azione e la responsabilità dell’individuo della specie umana nella gestione degli atti e delle decisioni giudiziarie.

In particolare, la Corte colombiana ha evidenziato come i giudici effettuo, nell’esercizio della loro attività, dei giudizi di valore che non possono essere delegati alle macchine. Più nello specifico, nella decisione si esclude che l’IA possa sostituire il ragionamento logico che un essere umano compie nell’interpretazione dei fatti, delle prove e delle ragioni sottese alla decisione stessa.

Sono molteplici le criticità che i giudici costituzionali colombiani hanno messo in luce sull’utilizzo dell’IA in sede giudiziaria, a partire da quelle di natura tecnico-scientifica e di sicurezza informatica. Tuttavia, riteniamo che quelli più insidiosi riguardino la tutela dei diritti fondamentali in sede giurisdizionale. Al riguardo, vengono in rilievo alcuni tratti essenziali del diritto ad un giusto processo, fra tutti:

a) la garanzia del giudice naturale precostituito per legge, perché può trattarsi solo di un giudice (umano) al quale la Costituzione o la legge hanno attribuito la conoscenza di una certa questione. Tale non può essere – almeno allo stato – una macchina, il cui utilizzo in maniera integralmente sostitutiva comprometterebbe tanto l’indipendenza, quanto l’imparzialità del giudice, dal momento che non è



possibile verificare l'influenza di eventuali pregiudizi che caratterizzano i *dataset* su cui si basa il funzionamento dell'IA⁵³ (tralasciando l'ulteriore criticità che si determina quando i *software* sono di titolarità di soggetti privati, come nel caso di specie);

b) l'obbligo di motivazione delle decisioni giudiziarie, perché, per ragioni analoghe a quelle analizzate fino ad ora, non sarebbe ammissibile l'inserimento nelle decisioni giudiziarie di testi generati dall'IA senza supervisione umana. In tal caso, verrebbe meno la responsabilità del giudice rispetto all'obbligo di motivare la propria decisione, che risulterebbe il frutto di processi logici – spesso opachi – elaborati dalla macchina, che potrebbero anche riflettere pregiudizi derivanti dai dati con cui è stata addestrata (inesatti, non aggiornati, parziali), rendendo erronea o irragionevole la motivazione;

c) la corretta valutazione dei fatti e delle prove per la determinazione del libero convincimento del giudice, all'esito del contraddittorio tra le parti in condizione di parità, davanti a un giudice terzo e imparziale. Anche qui, il ruolo del giudice sembrerebbe insostituibile, dal momento che delegare tale compito a un'IA comporterebbe, anzitutto, una serie di rischi quali valutazioni distorte o incomplete, violazioni della riservatezza delle parti coinvolte o la possibile esposizione di dati sensibili alle società private titolari del *software* (come nel caso *ChatGPT*). Non solo, perché, come osservato in dottrina, «una sentenza che traesse la propria legittimazione dal giudizio dei pari è [...] principio divenuto intrinseco al *due process* e, *mutatis mutandis*, al giusto processo. In tale logica, pare improponibile delegare la funzione giurisdizionale a macchinari dotati di AI. Anche se con il tempo si risolvessero gli ordini di problemi di natura tecnico-scientifica segnalati e se le

⁵³ Sul punto, cfr. G. MOBILIO, *I giudici e le nuove tecnologie per giudicare: una occasione per riscoprire i caratteri fondanti della funzione giurisdizionale*, cit., p. 486 ss., che evidenzia come «il mito dell'imparzialità della macchina risulta presto sfatato se solo si considera come durante la progettazione degli algoritmi avvengano una serie di scelte aventi ad oggetto sia questioni di politica criminale, sia questioni relative alle circostanze valutabili in sede di giudizio, che si traducono nella scelta delle informazioni che gli algoritmi prendono in considerazione o nel modo con cui esse vengono processate. Se queste operazioni sono falsate da pregiudizi, siano essi espliciti o impliciti, c'è il pericolo che la decisione proposta dall'algoritmo sia a sua volta affetta da "distorsioni" (c.d. *bias*), in grado di produrre effetti discriminatori nei confronti del destinatario della stessa».



macchine potessero dirsi addirittura superiori nel condurre interpretazioni giuridico-algoritmiche, rimarrebbe sempre la distanza fra il genere umano e quello artificiale; distanza che impedirebbe di parlarne come di nostri pari»⁵⁴.

D'altra parte, la Corte costituzionale osserva come simili tecnologie potrebbero essere legittimamente utilizzate nel sistema giudiziario:

a) se limitate ad aree di mera gestione amministrativa e documentale (come la diffusione dell'ordine del giorno per la programmazione dei procedimenti, la digitalizzazione e classificazione dei fascicoli o l'utilizzo di motori di ricerca giurisprudenziale);

b) come mero supporto alla funzione giurisdizionale, in attività che non comportano la produzione di contenuti, l'interpretazione dei fatti e delle norme o la soluzione dei casi (come l'individuazione di precedenti giurisprudenziali, l'analisi di dati statistici, l'individuazione di argomenti ripetitivi o l'interazione con altri utenti dell'amministrazione della giustizia);

c) la correzione e la sintesi dei testi, a condizione che questi siano successivamente supervisionati da un operatore giuridico (umano) per determinarne l'accuratezza, la coerenza e la corretta applicazione della normativa di riferimento.

