

Rischi algoritmici. Alcune questioni sulla digitalizzazione della giustizia¹

Piero Marra

Università degli Studi di Bari Aldo Moro

The recent European regulation on artificial intelligence, known as the AI Act, categorizes AI systems intended for use by judicial authorities in the investigation and interpretation of facts and law, as well as in law enforcement, as high-risk systems. However, starting from this framework, some issues arise even earlier regarding the assessment of the risks associated with the judicial use of algorithmic logic. In fact, before evaluating risks, they must first be identified—an activity that is neither simple nor automatic. In this context, numerous contradictions and challenges affect the digitalization of justice. This essay aims to illustrate the main issues, starting from the fundamentally rhetorical and dialectical nature of the judicial process.

Keywords: Digital Justice, Algorithmic Risks, Procedural Justice, Undesign Perspectives

¹ Questo lavoro è stato parzialmente sostenuto dal progetto FAIR – Future AI Research (PE00000013), nell'ambito del programma MUR del PNRR finanziato dal NextGenerationEU. Il giurista esperto potrà trovare alcune argomentazioni di questo saggio, soprattutto quelle riguardanti l'interpretazione e l'applicazione del diritto, come presupposte o, quantomeno, note e, probabilmente, troppo semplicistiche. Sono, invece, meno scontate per il non giurista, ad esempio l'esperto informatico, a cui queste pagine sono pure rivolte. Alla base, l'idea che le soluzioni ai molti quesiti che la rivoluzione tecnologica ci pone vanno cercati molto in profondità e nei concetti primi, ben oltre i regolamenti di settore.

1. Premessa

L'intelligenza artificiale (IA) ha un carattere trasformativo². Le ragioni sono interne (sul piano dei meccanismi) ed esterne (sul piano degli effetti) al modo in cui questo artefatto tecnologico funziona e inter-agisce con l'ambiente circostante. Si tratta, infatti, di un insieme di tecniche molto sofisticato che sfrutta correlazioni e nessi inferenziali tra grandi masse di dati per ridare informazioni differentemente strutturate³. Informazioni che a loro volta confluiscano in processi imitativi – e in alcuni casi emulativi – del ragionamento umano (o di una sua parte), per affrontare molti compiti e con esiti sorprendenti⁴. E i risultati cambiano a seconda dell'IA adoperata, perché non c'è solo un sistema di IA, ma ce ne sono molti: deterministici o non deterministici (se si guarda alla possibilità umana di prevedere o meno il comportamento e gli esiti del processo decisionale); simbolici, non-simbolici o sub-simbolici (se ci si riferisce alle capacità artificiali di manipolare e trasformare l'informazione); deduttivi, induttivi, abduttivi e connessionisti (in base al ragionamento per apprendere e prendere decisioni); debole e forte (a seconda del grado di autonomia della macchina nel replicare processi decisionali intelligenti). Questo solo per citare alcuni tipi. È chiaro che ogni tipologia potrà rendere, al sistema digitale in cui funziona, risultati eterogenei, tutti egualmente possibili e utilizzabili.

Se questo è il quadro sul piano informatico e informazionale, non è al contempo facile decifrare fino in fondo le conseguenze sul terreno della biosfera. Con vari livelli di autonomia per realizzare un set di obiettivi, l'IA pervade ogni campo di decisione e azione, dall'industria al commercio, dalla medicina alla robotica, passando per i trasporti e i mercati, azionari e monetari, fino ai sistemi dell'informazione e della comunicazione alla base di molti "metaversi". In sostanza, attraverso contenuti variabili nelle forme di predizioni, raccomandazioni e decisioni, l'IA influenza ogni ambiente, cambiando in profondità le relazioni, personali, professionali e sociali⁵. A ben vedere,

² Così Hight-Level Expert Group on Artificial Intelligence, nominato dalla Commissione europea per la redazione dei principi etici per una IA affidabile, nella prima versione del *Draft* pubblicato il 18 dicembre 2018.

³ Floridi 2014. Si veda anche Romeo 2024, 51-69. In parte, anche Boden 2019, 7.

⁴ Il termine "imitazione" è usato da Turing (1950, 433-460). Il termine "emulazione" viene usato da McCarthy, Minsky, Rochester, Shannon (1955).

⁵ La definizione dell'OCSE è la seguente: "An AI system is a machine-based system capable of influencing the environment by producing an output (predictions, recommendations, or decisions) for a given set of objectives. It uses machine and human-based data and inputs to (i) perceive real and/or virtual environments, (ii) abstract these perceptions into models through analysis in an automated manner (e.g., with machine learning) or manually, and (iii) use model inference to formulate options for outcomes. AI systems are designed to operate with varying levels of autonomy". OECD 2019, 15. Sui problemi definitori dell'intelligenza

non si tratta di un'azione priva di effetti sul piano ontologico, perché sono gli stessi ambienti umani che si strutturano (*enveloping*) per rendere effettiva l'interazione dell'uomo con la macchina. In questo modo, viene in-scritta intimamente la nostra socialità, reale e virtuale, non solo condizionandola, ma ri-ontologizzandola attraverso il linguaggio dell'aritmetica⁶. Per questo motivo, più che di semplice fattore di cambiamento, l'IA è un *fattore meta-morfico* che offre potenzialità e nasconde altrettanti rischi. A preoccupare è in particolare l'iter con cui si giunge agli *outputs*; in molti casi si tratta di un processo poco trasparente e molto oscuro alle stesse capacità di comprensione umana. Basti osservare i processi di decisione che sfruttano le tecniche di *machine learning* (nelle forme più note dell'apprendimento supervisionato, non supervisionato, con rinforzo) e di *deep learning* (ad esempio, reti neurali). Queste impediscono in radice la partecipazione dialettica dell'uomo e, ciononostante, finiscono per condizionarne le scelte, con effetti nella sfera pubblica e privata⁷.

