

# Discrezionalità artificiale e giudizio algoritmico

## Artificial Discretion and Algorithmic Judgment

LUIGI COMINELLI<sup>1</sup>

### Sommario

L'articolo esplora l'impatto della digitalizzazione e dell'intelligenza artificiale (IA) sul sistema giudiziario e amministrativo, concentrandosi su "discrezionalità artificiale" e "giustizia digitale". La digitalizzazione può semplificare gli adempimenti, ma una sua cattiva implementazione può complicare ulteriormente i processi. Le tecnologie IA, invece, presentano sfide qualitative: possono automatizzare l'apprendimento e la decisione, adattando il comportamento in base ai dati acquisiti. Tuttavia, l'IA comporta anche rischi di errori. Nel contesto giuridico, l'IA non sostituirà a breve l'analisi delle politiche pubbliche, ma piuttosto la supporterà, ad esempio, attraverso l'analisi degli umori del pubblico sui social network e la realizzazione di analisi di impatto più accurate. La "giustizia digitale" include l'uso di algoritmi per decisioni giudiziarie, sollevando preoccupazioni su *bias* e pregiudizi. Sebbene l'IA possa supportare il lavoro intellettuale, il decisore umano rimane insostituibile in molte situazioni, in particolare quando è necessaria discrezionalità. È necessario un uso etico e appropriato dell'IA, per evitare di pregiudicare l'equità delle decisioni e mantenere il controllo umano nelle decisioni critiche.

**Parole chiave:** diritto, algoritmo, discrezionalità, decisione, amministrazione, giurisdizione

### Abstract

The article explores the impact of digitisation and artificial intelligence (AI) on judicial and administrative systems, focusing on 'artificial discretion' and 'digital justice'. Digitisation, in a broad sense, can simplify tasks, but its poor implementation can complicate processes further. AI technologies, by contrast, present qualitative challenges: they can automate learning and decision-making by adapting behaviour based on accumulated data. However,

---

<sup>1</sup> Dipartimento di scienze Giuridiche "C. Beccaria", Università degli Studi di Milano. [luigi.cominelli@unimi.it](mailto:luigi.cominelli@unimi.it).



AI also entails the risk of errors. In the legal context, AI is unlikely to replace public policy analysis in the near future; instead, it is expected to support it, e.g., by analysing public sentiment on social networks and conducting more accurate impact assessments. ‘Digital justice’ includes the use of algorithms in judicial decisions, raising concerns about bias and prejudice. Although AI can support intellectual work, human decision-makers remain irreplaceable in many situations, particularly when discretion is required.

**Keywords:** Law, Algorithm, Discretion, Decision, Administration, Jurisdiction

## 1. Introduzione

La relazione/articolo di Monica Raiteri cita i due aspetti della semplificazione e della proceduralizzazione. Sono due valori entrambi apprezzabili scaturenti da principi costituzionali, ma il cui potenziale contrasto non è sempre letto e gestito correttamente dalle nostre istituzioni.

La digitalizzazione in senso lato implica l’abbandono della burocrazia cartacea e, se implementata con gli strumenti giusti, porta alla semplificazione degli adempimenti. Come sappiamo bene, non sempre è così: una digitalizzazione di cattiva qualità o realizzata con strumenti impropri, uniti a un modesto tasso di alfabetizzazione informatica nella popolazione<sup>2</sup>, cancellano ogni potenzialità semplificatoria e anzi creano ulteriore complicazione. Anche la proceduralizzazione del diritto a scopo di garanzia, paventa Raiteri, rischia di incidere negativamente sull’asimmetria di potere tra cittadino e istituzioni. Le dinamiche burocratiche descritte dai maestri della sociologia (Merton 1983; Weber 1995) si replicano in epoca moderna con le società democratiche (Coser 1974; Crozier 1962), ed ineluttabilmente producono logiche di chiusura e distanziamento che si scontrano con gli scopi dichiarati da qualsiasi riforma amministrativa (Cassese 2001) e con le esigenze di apertura indicate dalle dottrine del New Public Management (Peters e Pierre 1998).

Oggi la digitalizzazione indica però in un senso più specifico anche tecnologie che impattano sul contenuto dei nostri diritti. L’effetto della tecnologia sul mondo del diritto è un tema che sempre più dovrà interessare giuristi e cittadini. Gli avanzamenti nel campo dell’intelligenza artificiale generativa hanno alzato il livello dell’attenzione: il 63% degli avvocati interpellati a livello internazionale da LexisNexis ritiene che l’IA cambierà il

2 L’Italia, ad esempio, è al quart’ultimo posto tra i paesi dell’Unione europea per competenze digitali: Istituto Nazionale di Statistica, *Cittadini e competenze digitali*, <https://www.istat.it/it/files/2023/06/cs-competenzedigitali.pdf>, consultato il 15 maggio 2024.

modo in cui il diritto è insegnato e studiato; il 47% ritiene che vi sarà un impatto significativo o trasformativo sull'esercizio delle professioni forensi<sup>3</sup>. Complessivamente, il 77% crede che l'impatto sarà in generale positivo, permettendo al legale di concentrarsi sulle attività a più alto valore intellettuale come l'analisi del caso e la consulenza, ad aiutandolo nelle attività ripetitive o di natura amministrativa e segretariale. Le preoccupazioni riguardano la qualità e l'affidabilità del lavoro delle IA, e la necessità di un uso etico<sup>4</sup>. Non si tratterebbe sicuramente della “fine degli avvocati” preconizzata da Susskind (2023).