A ben vedere, pur essendo stata adottata dall'altra parte dell'Atlantico, tale decisione sembrerebbe fare propria la distinzione tra attività giurisdizionale in senso stretto e amministrazione della giurisdizione presente nel regolamento europeo sull'IA, ove solo la prima è classificata ad alto rischio. Difatti, leggendo il *considerando* n. 61 si scorge la medesima bipartizione, laddove, da una parte, i sistemi di IA destinati all'amministrazione della giustizia vengono classificati come sistemi ad alto rischio, dall'altra, si chiarisce che «non è tuttavia opportuno estendere la classificazione dei sistemi di IA come ad alto rischio ai 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».

⁵⁴ Così C. CASONATO, *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, cit., p. 124 s.



Tale distinzione è risultata determinante anche nel caso di specie. Difatti, la Corte costituzionale colombiana ha osservato come, nella vicenda sottoposta al suo scrutinio, non vi sia stata alcuna *sostituzione* dell'esercizio della funzione giurisdizionale da parte di *ChatGPT*, essenzialmente perché il sistema di IA era stato utilizzato solo dopo che la decisione era stata sostanzialmente presa dal giudice.

Se si guarda all'*iter* logico e metodologico seguito, infatti, il giudice: anzitutto ha ricostruito la fattispecie fattuale; in seguito, ha individuato i principi costituzionali e le altre norme applicabili al caso concreto, avvalorando le proprie considerazioni con riferimenti giurisprudenziali relativi a casi analoghi; infine, ha fornito la "propria" soluzione del caso sottoposto al suo giudizio.

Solo a valle dell'intero processo decisionale sopra descritto, il giudice ha interpellato *ChatGPT*, rivolgendo al *chatbot* alcune domande astrattamente determinanti per caso di specie, ma che sono servite – così ha affermato anche la Corte costituzionale colombiana – solo a confermare la correttezza delle soluzioni individuate⁵⁵, al fine di ottimizzare il tempo impiegato nella stesura della sentenza, comunque dopo una verifica circa la correttezza delle informazioni fornite dall'IA.

A ben vedere, quindi, con considerazioni che sembrerebbero perfettamente mutuabili nel sistema italiano, la Corte costituzionale colombiana ha ribadito la necessità che la decisione giudiziaria rimanga prerogativa umana.

Nella medesima direzione, infatti, sembra essere andato anche il legislatore italiano, con la presentazione, nel maggio del 2024, del disegno di legge recante «Disposizioni e deleghe al Governo in materia di intelligenza artificiale» (S. 1146), approvato dal Senato della Repubblica (20 marzo 2025) e, con modificazioni, dalla Camera dei deputati (25 giugno 2025) e attualmente tornato all'esame

⁵⁵ Tuttavia, non può trascurarsi l'ipotesi, messa in luce da A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, cit., p. 81, secondo cui la convenienza pratica del ricorso all'algoritmo farà sì che quest'ultimo tenda comunque a "catturare" la decisione giudiziale, per cui «per chi vede violato un suo diritto o una sua libertà, dimostrare che la decisione lesiva sia stata presa unicamente sulla base di un algoritmo è molto spesso una *probatio diabolica*; chi ha preso la decisione, infatti, avrà sempre la possibilità di sostenere che nella decisione si è solo avvalso dell'algoritmo, ma non è stato "sostituito" da esso, avendo autonomamente [...] aderito alla valutazione effettuata dall'algoritmo».



del Senato per la terza lettura. Ebbene, all'art. 15 (dedicato proprio all'«Impiego dei sistemi di intelligenza artificiale nell'attività giudiziaria»), il d.d.l. italiano sull'IA stabilisce che «1. Nei casi di impiego dei sistemi di intelligenza artificiale nell'attività giudiziaria è sempre riservata al magistrato ogni decisione sull'interpretazione e sull'applicazione della legge, sulla valutazione dei fatti e delle prove e sull'adozione dei provvedimenti. 2. Il Ministero della giustizia disciplina gli impieghi dei sistemi di intelligenza artificiale per l'organizzazione dei servizi relativi alla giustizia, per la semplificazione del lavoro giudiziario e per le attività amministrative accessorie».

Questa soluzione risulta in linea con le previsioni dell'*AI Act* sull'utilizzo di sistemi di IA ad alto rischio, specie per quanto riguarda la necessità che essi siano costantemente supervisionati dall'essere umano. Ma, soprattutto, la previsione proposta dal legislatore italiano sembrerebbe coerente con la complessità e la profondità del ragionamento giuridico che, assieme ai necessari giudizi di valore che lo caratterizzano, non possono essere ridotti alla logica computazionale degli algoritmi, al fine di preservare l'autonomia decisionale dei giudici e il carattere umano della motivazione, elementi imprescindibili per garantire l'equità e la legittimazione del sistema giudiziario, oltre che una piena tutela dei diritti coinvolti nella decisione.