Le maggiori criticità si mostrano quando le abilità statistico-correlazionali degli automi toccano la dimensione pratica delle scienze "poietiche" e cioè di quegli ambiti che, partendo dalla conoscenza e descrizione dei sistemi di riferimento, li disegnano, costituiscono o trasformano secondo propri e originali meccanismi regolatori anche quando le regole che ne discendono sono applicate⁸. Ciò accade, in modo particolare, con il "diritto giurisdizionale" [*Richterrecht*] o dei giudici. Non a caso, in questo ambito l'uso dei sistemi di IA è considerato come altamente rischioso. Ce lo ricorda il Regolamento 2024/1689 del Parlamento Europeo e del Consiglio del 13 giugno 2024 sull'intelligenza artificiale, più noto come *AI Act*. In base al regolamento, i sistemi

artificiale rinvio a Russell, Norvig, 2010, 2, ma anche Floridi 2023, 12.

⁶ Floridi 2020.

⁷ Le questioni giuridiche dell'IA sono molteplici. Un primo profilo di indagine riguarda la regolazione dei sistemi di IA: in questo campo di riflessione siamo più che altro interessati alla progettazione, distribuzione e uso di IA che consentano intervento e sorveglianza umani, robustezza tecnica e sicurezza, tutela della vita privata e governance dei dati, trasparenza, diversità, non discriminazione ed equità, benessere sociale e ambientale e responsabilità (si tratta della visione antropocentrica esposta negli *Orientamenti etici per un'intelligenza artificiale affidabile* elaborati da AI HLEG nel 2019, poi ripresi nel Libro Bianco sull'intelligenza artificiale del 2020 e, peraltro, anche nel considerando n. 27 dell'*AI Act* che ne fornisce una definizione). In questo modo siamo più interessati a conoscere i regimi di responsabilità legati all'IA. Il problema regolatorio è, tuttavia, più ampio e non coinvolge solo gli effetti della tecnologia nella biosfera. Ci sono, infatti, alcune fattispecie che trovano la loro origine, si costituiscono, funzionano o producono effetti direttamente o solamente nell'infosfera. In questo senso oltre alle *tecnologie regolate*, si possono pensare *tecnologie regolanti* nelle forme di artefatti regolativi digitali che si rivolgono a uomini nella realtà, ad avatar in quella virtuale o ad altre tecnologie. Parte di questa riflessione sono in Ricci, Rossetti 2024.

⁸ Floridi 2021, 113-134.

di IA adoperati per l’“amministrazione della giustizia e i processi democratici” sono “[...] 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”⁹. Non è però spiegato cosa vuol dire interpretare il diritto e il fatto e applicare la legge a una serie concreta di *situazioni*.

Rimane sullo sfondo la questione di sapere in che misura l’uso dei sistemi di IA realizzi pienamente l’aspirazione umana alla certezza del diritto nel contesto di una giustizia “digitale” [*digital justice* o *e-justice*]. Qui il termine può ricomprendere varie manifestazioni, dalla *cyberjustice*, alla *automated justice*, passando per l’*algorithmic justice* e la *predictive justice*¹⁰. In ultima istanza, il tema rinvia all’altra questione, ben più ampia, di verificare sino a che punto la certezza del diritto come calcolo sia desiderabile per la vita stessa che è anche diritto¹¹. Sono i punti che vorrei esplorare in questo saggio, partendo da alcune domande decisive per una logica del processo.

2. Metodologia-giuridica critica

È possibile derivare da una proposizione (la norma è una proposizione un po’ particolare) l’unica soluzione *possibile* per un medesimo problema giuridico? Garantendo l’unica soluzione possibile si potrebbe realizzare la certezza del diritto, forse in tempi rapidissimi, con decisioni “uguali” per tutti ed in ogni tempo. Ora, tralascerò l’ipotesi, pure pensabile, di tanti esiti differenti quanti sono i sistemi e le tecniche di IA adoperati. Tenterò, invece, di mostrare come al processo non basti il sapere, né l’informazione che si disperde

⁹ Ne è consapevole il legislatore europeo che ha adottato una regolamentazione di tipo precauzionale con un approccio *risk-based*. L’azione giurisdizionale è considerata attività ad alto rischio ammessa purché non sacrifichi le libertà fondamentali. V. Regolamento (UE) 2024/1689 del Parlamento Europeo e del Consiglio del 13 giugno 2024 che stabilisce regole armonizzate sull’intelligenza artificiale e modifica i regolamenti (CE) n. 300/2008, (UE) n. 167/2013, (UE) n. 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull’intelligenza artificiale). L’Allegato III individua i sistemi ad altro rischio e al punto 8 quelli riguardanti l’“Amministrazione della giustizia e processi democratici”. Il riferimento al regolamento europeo non deve, tuttavia, lasciare intendere che questo lavoro si limiti a considerazioni sull’applicazione giudiziale dell’IA nel contesto europeo. Si intende richiamare l’*Ai Act* solo perché si tratta del primo testo normativo, a carattere sovranazionale, di questo tipo. D’altronde, non tutti i sistemi giuridici hanno adottato l’approccio regolatorio europeo. Si veda Finocchiaro 2024, 107-125.

¹⁰ Alcune delle categorie della giustizia digitale illustrate sono in Palmirani, Sapienza, Bomprezzi 2022, 1-36.

¹¹ Parafraso un’espressione di Lombardi Vallauri 1998, 163.

nelle molte rappresentazioni del soggetto. Per il giudizio, infatti, occorre fare i conti con una conoscenza “finalizzata” della regola e del fatto giuridico.