Tra gli effetti dell'automazione del lavoro intellettuale c'è lo stesso “effetto di sostituzione” della forza lavoro umana da parte del capitale che già da qualche secolo interessa il lavoro manuale (Acemoglu e Restrepo 2019, 2020). Con taluni tipi di automazione a basso valore aggiunto, tale effetto rischia semplicemente di fare scomparire posti di lavoro senza creare contestualmente opportunità e campi di impiego alternativi (Acemoglu e Restrepo 2020, p. 3). Tra gli esempi di automatizzazione positivi, che cioè non minacciano la centralità del fattore produttivo umano, vi sono l'*educazione*, dove le modalità di interazione dell'insegnante con la classe non sono cambiate negli ultimi secoli, e dove potrebbe essere utile potenziare l'insegnamento con attività dirette a singoli o sottogruppi, oppure l'*assistenza sanitaria*, dove non mancano le opportunità di svolgere compiti ausiliari di cura, rispetto ai quali il lavoro umano non sempre riesce a tenere il passo, per finire con la *realtà aumentata*, grazie alla quale il lavoro umano si gioverebbe di un incremento delle abilità percettive e di monitoraggio (Acemoglu e Restrepo 2020, p. 7). Una buona automazione incrementa notevolmente la produttività, e contribuisce indirettamente alla creazione di domanda per compiti non automatizzabili nei quali il fattore umano ha un vantaggio competitivo, oltre a creare nuovi impieghi nel settore dell'automazione stessa. Sono quindi paradossalmente le forme di automazione che creano minore incremento di produttività e dunque meno innovative, a creare un rischio di sostituzione.

Nel campo delle professioni intellettuali, le ricerche empiriche sulla consulenza professionale hanno evidenziato che se l'applicazione dell'IA ai *large language models* tende a generare output di qualità media percepita più alta e in un tempo minore, l'intervento umano è ancora quello che può generare soluzioni più innovative e salti di paradigma (Bossieux et al. 2023, p. 5). In questi contesti, il noto problema delle “allucinazioni” dei *large language models*, cioè il rischio che l'IA restituisca risposte verosimili ma fondate su circostanze del tutto inventate, può essere trasformato in uno strumento di generazione di idee a supporto della creatività umana (Girotra et al. 2023, p.

---

3 Lexis Nexis, International Legal Generative AI Report, 2023, pp. 3-4.

4 Ivi, p. 11.

4). L'intervento umano è comunque essenziale nella valutazione e selezione delle opzioni generate.

Un altro recente studio ipotizza che i maggiori benefici delle IA generative riguardano i lavoratori intellettuali che si collocano sotto la media delle performance (Dell'Acqua et al. 2023, p. 11). Uno dei rischi impliciti in questo caso sarebbe quello della sostituzione dei lavoratori ancora in fase di training, come i più giovani (Dell'Acqua et al. 2023, p. 18). Alcuni esperimenti hanno provato che chatGPT e sistemi generativi concorrenti sono in grado di arrivare molto vicino a passare (talvolta riuscendoci) gli esami di abilitazione alla professione medica, senza nessuno tipo di training specializzato (Kung et al. 2023).

## 2. Automazione, intelligenza artificiale e diritto

Problemi parzialmente diversi pone l'uso dell'intelligenza artificiale nell'azione amministrativa e di governo. Il fatto che queste ultime siano improntate alla discrezionalità ha spinto taluni a coniare l'espressione "discrezionalità artificiale" (Young, Bullock, and Lecy 2019), applicabile ad almeno quattro aspetti fondamentali delle attività amministrative: analisi, previsione, decisione e monitoraggio<sup>5</sup>. Anche se ogni attività decisionale può già da tempo giovare degli strumenti informatici e della loro potenza di calcolo e di organizzazione delle informazioni, l'intelligenza artificiale modifica il processo da un punto di vista qualitativo in virtù di due aspetti: 1) è progettata per automatizzare l'apprendimento e la decisione grazie alla rappresentazione matematica e astratta dei problemi; 2) quanto più acquisisce nuovi dati, tanto più adatta il proprio comportamento modificando le euristiche decisionali (Young, Bullock, e Lecy 2019, p. 301).

Come per ogni sistema automatizzato, sussiste sempre il rischio di errori. Un celebre esempio è quello dell'autorità olandese per il welfare dell'infanzia, la quale ricorse a un sistema automatizzato che irrogò automaticamente sanzioni sulla base di un semplice sospetto di frode determinato da un algoritmo statistico. Oltre mille bambini furono temporaneamente sottratti alla potestà dei genitori, e tra questi ultimi vi furono dei casi di suicidio. Un

---

5 Tra gli esempi virtuosi di Intelligenza artificiale e discrezionalità artificiale c'è il caso dell'Estonia, dove il fondo per la disoccupazione usa ad esempio sistemi predittivi per determinare le probabilità che una persona disoccupata riesca a trovare un nuovo lavoro in base a determinate variabili; il Ministero dell'ambiente determina la specie degli alberi e l'altezza media delle piante nelle foreste, e individua i casi di abbattimenti illegali; l'ente di gestione delle strade e la polizia di frontiera mappano i segnali stradali e tentano una previsione degli incidenti stradali; le autorità di vigilanza fiscale selezionano le imprese per le verifiche fiscali sulla base delle dichiarazioni presentate e della probabilità statistica che siano stati effettuati pagamenti in nero (Ebers e Trasberg 2023, p. 11).

altro celebre caso fu quello del *Robodebt* australiano, che inviò erroneamente avvisi di pagamento a circa 400.000 cittadini (Ebers e Trasberg 2023, pp. 19, 21).

In taluni casi, i sistemi automatizzati di tracciatura e riconoscimento hanno fornito l'occasione per un pronunciamento dei tribunali costituzionali. Nel 2018 la Corte costituzionale tedesca ha dichiarato l'illegittimità dei sistemi di riconoscimento delle targhe automobilistiche in tre Länder, sulla base del diritto a muoversi liberamente senza che gli spostamenti siano registrati arbitrariamente dalle autorità pubbliche e senza che il cittadino debba provare la propria integrità o si senta sotto costante sorveglianza (Ebers e Trasberg 2023, pp. 19, 23).

Ebers e Trasberg si chiedono se l'utilizzo dell'AI da parte di un ente pubblico nell'esercizio dei suoi poteri non configuri una delega di poteri pubblici. Il rischio è che il decisore umano, in cerca come è di parametri "oggettivi" che limitino la sua discrezionalità, si affidi supinamente al suggerimento dell'AI, finendo così per pregiudicare l'equità della decisione. La normativa tedesca sull'assunzione automatizzata di decisioni stabilisce che un atto amministrativo possa essere adottato integralmente in maniera automatizzata solo quando ciò sia previsto dalla legge e quando non siano richiesti discrezionalità o valutazioni<sup>6</sup>. Dobbiamo quindi escludere che nella tradizionale discrezionalità amministrativa possa essere compresa la scelta stessa di affidare motivatamente una decisione a un sistema automatico?

Il tema che tuttavia mi preme espressamente esplorare è quello della "giustizia digitale" citato da Raiteri, ossia degli algoritmi che prendono decisioni di rilevanza amministrativa o giudiziaria. Sussiste il rischio concreto che tali metodi di decisione finiscano per tradursi in una profezia che si autoavvera, e che cristallizzino pregiudizi e bias preesistenti, legittimando in ultima analisi lo status quo.

Qui occorre però fare una distinzione essenziale tra algoritmi come criteri di decisione automatica basati su regole definite a monte, e algoritmi come sistemi di "deep machine learning", che si evolvono autonomamente imparando dagli esempi. Il problema posto dai secondi è meno impellente di quanto non appaia, dal momento che il deep learning è ancora largamente perfettibile, e anche nel lungo periodo potrebbe non raggiungere mai livelli di intelligenza generale comparabili a quelli umani (Faggin 2022). Sul punto esistono opinioni discordanti anche tra i maggiori esperti, ma è importante puntualizzare che i *large-language models* divenuti di recente popolari grazie all'illusione di un'intelligenza naturale che producono, sono in grado di svolgere solo dei task limitati, e non sono certo al livello dell'in-

---

<sup>6</sup> Si veda il nuovo art. 35a del codice degli atti amministrativi della Repubblica Federale tedesca, introdotto nel 2017 ("Gesetz zur Modernisierung des Besteuerungsverfahrens," Vol. 18 July 2016, BGBl. I, p. 1679 - no. 35).

telligenza umana, anche a prescindere dai palesi errori e delle allucinazioni di cui soffrono<sup>7</sup>.