2.1. *Quaestio iuris*

Stando alle visioni logiciste della scienza giuridica, ogni controversia è valutabile *more geometrico*: da una disposizione deriva una sola soluzione. È invece opposta l’idea di chi considera l’interpretazione e applicazione del diritto come l’esito di un pascaliano *esprit de finesse* per la scelta della soluzione “politicamente” (e cioè valutativamente), *migliore* tra quelle “logicamente” (meglio, logico-giuridicamente) *possibili*. E quest’ultimo il modo di intendere l’attività pratica del *giurista critico* che deve: *i*) riconoscere il diritto vigente; *ii*) interpretare il diritto vigente; *iii*) scegliere la soluzione migliore possibile tra le interpretazioni del diritto vigente. Introdurrò i vari passaggi, seppure semplificandoli, con l’intento di dimostrare il carattere essenzialmente retorico (in senso ampio “politico”) del diritto già nella fase di individuazione della norma da applicare. È la nostra *quaestio iuris*.

i) Se il *diritto* può intendersi come l’insieme di regole (generali e individuali, di condotta e di struttura) che organizza e indirizza le interazioni in un corpo sociale, meno chiaro è il senso del *diritto vigente*. Mi limito alle sue teorie note e nelle versioni volutamente ingenue e semplicistiche: giusnaturalismo, positivismo formalista, positivismo realista. Per il giusnaturalismo, la norma giuridica è quella dotata di valore (*vel, axiological validity*). Secondo il positivismo formalista (*vide* Hans Kelsen) l’esistenza di una norma deriva da un atto di posizione al di là del suo contenuto (*vel, systemic validity*). Infine, il positivismo realista (*vide* Alf Ross) fa dipendere l’esistenza di una norma dal comportamento dei destinatari che o l’hanno osservata (effettività statica) o si può presumere/prevedere che verrà osservata in futuro (effettività dinamica) da un numero sufficiente di utenti (*vel, factual validity*). Insomma, una norma può essere vigente in base al suo *valore*, alla *validità formale* e alla sua *effettività*. E può anche darsi che la vigenza di una norma dipenda contemporaneamente da due di questi significati, indifferentemente combinati tra loro; e se si guarda al comportamento effettivo dei destinatari, l’effettività sarà o statica o dinamica¹². Restando nell’ambito del diritto positivo, le norme potranno quantomeno essere: formulate o non formulate; valide o non valide; effettive o non effettive (staticamente o dinamicamente). È ad esempio possibile che una norma sia non formulata eppure effettiva. Tolto quest’ultimo

¹² Non può trascurarsi che il giuspositivismo contemporaneo così come il formalismo non siano riducibili soltanto ad una visione logicista *stricto sensu*. Si veda, tra i tanti scritti che potrebbero essere menzionati, Scarpelli 1997, ma anche l’antologia a cura di Schiavello, Velluzzi 2005.

caso, in cui si rende difficile l'interpretazione di una disposizione inespressa (si badi, una regola inespressa è comunque esistente... vedi la consuetudine), una volta riconosciuto il diritto vigente nel ventaglio di possibilità, questo va appunto interpretato e applicato. Già a questo livello, sorgono i primi problemi perché resta la domanda: Cos'è il *diritto vigente*? Poniamo che si propenda per la validità fattuale (ho invece escluso il caso, molto complicato, di una validità che derivi dal valore di una norma). A questo punto, andrebbe quantomeno fissata una misura sociologicamente rilevante per stabilire se una certa norma sia sufficientemente rispettata o meno, sino all'ipotesi estrema di una norma osservata senza eccezione alcuna. In tutti questi casi, il *riconoscimento* del diritto vigente presuppone una scelta particolare e si fa, in tal modo, politica.

ii) Ora, supponiamo di avere un'idea chiara di cosa sia il diritto vigente. Un sistema di intelligenza artificiale non esiterebbe a considerare vigente solo e soltanto l'insieme delle norme formulate per iscritto, valide perché approvate da un'autorità competente ed effettive staticamente cioè al momento della loro applicazione. Si provi a sperimentarlo, rivolgendosi ad un modello conversazionale di tipo generativo¹³. La valutazione politica non è,

¹³ I modelli generativi sono definiti dall'art. 3, n. 63, dell'*Ai Act*. In realtà, la risposta ipotizzata non è affatto di fantasia. Ho banalmente chiesto a ChatGpt, sistema Large Language Model particolarmente evoluto, "Che cos'è il diritto vigente?" La risposta è stata la seguente: "Il diritto vigente è l'insieme delle norme giuridiche attualmente in vigore in un determinato ordinamento giuridico e applicabili ai rapporti sociali. Esso comprende leggi, regolamenti, decreti e altre fonti del diritto che non siano state abrogate o dichiarate incostituzionali. In altre parole, il diritto vigente è l'insieme delle regole che devono essere rispettate in un dato momento e luogo, stabilite dallo Stato o da altre autorità competenti.

Caratteristiche del diritto vigente:

Efficacia attuale → Si applica nel momento presente, a differenza del diritto abrogato o storico.

Applicabilità → Deve essere rispettato da cittadini, imprese e istituzioni.

Gerarchia delle fonti → È regolato da un ordine gerarchico (Costituzione, leggi, regolamenti, ecc.).

Possibilità di modifica → Può essere cambiato, abrogato o sostituito da nuove norme.

Un esempio concreto di diritto vigente è il Codice Civile italiano, che contiene norme applicabili oggi nei rapporti tra privati, oppure la Costituzione Italiana, che stabilisce i principi fondamentali del Paese" [Data di consultazione, 13/02/2025].

Ora, non mi interessa approfondire ulteriormente il tema con il mio interlocutore conversazionale; probabilmente potrebbe dare risposte sorprendenti e correggere il tiro, fornendo repliche più complete rispetto a quelle che potrebbe rendere un operatore umano. Né mi interessa sapere se altri applicativi – magari specializzati nelle questioni giuridiche – possano fornire risposte diverse allo stesso quesito. In questo articolo, intendo mostrare come un sistema di intelligenza artificiale sia in grado di opzionare solo uno degli esiti possibili. E tale risposta è, per questa ragione, incompiuta. In ogni caso, il risultato appena descritto potrebbe essere differente da quello fornito dallo stesso sistema interrogato in un altro tempo, o da altri sistemi differenti. Il che, in astratto, porterebbe a ritenere che anche gli applicativi LLM presuppongono (e richiedono) una scelta tra le tante possibili.

ancora una volta, eclissata. Invero, dopo avere scovato il diritto vigente, ci si accorge subito che questo è in sé lacunoso e cioè continua a non offrire una soluzione univoca, ma ne può offrire molte (lacune *intra legem*, frutto del dubbio sul significato di una norma che è validamente posta); a volte, può non offrirne alcuna (lacune *praeter legem*, a causa di una precisa assenza di una norma). Così, anche una disposizione generalissima, ma altrettanto chiara come quella espressa dall'art. 1 della Costituzione italiana ("*L'Italia è una Repubblica democratica, fondata sul lavoro*") può porre altrettanti problemi di senso e significato.

Ebbene, se una norma manca del tutto, la *lacuna* può essere colmata dall'ordinamento giuridico (l'insieme delle norme che pongono gli stessi problemi di riconoscimento visti prima); altre volte sarà il giudice a integrare le norme. Ma quando il giudice risolve le lacune, finisce in fin dei conti per applicare solo *indirettamente* le norme. Pensiamo ai procedimenti della logica giuridica e cioè operazioni sintattiche (analogia, argomento *a contrario*, presunzioni, finzioni, costruzione dogmatica ...) di derivazione di una norma da altre non direttamente riferibili al caso sottoposto a giudizio. Queste operazioni non sono affatto rigorosamente logiche e sono il più delle volte teleologicamente orientate. E dunque, anche in assenza di lacune *praeter legem* (nel caso in cui immaginassimo un ordinamento completo senza vuoti ovvero stabilendo l'intervento *ubiquo* del legislatore ogni volta che si rendesse necessario) il diritto vigente continua a fornire molte soluzioni possibili. Due almeno i motivi. Innanzitutto, per la vaghezza e/o indeterminatezza e/o ambiguità del linguaggio naturale i cui termini non sono perfettamente definiti nel senso della logica formale¹⁴; in secondo luogo, perché esistono molti metodi di interpretazione i cui risultati sono altrettanto diversi. Sono illuminanti alcune considerazioni di L. Lombardi Vallauri nel suo *Corso di filosofia del diritto* (1981) che perviene ad ipotizzare un ventaglio di almeno 144 usi possibili di ogni proposizione normativa¹⁵. E non c'è un criterio oggettivo o universalmente accettato su quale metodo adottare. Non sarebbe sufficiente nemmeno una norma sull'interpretazione (nell'ordinamento italiano è l'art. 12 delle disposizioni premesse al codice civile del 1942), perché anche questa va a sua volta interpretata con metodi non necessariamente coincidenti con quelli che essa prevede. Ancora una volta, anche l'interpretazione subisce gli effetti della volontà particolare di una scelta.

¹⁴ Parlo di vaghezza e/o indeterminatezza e/o ambiguità in senso ampio, ma i termini indicano fenomeni differenti. Per un maggiore approfondimento, rinvio a Luzzati 1990, Lombardi Vallauri 1998, Velluzzi 2006.

¹⁵ Lombardi Vallauri 1981, 55-58 e ancora Lombardi Vallauri 1982, 303-315.

iii) In conclusione, molte sono le incertezze. La prima riguarda il riconoscimento della norma vigente: nella *pratica* si danno più norme possibili. La seconda riguarda l'interpretazione e i procedimenti della logica giuridica: nella *pratica*, pur avendo scelto la norma possibile, essa ha più di un significato possibile. Ed allora, se da un lato il diritto vigente è l'insieme di soluzioni che devono cadere in un ambito circoscritto, vuoi ideale-universale, vuoi positivo-particolare, dall'altro non si sa in quale punto di questo ambito deve cadere la decisione. E però, l'incertezza sugli esiti possibili non sottrae il giudice dall'obbligo (che è anche una necessità) di decidere (è il divieto di *non liquet*); eppure non può decidere sulla base del solo diritto positivo vigente.

In tal senso, il diritto applicato alla realtà con la metodologia del giurista altro non è che la soluzione migliore tra quelle logico-giuridicamente possibili entro il diritto vigente individuato dal soggetto. Il diritto positivo, però, non fornisce alcun criterio oggettivo per scegliere. Ciò vuol dire che la soluzione è frutto di scelta, non certamente del capriccio, ma comunque l'esito di valutazioni specifiche. È il dramma del giudizio.

2.2. *Quaestio facti*

Accanto all'interpretazione e applicazione del diritto e della legge, vi è l'accertamento del fatto. L'*Ai Act* si riferisce genericamente all'"interpretazione del fatto". Non è tra l'altro semplice distinguere la "questione di diritto" dalla "questione di fatto" in ambito giurisprudenziale¹⁶. D'altra parte, non tutti i fatti sono rilevanti per il processo. Ad esempio, l'art. 118, comma 1, delle "Disposizioni per l'attuazione del codice di procedura civile" italiano impone al giudice di esporre i fatti rilevanti per la causa. E già a questo livello è richiesto un *giudizio di rilevanza* che ancora dipende dal soggetto. Ma tale giudizio va oltre la selezione dei fatti rilevanti tra quelli possibili, e riguarda anche la prova dei fatti. Ecco perché si dovrebbe parlare di "giudizio sui fatti" più che di "prova dei fatti"¹⁷. Il processo, in effetti, si trova innanzitutto a confrontarsi con fatti che sono già accaduti. Ma "provare" vuole dire avere esperienza di qualcosa. In tal senso, non potendo avere esperienza di un fatto già accaduto non resta che la valutazione del fatto. Ma quali sono i fatti rilevanti per il diritto?

Innanzitutto, i fatti che contano per il diritto non sono soltanto "fatti empirici" [*external facts*]. Il più delle volte sono "fatti determinabili valutativamente" [*facts evaluatively determined*]. Qui il confine tra questione di diritto e di fatto è molto labile: Cosa significa "danno morale", "minaccia", "pubblico pudore", comportamento conforme a "buona fede" o la "sofferenza sogget-

¹⁶ Perelman 1961, 268-278.