Il punto di riferimento per l'intelligenza artificiale oggi è l'*AI Act* dell'Unione europea<sup>8</sup>:

*Art. 3 (1) 'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environment.*

In passato la Commissione europea<sup>9</sup>, e successivamente un gruppo di esperti da questa nominati, avevano offerto una definizione ancora più esauriente:

*Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.*

*As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems)."<sup>10</sup>*

7 Basti citare l'esempio dei sistemi autonomi di guida di un veicolo, cioè di un'attività che richiede un'intelligenza tutto sommato limitata. Nella gran parte dei paesi, per guidare è sufficiente un esame di abilitazione dopo il compimento dei 16/18 anni di età, senza nessun particolare titolo di studio. Eppure, i sistemi autonomi di guida non sono ancora abbastanza sicuri da essere autorizzati all'uso generalizzato, se non in contesti normativi e ambientali molto specifici e con la supervisione umana.

8 *AI Act*, 13 marzo 2024: Risoluzione legislativa del Parlamento europeo del 13 marzo 2024 sulla proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)).

9 Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni: L'intelligenza artificiale per l'Europa, Brussels, 25.4.2018 COM(2018) 237 final.

10 A definition of AI: Main capabilities and scientific disciplines. High-Level Expert Group on Artificial Intelligence, 2019, [https://ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=56341](https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56341). Sembra tuttavia mancare in questa definizione, a dimostrazione della natura

L'approccio regolamentare dell'AI Act si basa su una definizione di livelli di rischio che a seconda del potenziale impatto di ogni singola applicazione richiederanno cautele diverse. Anche nel caso di *rischio alto*, le relative applicazioni dovranno essere fortemente limitate, ma non necessariamente vietate.

Particolarmente interessanti a proposito dell'attività amministrativa e giurisdizionale sono due dei "considerando" introduttivi. Il n. 58 considera le prestazioni e i servizi essenziali di assistenza pubblica, ed afferma in maniera piuttosto apodittica che nella decisione sulla concessione o sulla revoca, la delega all'IA dovrebbe essere comunque possibile, purché ciò non comporti un rischio alto.

In maniera più puntuale, il considerando n. 61 classifica come intrinsecamente ad alto rischio "i sistemi di IA destinati all'amministrazione della giustizia e ai processi democratici", in particolare quando svolgono "attività di ricerca e interpretazione dei fatti e del diritto e nell'applicazione della legge a una serie concreta di fatti", e quando sono usati da "organismi di risoluzione alternativa delle controversie [che] producono effetti giuridici per le parti". Il testo afferma che il potere decisionale e l'indipendenza dei giudici non possono in ogni caso essere sostituiti. A rischio "non alto" dovrebbero essere considerate tutte le attività amministrative puramente accessorie come l'anonimizzazione delle decisioni e la comunicazione tra il personale.

Cosa questo possa significare in concreto non è sempre chiaro, ed è ipotizzabile che in un secondo momento l'Ue provvederà a normare in maniera più casistica. Peraltro, nel contesto normativo e giudiziario attuale, il ricorso a sistemi automatici e algoritmici semplici e basati su dati pubblici potrebbe in molti campi costituire un progresso rispetto ai sistemi decisionali intuitivi, rimessi totalmente alla discrezionalità e soggetti a bias sistematici (Kahneman and Tversky 2000), e potrebbe addirittura contribuire alla riduzione di discriminazioni e disuguaglianze.

Accanto ai bias, non possiamo nemmeno trascurare l'enorme impatto del *rumore di fondo*, rispetto a cui esistono raccolte di dati empirici che testimoniano della pericolosità e dell'irrazionalità dei processi decisionali pacificamente praticati in campo legale, giudiziario, contrattuale, medico-clinico e ingegneristico (Kahneman, Sibony e Sunstein 2021). Il rumore di fondo è costituito da ogni istanza di variazione randomica involontaria (si pensi alla distribuzione imprecisa delle freccette sul bersaglio). Rappresenta qualcosa di diverso dai bias, e fino ad ora non è stato altrettanto studiato: oltre a tutto è facilmente confondibile con i bias, perché tende a sovrapporsi e confondersi.

---

inesorabilmente limitata di ogni attività definitoria, il riferimento a quei sistemi intelligenti che si ipotizza possano essere a loro volta concepiti da altri sistemi intelligenti.

È ad esempio ampiamente accettato che la decisione sull'affidamento di un minore o la durata di una sentenza detentiva siano soggetti a un'ampia variabilità anche in casi simili, e ciò è difficilmente spiegabile anche controllando l'effetto che il bias del singolo giudice imprime. Questo rumore di fondo crea decisioni irragionevolmente casuali, e nel campo della giustizia rappresenta un magnifico esempio della combinazione tra materialità e irrazionalità<sup>11</sup>. In campo medico, uno studio condotto nel 1964 su 91 pazienti esaminati da 10 psichiatri esperti scopre che le diagnosi concordavano solo nel 57% dei casi (Kahneman, Sibony e Sunstein 2021, p. 229; Aboraya et al. 2006).