¹⁷ A rilevarlo è Incampo 2016, 176-177.

tiva” nei casi di risarcimento del danno biologico macropermanente? In altri casi, i fatti possono identificarsi con la loro documentalità (si pensi alle scritture contabili per i reati di “falso in bilancio”) oppure, più in generale, a “fatti istituzionali” [*institutional facts*]¹⁸ qualificati come tali da norme (un tipo di contratto, ad esempio, è un fatto che dipende dalla qualificazione data dal diritto stesso). Qui l’accertamento del fatto è il più delle volte il frutto dell’interpretazione ben oltre la letteralità del documento (basti guardare ai casi in cui il fatto determinante è la stessa intenzione dei contraenti secondo la regola fissata dall’art. 1362 del codice civile italiano). Alcuni fatti rilevanti per il diritto nemmeno ci sono (condotte omissive o le posizioni di pericolo astratto), per non parlare della prova dell’elemento soggettivo (colpa, dolo, dolo eventuale...) e cioè fatti interni alla base del giudizio di imputazione. Infine, per il processo possono contare anche fatti che devono ancora accadere, come quelli alla base dei giudizi prognostici in materia cautelare. Per queste ragioni, anche l’accertamento del fatto non eclissa il giudizio sul fatto.

E gli artefatti tecnologici? Questi possono entrare nel processo come fatto da provare o come mezzo di prova. Nel primo caso, rientrano molte fattispecie giuridiche che derivano da interazioni informazionali tra soggetti (reali o virtuali) che richiedono strumenti tecnologici per essere realizzati. A questo livello, le domande che la sfida tecnologica pone al diritto sono inedite: Cosa potrà significare “ingiusto profitto” se riferito alle monete virtuali (*bitcoin*)? E il “falso” per i *token* di beni digitali (si tratta di NFT, *non-fungible token*)?¹⁹ Per non parlare del “danno ingiusto” nei casi di responsabilità extracontrattuale dell’IA. Insomma, i fatti qui sono più che altro il frutto di operazioni algoritmiche e, in ogni caso, della loro percezione da parte del soggetto. In questo senso, somigliano molto ai “fatti” di cui ci parla G.E.M. Anscombe in *On Brute Facts* (1958). L’artefatto tecnologico può essere considerato come un fatto ideato dall’uomo e sviluppato secondo il linguaggio della programmazione i cui effetti si producono anche nella biosfera²⁰. Il valore di mezzo di prova, invece, è riconosciuto da certi istituti del codice di procedura penale italiano. Nell’ordinamento italiano era già accaduto, nel 2008 (ad opera del D.lgs. 48/2008), per l’ispezione informatica (art. 244, comma 2), la perquisizione informatica (art. 247, comma 1bis), il sequestro di dati informatici presso i fornitori di servizi informatici, telematici e di telecomunicazioni (art. 254 bis) per arrivare all’uso di strumenti avanzati come il captatore informatico (art. 267, comma 2bis, c.p.p. introdotto dal D.Lgs. 216/2017) che consente di

¹⁸ Si tratta di un concetto riconducibile icasticamente a Searle 1996.

¹⁹ Alcuni di questi interrogativi sono in Incampo, 2025.

²⁰ Anscombe 1958, 69-72. L’idea di collocare gli artefatti digitali in un’istituzione socio-digitale, riprendendo le tesi di Anna Beckers e Gunther Teubner, è di Loddo 2024.

controllare un dispositivo da remoto a fini investigativi. Infine, alcuni strumenti mantengono un carattere ibrido, sono cioè mezzi di prova che al contempo offrono all'operatore del diritto e al giudice informazioni direttamente utilizzabili nel processo. È il caso dei *tools* di IA per il riconoscimento facciale, capaci di dare vita ad indagini scientificamente avanzate oltrepassando l'intuizione degli operatori.

Ora, anche in questi casi mi sembra non si possa escludere la valutazione del giudice. Senza entrare nel dibattito sul carattere tipico e atipico della prova algoritmica come strumento probatorio scientifico-tecnico, la natura riproduttiva di tale mezzo è allo stato affermata da gran parte della giurisprudenza. Con la sentenza n. 23090 del 2020, la Suprema Corte di Cassazione ha tra l'altro avuto modo di evidenziare che "L'individuazione di un soggetto – sia personale che fotografica – da parte della polizia giudiziaria in fase di indagini è una manifestazione riproduttiva di una percezione visiva e rappresenta una specie del più generale concetto di dichiarazione, sicché la sua forza probatoria non discende dalle modalità formali del riconoscimento, bensì dal valore della dichiarazione confermativa assunta in sede di deposizione testimoniale". D'altro canto, se è vero che strumenti di riconoscimento facciale sono capaci di registrare un fatto per iterarlo nel tempo, è altrettanto vero che a ripetersi è più che altro la percezione di esso. Per questo motivo, stando alla Cassazione, servono ancora i testimoni.

Ecco che le predizioni algoritmiche alla base della prova algoritmica non superano la differenza tra provare un fatto e il giudizio sul fatto viste prima. Il fatto passato e quello futuro sono percepiti direttamente e proiettati nel presente al momento dell'accertamento e si identificano con gli *outputs* resi dal sistema. Per richiamare ancora l'*automatic facial recognition system*, esso calcola solo la probabilità di *matching* tra l'immagine di una persona e altre inserite nel *data set* del sistema. Così il riconoscimento facciale potenziato dall'uso di IA sembra restituire viva esperienza dell'accaduto. Ma ciò è sufficiente per escludere il giudizio? Certamente no. Invero, i fatti rilevanti per il diritto e per il giudizio sono fatti istituzionali, cioè, fatti che subiscono le trasformazioni poietiche e finalizzate del diritto e dell'intenzionalità del soggetto con cui si confrontano. In questo senso, non contano solo i fatti bruti. Una rigorosa logica della prova deve, ad esempio, tenere conto dell'imputabilità, della causalità e dell'elemento psicologico sulla scorta delle *evidenze* che emergono nel giudizio. D'altronde, nemmeno la visione diretta delle azioni in un gioco sportivo riesce a conferire certezza alle valutazioni dell'arbitro, perché non è esclusa la percezione dell'informazione desunta dal fatto. Dunque, anche dinanzi ad un sistema perfettissimo, la prova di un fatto resta sempre un giudizio che si trova di fronte un ventaglio di possibilità: non una,

ma molte soluzioni. Più che il fatto, rileva il suo significato che va costruito e ricostruito da parte del soggetto.