Il rumore di fondo provoca incertezze anche nella pratica commerciale: una meta-analisi evidenzia che le previsioni tecniche sulla durata dei lavori arrivano a differire anche del 71% (Kahneman, Sibony e Sunstein 2021, pp. 8ss). I premi di polizza fissati dagli assicuratori hanno tra loro una differenza mediana del 55%, ossia cinque volte quanto stimato dal management dell'azienda (il quale a sua volta soffre evidentemente di un *agreement illusion*), e la differenza tra le stime dei liquidatori è del 43%.

Tutte queste variazioni nell'esito decisionale sono spiegabili solo ricorrendo a fattori extra-razionali incontrollabili, ai quali si potrebbe almeno in parte rimediare ricorrendo maggiormente a regole anche automatiche (che sono più lunghe e costose in fase di elaborazione), e limitando invece il ricorso agli standard generali (che richiedono discrezionalità e quindi tempi più lunghi in fase di applicazione). Con questi ultimi, ci si affida a valutazioni cliniche basate su una considerazione intuitiva e non sistematica dei precedenti (Grove and Meehl 1996), introducendo appunto un rumore irrazionale e iniquo: casi diversi sono trattati allo stesso modo, e casi identici sono trattati in modo diverso. Esempi ben diffusi e funzionanti di regole automatiche che orientano le decisioni giudiziarie sono quei criteri semplificati forniti attraverso linee guida, tabelle dal danno biologico o *sentencing guidelines*.

### 3. Intelligenza artificiale e decisione

Nella percezione comune, le nuove applicazioni dell'intelligenza artificiale sembrano arrivate a maturazione e dunque pronte a svolgere compiti sensibili e di evidenza pubblica. Come si è argomentato, tale risultato non è ancora definitivamente acquisito, e nei compiti per i quali sia richiesto uno standard anche ordinario di diligenza, non si sono ancora visti casi impor-

---

11 Simonelli (2020) ricorda come lo stesso Weber sottolineasse l'appetibilità di un sistema decisionale giuridico di tipo meccanicistico e deterministico.

tanti di delega nell'assunzione di decisioni senza una importante supervisione del soggetto umano.

Posto tuttavia che anche questo ultimo passaggio potrà avvenire a breve, e che dunque si dovrà iniziare a ragionare in senso proprio di responsabilità, di diligenza, e di imputabilità con riferimento a un'intelligenza artificiale, si apre il tema dell'*accountability* democratica di un sistema normativo che in ultima istanza potrebbe definirsi autoreferenziale e autopoietico. Questo appare particolarmente preoccupante quando la massa di dati sui quali l'IA si esercita e impara è chiaramente sbilanciata verso specifici settori della popolazione e particolari aree del mondo, e prende dunque a modello una realtà che non è rappresentativa. Nonostante non si possa certamente ascrivere all'IA un auto-interesse, manca in essa anche quella capacità di auto-critica riflessiva che conduce a soppesare in un secondo momento il proprio giudizio e rivalutarlo alla luce di altre prospettive o del trascorrere del tempo.

In campo giudiziario, uno degli studi più citati scoprì che il software utilizzato per prevedere la possibilità di una recidiva e per guidare dunque la decisione del magistrato di sorveglianza nella concessione dei benefici, era appena marginalmente migliore rispetto alla decisione presa da un soggetto senza particolari competenze nel campo giustizia penale (65% contro il 62%) (Dressel e Farid 2018). Uno studio successivo ha tuttavia sollevato dei dubbi sul setting sperimentale, che aveva di fatto aiutato la decisione del soggetto umano, orientandola verso gli elementi più importanti e migliorandola attraverso un feedback continuo, cosa che nel contesto decisionale reale non si verifica (Lin et al. 2020).

Più la decisione è contestualizzata e ricca di dati, più il ragionamento e la decisione umana divengono infatti ondivaghi e incoerenti, proprio come nel caso del rumore. Lo studio sembra in sostanza arrivare alla conclusione che nei contesti giudiziari reali, l'umano potrebbe arrivare a una predittività simile a quella di un algoritmo relativamente semplice, trasparente e correttamente istruito, solo con enorme sforzo di tempo e di training, e che quindi l'IA come strumento di supporto sarebbe nell'interesse di tutti, oltre che dei principi di economia ed efficienza dei procedimenti giudiziari e decisionali.

La soluzione di recente proposta da Lobel (2022) è di definire e regolare come un *commons* l'utilizzo che i sistemi di intelligenza artificiale fanno dei dati. D'altro canto, una volta prevenuti i rischi più importanti, l'intelligenza artificiale offre potenzialità innumerevoli di individuazione e di riduzione dei comportamenti discriminatori in maniera anche molto più efficace rispetto al contenzioso fondato su norme di carattere generale.

Gli esempi di queste discriminazioni sottili, che sfuggono perlopiù ai meccanismi del diritto, sono numerosi e ormai noti: dal gap retributivo tra uomo e donna (quantificabile tra il 20-30% a parità di mansioni ed esperienza), alla discriminazione dei profili lavorativi in fase di selezione del

personale, per finire con l'iniquità nella selezione di candidati ugualmente qualificati che applicano per l'ammissione a un corso di master o di dottorato (largamente basata in ultima istanza sull'università di provenienza). Si tratta come è ormai chiaro anche di bias impliciti, che come tali operano su soggetti talvolta inconsapevoli di perpetrare una discriminazione.

Nel caso delle politiche di assunzione, molti enti in passato ritenevano che fossero più effettivi il passaparola e la conoscenza diretta, ma tendevano in questo modo a reclutare gruppi di dipendenti sempre più omogenei, privilegiando chi avesse le stesse caratteristiche. Se anche questo metodo fosse semplicemente sostituito da sistemi algoritmici e automatizzati addestrati sulla base degli attuali dipendenti, si finirebbe per contribuire al perdurare della discriminazione. D'altro canto, gli stessi algoritmi sono in grado di eliminare negli annunci per le nuove assunzioni quei termini che sono repulsivi per un certo gruppo di persone. Si è notato ad esempio che l'uso di parole come *partnership* o *passione per l'apprendimento* riuscissero ad attrarre più donne. Rischi analoghi possono derivare dai software per il riconoscimento facciale, o da sistemi di intelligenza artificiale che inferiscono il livello di salute di una popolazione di riferimento dai livelli di spesa sanitaria.

Sappiamo invece che la diversità può portare vantaggi economici anche in termini di appetibilità del prodotto e di produttività del lavoratore<sup>12</sup>. In ultima analisi, è comunque l'istituzione sapiente che deve mettere in atto meccanismi di igiene decisionale che impediscano al decisore (giudice o altri) di impigrirsi e appiattare la sua decisione sulla probabilità statistica di un fatto, e di orientare e addestrare in questo senso gli algoritmi decisionali, che rispetto all'output finale restano pur sempre neutri. Sarà poi opportuno che la legislazione o la normativa regolamentare esplicitino il principio per cui non è detto che a una probabilità di recidiva più alta, considerando le circostanze personali, debba necessariamente corrispondere una decisione di diniego di un beneficio o di inasprimento della sanzione.

Così come è intuibile che la decisa preferenza a sottoporsi esclusivamente a un decisore umano e l'istintiva refrattarietà a sottoporsi a un giudizio automatizzato (cd "algorithm avoidance"), anche quando è chiaro che questo sia di qualità superiore, possano essere contrastate consentendo al cittadino di fornire un feedback che modifichi, magari anche di poco, l'output del sistema, restituendo una sensazione di controllo sul processo (Young, Bullock, and Lecy 2019, p. 311).

---

12 Lobel (2022, p. 165, p. 60) mette in evidenza la consistenza di questo "premio" sia nei risultati dei film con un cast più diversificato (in media 130 milioni di dollari di incassi in più rispetto alla media), sia nei risultati delle società con un consiglio di amministrazione più inclusivo (profitti superiori del 91% e valutazioni delle partecipazioni azionarie maggiorate del 36%).

#### 4. Conclusioni

Tornando alle tesi di Raiteri, mi sento di condividere la sua cauta apertura anche rispetto a questo tipo di digitalizzazione, il cui rifiuto ideologico finirebbe per essere controproducente. Circa la validazione umana ex-post sulla decisione automatica, che secondo Raiteri richiederebbe “una capacità computazionale [del decisore umano] superiore a quella dello strumento tecnologico”, credo che l'*arrêt* giurisprudenziale citato sia interpretabile anche nel senso che non sia necessariamente richiesta una superiorità computazionale ma semplicemente cognitiva (qualcosa di più e di meno al tempo stesso), trovandosi il decisore umano ancora su un piano di superiorità sia a livello di intelligenza generale che di sensibilità politica. Ciò rifletterebbe la natura necessariamente semantica dell'attività di interpretazione giuridica, a fronte della natura sintattica dei Big Data (Ferrari 2020).

Condivido inoltre il timore dell'Autrice che la proceduralizzazione, spinta oltre un certo limite, si ritorca contro l'apertura e il controllo democratico, creando un sovraccarico cognitivo che di fatto allontana il cittadino dalla dinamica della decisione pubblica. È un problema che spesso ritorna<sup>13</sup>, e che testimonia dell'incapacità del potere pubblico di imparare dai propri errori, insistendo ulteriormente sul concetto che più regolamentazione equivalga a più diritti.