In conclusione, nell'interpretare la regola, nel ricostruire il fatto, nell'individuare i moventi e l'elemento psicologico nonché la personalità delle parti e nell'accertare i nessi di causalità, il dubbio permane. È questa l'idea, tra l'altro, sottesa al principio del "libero convincimento" e dell'"oltre ogni ragionevole dubbio" come misura d'ordine della prova processuale²¹. Non si può, però, negare che alcuni fatti richiedano strumenti tecnologicamente avanzati per verificarli (si pensi al riconoscimento del *deep fake*), ma si tratta pur sempre del risultato, seppure scientifico, del calcolo di probabilità. Il rischio è quello di confondere i mezzi con i fini, gli strumenti con gli esiti sempre interpretabili dal soggetto.

3. Logica incompiutezza dei sistemi algoritmici

Ho provato ad esemplificare le principali questioni che rendono necessaria, nella pratica giudiziale, una metodologia giuridica-critica. Ciò mi è servito per mostrare come il processo non sia riducibile ad operazioni sintattiche alla base di correlazioni e inferenze, né tantomeno agli esiti probabilistici delle predizioni. In altri termini, bisogna "conoscere il diritto" prima di vederne le sue evoluzioni²². A questo punto c'è da vedere se l'uso di tecniche di analisi e programmazione computazionale altamente sofisticate possa trasformare una metodologia critica (come tale insidiata da margini di discrezionalità e di indeterminatezza) in una logica pre-critica, calcolabile, parametrata a criteri oggettivi di valutazione e assolutamente certa. Non è un passaggio così immediato. Il punto forse un po' trascurato è che l'intelligenza artificiale come scienza ultramoderna si rapporta ad un oggetto (il diritto, appunto) nato già in epoca arcaica, e all'esperienza che di esso ha inevitabilmente il soggetto nei tanti casi unici e irripetibili della vita. Alcune risposte si possono rintracciare osservando la natura degli *outputs* resi dall'IA. Gli esiti algoritmici, infatti, somigliano molto al calcolo di probabilità inferenziale basato su massime di esperienza o su leggi scientifiche. Si tratta di saperi decisivi per il giudizio, eppure entrambi i criteri non riescono ad accertare la verità dei fatti che entrano nel processo.

Le massime di esperienza sono il prodotto di "generalizzazioni di luoghi comuni" [*commonplace generalizations*]. Anche le opinioni basate sui precedenti giurisprudenziali possono contribuire a fissare *standard* valutativi di alcune condotte normativamente previste (ne è un esempio la clausola

²¹ Nobili 2003, 72.

²² Prendo in prestito l'espressione di Schiavello 2023.

del “buon padre di famiglia” la cui definizione dipende strettamente dalle elaborazioni della giurisprudenza). Affinché la massima possa valere come mezzo di prova occorre, tuttavia, stabilirne il grado di attendibilità, plausibilità, probabilità al fine di superare ogni ipotesi ricostruttiva antagonista²³. E qui il primo problema che riguarda la verifica della fonte della massima d’esperienza. Se la fonte è riconducibile a sistemi non deterministici che operano come scatole nere [*black-box*], la fonte non sarà pienamente verificabile. Gli esiti algoritmici, inoltre, possono soffrire problemi di generalizzazione in eccesso o in difetto (si tratta, rispettivamente, dell’*overfitting* e dell’*underfitting*); dall’altro lato, pur superando questi problemi attraverso *training* specifici del modello di addestramento, è la generalizzazione in sé che impedisce l’univocità dei risultati, perché ogni generalizzazione ammette delle eccezioni. È, infatti, motivo di errore la confusione tra il piano generale e quello particolare. Il rapporto causale si istituisce sempre fra *eventi singoli* e sarebbe fallace pretendere che la spiegazione di un evento storico si possa ottenere sulla scorta di una legge generale: la *logica delle asserzioni singolari* differisce dalla *logica delle asserzioni generali*. Così, la spiegazione causale di un *evento singolo* già accaduto non può correttamente fondarsi nemmeno se ricondotto ad una *legge scientifica generale* (tali sono anche le correlazioni statistiche). Né si possono sottovalutare le anomalie rispetto ai modelli conoscitivi che possono verificarsi nelle serie causali. A mancare è in definitiva la *simultaneità* tra l’accadimento del fatto e la sua conoscenza. Sicché, il fatto non può mai essere accertato *in sé e per sé*. Anche in questo campo di riflessione, i risultati della statistica (comprensiva dell’apporto neuroscientifico) sono sempre insoddisfacenti: la verifica scientifica di un “antecedente” nei sistemi complessi porta in genere a stabilire una condizione soltanto *sufficiente* e non anche *necessaria* di un accadimento²⁴.

Restano pure insufficienti i risultati stocastici dei modelli di apprendimento automatico e profondo, ma anche *general purpose*. Il rischio è quello di far entrare nel giudizio il concetto di *causabilità* piuttosto che di *causa specifica* di un fatto particolare. I problemi, tuttavia, sono almeno due: il primo incide su consolidati paradigmi epistemologici, in particolare sul giudizio penale ove vige il “principio di tassatività e determinatezza della fattispecie”; il secondo rimane l’insufficienza del calcolo di probabilità nell’effettivo accertamento causale tra fatto ed evento. Un rischio che sarebbe molto concreto se ci si abbandonasse alle acrobazie, seppure scientifiche, di macchine evolute. Con l’effetto di perdersi e perdere il senso del diritto e delle relazioni che fonda. Il punto è proprio questo: la probabilità non è mai oggettiva, ma è sempre

²³ Incampo 2016, 193

²⁴ Incampo 2017, 69-72.

frutto di risultati che si avvicinano alla verità e, dunque, non dà mai l'unica soluzione possibile. Va, tra l'altro, osservato che un calcolo soddisfacente di probabilità non può svilupparsi su casi che si ritengano unici. In ognuno di questi casi, la probabilità, nel lancio di un dado, sarebbe sempre quella iniziale di uno a sei. E magari, in situazioni eccezionali, il dado resta fermo su uno spigolo dei lati. In secondo luogo, il processo si confronta *ex post* con casi unici e, dunque, sempre diversi. Anche la serialità delle cause, a cui pure si vorrebbe limitare l'uso dell'IA, è nella pratica una chimera.

C'è infine un ultimo aspetto. L'incompletezza dei sistemi di IA oltre che dal punto di vista di una metodologia critica, è anche sul piano della logica. A metà degli anni '30, Alan Turing stava lavorando al problema della decisione in logica [*Entscheidungsproblem*], posto nel 1928 dal pensiero finitario di David Hilbert i cui esiti vennero codificati dai noti teoremi di incompletezza di Kurt Gödel²⁵. La questione era se all'interno di un sistema assiomatico del primo ordine si potesse decidere attraverso una procedura meccanica, per qualsiasi formula di quell'ordine, e se tale formula fosse un teorema del primo ordine oppure no. Già i teoremi di Gödel avevano dimostrato l'impossibilità dei sistemi assiomatici di dedurre tutte le formule o proposizioni vere che vi appartengono. Un sistema di formalizzazione dell'aritmetica elementare è in grado di comprendere insieme, in base a regole di trasformazione del sistema, sia proposizioni aritmetiche, sia "proposizioni formalmente indecidibili"²⁶. C'è in estrema sintesi un livello di comprensione che sfugge a qualsiasi formalizzazione e il cui risultato è *indecidibile* dal punto di vista del calcolo. La potenza di calcolo di una macchina universale, dunque, non implica anche la capacità di fondare tutte le proprietà contenute nei propri assiomi. E, dunque, i sistemi logici non sono capaci di fondarsi da soli, ma hanno bisogno di un *giudizio di con-valida* esterno ad essi.

Non è tuttavia escluso che si possa considerare la calcolabilità algoritmica come valore metapositivo assoluto fondante un sistema giuridico-positivo. È una possibilità ideologicamente sostenibile dal momento che l'informatica è andata oltre i sistemi giuridici esperti per progettarne di nuovi che generano essi stessi il modello di conoscenza sulla base della statistica e della correlazione. Il valore potrebbe essere così formulato: "È norma l'esito e solo l'esito del calcolo". È un'ipotesi radicale che potrebbe rispondere ad una precisa scelta di politica in cui il giudice delega l'autonomia funzionale dell'algoritmo senza alcuna discussione²⁷. Sarebbe come affidare la giustizia

²⁵ Turing 1937, 230-265, Hilbert 1926, 161-190, Hilbert 1923, 151-157, Gödel 1931, 173-198.

²⁶ Sulla rilevanza dei teoremi per una fondazione dei sistemi giuridici, si veda Incampo 2016, 221-223 e prima Incampo 1996, 330-351. Per alcune applicazioni dei teoremi di Gödel alla costruzione dei sistemi giuridici, Agrusti 2013, 1-17.

²⁷ Va comunque rilevato che una tale pratica sarebbe oggi vietata ai sensi dell'art. 14

al lancio dei dadi (o al lancio dei dati). Corrisponderebbe, d'altronde, ad una procedura poco dispendiosa e perfettamente imparziale. Potremmo chiamarlo l'algoritmo di Bridoye, il protagonista di alcuni racconti di Rabelais, che amava risolvere le dispute lanciando i dadi e poi motivare assecondando le argomentazioni della parte baciata dalla sorte. Si potrebbe, d'altro canto, immaginare un algoritmo compiacente con logiche particolari stabilite *ex ante* dal progettista o dallo sviluppatore. Ad esempio, il sistema potrebbe essere programmato in modo che gli esiti siano sempre favorevoli alle classi meno abbienti o, all'opposto, alle grandi *companies* del tech. Sarebbe un algoritmo meno imparziale eppure potrebbe, dal punto di vista logico, adottare una decisione assolutamente coerente con il sistema metapositivo *particolare* scelto e le previsioni sarebbero addirittura più facilmente calcolabili. Non credo, tuttavia, che i problemi siano anche da questo punto di vista risolti. Senza considerare le questioni di legittimità della fonte (non normativa) usata dal sistema, sarà sempre e comunque necessario fissare una misura di reddito per distinguere le classi e così anche un criterio quantitativo non potrebbe prescindere da una valutazione perché non è mai esatto; ed ancora, occorrerà individuare il preciso comparto di investimento tra i tanti ed eterogenei oggetti sociali indicati in uno statuto societario; e così persino il lancio dei dadi potrebbe ridare più di una risposta²⁸. In fin dei conti, c'è sempre un segmento della decisione che si sottrae alla formalizzazione.

Resta forse l'ultimo tassello. Nulla impedisce che si arrivi all'automazione con la conseguente delegazione piena all'algoritmo del processo decisionale. È il paradigma dell'*automated justice*. Il rischio, tuttavia, è che una certa politica del diritto snaturi il carattere politico (critico) del diritto, aprendo la strada al "nichilismo giuridico". Resta dunque un'ultima domanda che si collega ad una genealogia della giuridicità: È davvero giustificabile una giustizia digitale automatizzata? Quali i rischi? Sono i passaggi conclusivi.