Ancora più che in altri terreni di ricerca, è inoltre necessario un ulteriore sforzo di dialogo con le altre scienze sociali, dall'economia, alla psicologia, e certamente anche con le discipline tecnico-scientifiche, poiché attingere da una conoscenza specialistica in maniera sempre e solo mediata, ho il timore che finisca per produrre una visione distorta su aspetti tecnici apparentemente marginali, ma che in realtà potrebbero poi finire per rivelarsi centrali.

Pragmaticamente, sul problema del rumore e degli errori sistematici umani, l'osservazione e la pratica suggeriscono l'adozione di prassi di igiene decisionali che ogni istituzione potrebbe e dovrebbe recepire: il ricorso a un terzo neutrale che, senza poter intervenire nel merito, osservi dall'interno il processo decisionale e ne corregga i difetti; l'affidamento di fasi diverse di uno stesso processo decisionale a soggetti distinti, il cui output possa poi in un secondo momento essere aggregato olisticamente; l'utilizzo, per quanto possibile, di giudizi comparativi graduati e in scala piuttosto che di giudizi assoluti.

Dal punto di vista dell'alfabetizzazione decisionale, è utile il ricorso alla statistica bayesiana per comprendere ad esempio il concetto della *baseline*. L'algoritmo non è sempre e necessariamente inserito in una black-box: è possibile fare ricorso ad algoritmi aperti e controllabili, che semplificano meccanicamente la decisione, come del resto già si è fatto in campo giudiziario con strumenti come linee-guida e tabelle schematiche, specialmente

13 Basti qui pensare al movimento contro il formalismo procedurale (Genn 1993).

in quei campi dove la disomogeneità tra le decisioni è particolarmente scandalosa per la sua imprevedibilità.

Benché imperfetti, si è anche argomentato che alcuni algoritmi “stupidi”, applicati meccanicamente e senza la capacità di imparare, e passibili di correzione solo su input umano, siano comunque meglio dei giudizi clinici (60% di accuratezza contro 55%) poiché eliminano del tutto quanto meno il sopracitato rumore (Kahneman, Sibony e Sunstein 2021). Il problema dell’algoritmo come profezia che si autoavvera è del resto né più né meno lo stesso problema che affligge le decisioni umane, per quanto ci si illuda che un decisore istruito, illuminato e bene intenzionato sia in grado di evitarlo (Danziger, Levav e Avnaim-Pesso 2011). La decisione sulla praticabilità deve quindi basarsi su una comparazione con la procedura preesistente e con i rischi a essa connessi, e non su petizioni di principio di natura assoluta. È presumibile che l’intelligenza artificiale non potrà soppiantare a breve l’analisi delle politiche pubbliche, ma che la supporterà ad esempio nel cogliere gli umori del pubblico attraverso i social network, e si è parlato in proposito di “intelligenza [naturale] aumentata” (Longo 2022).

Occorre naturalmente rispondere a legittime obiezioni e ansietà, alcune delle quali già affrontate a proposito dei bias razziali o discriminatori e della *predictive policing* (Boussioux et al. 2023). La soluzione al momento ipotizzabile e accettabile per il diritto può essere quella di fornire meccanismi e procedure di decisione ibrida, che lascino il controllo finale del risultato decisionale all’umano, alla cui sensibilità e intuito politico dovrà anche essere demandato il compito di stabilire quando è il momento di introdurre un elemento di novità e deviare dalle linee guida algoritmiche, perché un nuovo valore politico-sociale richiede che tutte le variabili coinvolte vadano soppesate nuovamente.

È poi necessario considerare che la creazione di un sistema algoritmico adeguato può essere ben più costosa di un sistema basato sulla decisione individuale clinica, e richiede un lavoro di sviluppo molto maggiore. Le decisioni intuitive sono più economiche: richiedono solo un decisore ispirato e carismatico (o supposto tale). Su questi temi sarebbe preferibile in generale non affidarsi ciecamente ad un principio di precauzione fondato su un’occulta presunzione di pericolosità, ma valutare caso per caso se in un determinato contesto i rischi e i difetti di una decisione automatica e algoritmica siano maggiori rispetto alla procedura decisionale in atto.

## Bibliografia

- Aboraya, A., Rankin, E., France, C., El-Missiry, A., John, C. (2006). The Reliability of Psychiatric Diagnosis Revisited: The Clinician's Guide to Improve the Reliability of Psychiatric Diagnosis. *Psychiatry*, 3, 1, pp. 41–50.
- Acemoglu, D., Restrepo, P., (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33, 2, pp. 3–30. <https://doi.org/10.1257/jep.33.2.3>
- Acemoglu, D., Restrepo, P., (2020). The wrong kind of AI? Artificial intelligence and the future of labour demand. *Cambridge Journal of Regions, Economy and Society*, 13, 1, pp. 25–35. <https://doi.org/10.1093/cjres/rsz022>
- Boussioux, L., N. Lane, J., Zhang, M., Jacimovic, V., Lakhani, K. R., (2023). The Crowdless Future? How Generative AI Is Shaping the Future of Human Crowdsourcing. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4533642>
- Cassese, S. (2001). L'esplosione del diritto. Il sistema giuridico italiano dal 1975 al 2000. *Sociologia Del Diritto*, 28, 1, pp. 55–65.
- Coser, L. A. (1974). *Greedy Institutions*, New York, Free Press.
- Crozier, M. (1962). Administration et bureaucratie. In *Transactions of the Fifth World Congress of Sociology*. 2. Washington, DC: International Sociological Association.
- Danziger, S., Levav, J., Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, 108, 17, 6889–6892. <https://doi.org/10.1073/pnas.1018033108>
- Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Lakhani, K. R., (2023). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4573321>
- Dressel, J., Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances*, 4, 1, pp. 1–5. <https://doi.org/10.1126/sciadv.aao5580>
- Ebers, M., Trasberg, H. (2023). Delegation of Administrative Powers to AI Systems, in Ebers, M., Tupay, P.K., eds., *Artificial Intelligence and Machine Learning Powered Public Service Delivery in Estonia*, Cham, Springer, pp. 85–102. [https://doi.org/10.1007/978-3-031-19667-6\\_5](https://doi.org/10.1007/978-3-031-19667-6_5)
- Ebers, M., Tupay, P. K., Juksaar, J., Kohv, K. (2023). The Promise and Perils of AI and ML in Public Administration, in Ebers, M., Tupay, P.K., eds., *Artificial Intelligence and Machine Learning Powered Public Service Delivery in Estonia*, Cham, Springer, pp. 7-33. [https://doi.org/10.1007/978-3-031-19667-6\\_2](https://doi.org/10.1007/978-3-031-19667-6_2)

- Faggin, F. (2022). *Irriducibile. La coscienza, la vita, i computer e la nostra natura*, Milano, Mondadori.
- Ferrari, V. (2020). Note socio-giuridiche introduttive per una discussione su diritto, intelligenza artificiale e big data, *Sociologia del Diritto*, 47, 3.
- Genn, H. (1993). Tribunals and Informal Justice. *The Modern Law Review*, 56, 3, pp. 393–411.
- Girotra, K., Meincke, L., Terwiesch, C., Ulrich, K. T. (2023). Ideas are Dimes a Dozen: Large Language Models for Idea Generation in Innovation. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4526071>
- Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The Clinical-Statistical Controversy, *Psychology, Public Policy, and Law*, 2, 2, pp. 293–323. <https://doi.org/10.1037/1076-8971.2.2.293>
- Kahneman, D., Sibony, O., Sunstein, C. R., (2021). *Noise: A Flaw in Human Judgement*, New York, Little, Brown and Company.
- Kahneman, D., Tversky, A., (2000). Choices, Values, and Frames, in Kahneman, D., Tversky, A., eds., *Choices, Values, and Frames*, Cambridge, Cambridge University Press.
- Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Tseng, V., (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health*, 2, 2, e0000198. <https://doi.org/10.1371/journal.pdig.0000198>
- Lexis Nexis. (2023). *International Legal Generative AI Report*.
- Lin, Z. J., Jung, J., Goel, S., Skeem, J. (2020). The limits of human predictions of recidivism, *Science Advances*, 6, 7, pp. 1–8. <https://doi.org/10.1126/sciadv.aaz0652>
- Lobel, O. (2022). *The Equality Machine. Harnessing Digitally Technology for a Brighter, More Inclusive Future*, New York, PublicAffairs.
- Longo, J. (2022). When artificial intelligence meets real public administration. *Canadian Public Administration*, 65, 2, pp. 384–388. <https://doi.org/10.1111/capa.12465>
- Merton, R. K. (1983). *Teoria e struttura sociale. I. Teoria sociologica e ricerca empirica*, Bologna, Il Mulino.
- Peters, B. G., Pierre, J., (1998). Governance without government? Rethinking public administration, *Journal of Public Administration Research and Theory*, 8, 2, pp. 223–244.
- Simonelli, M. A. (2020). Algoritmo versus ermeneutica giudiziaria. L'utopia della 'matematizzazione' del diritto, in *Annali del Dipartimento Giuridico dell'Università degli Studi del Molise*, Napoli, Edizioni scientifiche italiane, pp. 365–375.
- Susskind, R. (2023). *Tomorrow's Lawyers. An Introduction to your Future*, Oxford, Oxford University Press.

- Weber, M. (1995). *Economia e società. IV. Sociologia politica*, Torino, Edizioni di Comunità.
- Young, M. M., Bullock, J. B., Lecy, J. D., (2019). Artificial Discretion as a Tool of Governance: A Framework for Understanding the Impact of Artificial Intelligence on Public Administration, *Perspectives on Public Management and Governance*, 2, 4, pp. 301–313. <https://doi.org/10.1093/ppmgov/gvz014>