4. Chi sa valutare il rischio?

L'uso dei sistemi di intelligenza artificiale nella giustizia è vario. Coinvolge l'attività di controllo, gestione e pianificazione del *workflow* delle cause;

dell'*Ai Act* che, al paragrafo 4, individua misure che il *deployer* deve considerare in modo che alle persone fisiche sia consentita la sorveglianza umana (ove mai concretamente ed effettivamente realizzabile, come si vedrà, sinteticamente, nel paragrafo successivo). È chiaro che si tratta pur sempre di un atto a contenuto normativo valido in un'area specifica e, peraltro, potenzialmente rivedibile. Invero, l'obiettivo del paragrafo è proprio quello di rilevare l'incompletezza, da un punto di vista logico, dei sistemi di IA quando usati in quello specifico ambiente (oggi, si usa il termine, ecosistema) che è rappresentato dallo spazio processuale, al di là del contenuto di specifiche regolamentazioni.

²⁸ Per riprendere un'immagine di Incampo 2016, 213.

l'anonimizzazione e pseudonimizzazione delle sentenze; la consultazione ragionata delle decisioni giudiziali; l'estrazione di dati utili per il contrasto alla criminalità organizzata che proprio sullo scambio coordinato di informazioni si basa; il supporto alla decisione, persino sulla valutazione della pericolosità sociale del soggetto. Con la crescente sofisticazione dei sistemi di IA non è difficile immaginare la possibilità (ancora remota) di automatizzare ogni processo decisionale per assecondare l'aspettativa di maggiore certezza e prevedibilità degli esiti. Il punto è che gli algoritmi alla base delle IA oltre a non dare risultati oggettivi, non sono affatto immuni da *bias* cognitivi, fattori discriminanti ed iniquità. Così possono influire negativamente sulla sicurezza e sui diritti fondamentali e, per tale ragione, vengono considerati sistemi ad "alto rischio". Tali sono anche quei sistemi adoperati nell'ambito giudiziale.

Guardando alla risposta dell'Unione europea, in base all'*Ai Act*, la valutazione dei rischi associati a tali sistemi avviene prima del loro uso e segue anche durante il relativo ciclo di vita. I sistemi di IA classificati come ad "alto rischio" sono, ad esempio, sottoposti ad una valutazione dell'impatto sui diritti fondamentali, devono usare dati di qualità e garantire adeguate misure di sicurezza. Ad ogni buon conto, per mitigare i rischi occorre identificarli. Ma cosa vuol dire esattamente che un sistema è rischioso se usato da un'autorità giudiziaria o 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? La particolarità dell'uso giudiziale dei sistemi di IA sta tutto in questo delicato passaggio.

Si è visto come la valutazione sulla violazione dei diritti, così come la relativa protezione, non può che avvenire in un giudizio reso necessariamente *ex post* rispetto al caso da affrontare. Anche il giudizio cautelare è sempre reso dopo una situazione giuridica controversa. Il cortocircuito è alquanto evidente. I sistemi di IA per l'uso forense, in quanto *tools* ad alto rischio, sono in teoria valutati *ex ante* attraverso il rilascio di certificazioni che ne attestino la conformità. Nella pratica, tuttavia, possono dirsi rischiosi solo quando si confrontano con i tanti fatti che accadono. Dunque, una valutazione *ex ante* dei rischi che coinvolgano i diritti fondamentali è insufficiente se non si raffronta con l'uso giudiziale dei sistemi. E tale valutazione non può che essere affidata alle Corti ed all'attività essenzialmente dialogica e cooperativa che in esse viene espletata. Ma qui un primo paradosso: se la valutazione è concretamente possibile solo dopo l'accadimento di un fatto, la stessa sarà soggettiva e ancora legata alla scelta tra le tante soluzioni possibili sulla conformità o meno del sistema. E poi, con quali mezzi si può valutare il rischio? Non potrà negarsi, in questo campo, un ruolo rilevante della tecnica, ma ancora un secondo paradosso: se tale valutazione avviene con il supporto

di un'IA, è essa stessa il risultato di un'attività rischiosa. Sono d'altronde oscuri i processi di ideazione, progettazione e realizzazione di questi sistemi. La sensazione è che la dimensione pubblica della giustizia ceda il passo al controllo di poche società che dispongono di un certo sapere tecnologico e che addirittura auto-certificano l'affidabilità dei propri ritrovati (si tratta delle certificazioni *self-risk assesment*) senza incontrare un valido ostacolo quando sono utilizzati. È così in gioco la stessa legittimità della valutazione. Come uscirne?

L'intelligenza artificiale è certamente un "fatto sociale totale", per richiamare una nota categoria di Marcel Mauss²⁹. Ridisegna significativamente spazi (anche simbolici) come lo sono quelli legati alla giustizia e, soprattutto, plasma il comportamento degli operatori di un sistema come quello giurisdizionale. In questo modo si può dire che il "codice è la legge" [*Code is law*]³⁰. È, in ogni caso, interessante capire quali sono i meccanismi che trasferiscono la normatività all'algoritmo. L'algoritmo non ha in sé alcun senso "normativo", semmai "regolativo". È, cioè, in grado comunque di regolare comportamenti. Dall'altro lato, il digitale cambia in profondità la stessa elaborazione del diritto, poiché anche l'informazione (e l'informatica che l'utilizza) sono capaci di produrre senso. La normatività, tuttavia, è concepibile solo dal punto di vista tipicamente umano di *procedere*. In questo senso, si può ritenere che l'IA cambi innanzitutto (seppure non esclusivamente) le modalità di azione dell'uomo. Il che vuol dire, per il diritto giurisdizionale, fare metodologia. Se siamo interessati al modo di procedere, l'incomunicabilità tra il piano poietico del diritto e quello essenzialmente mimetico della tecnologia può trovare un'utile sintesi a livello procedurale. Le *procedure* ci dicono come certe cose vanno fatte, prima ancora di giungere ai molteplici contenuti che si possono avere. D'altronde, è molto difficile pensare ad un programma che consenta all'automa giuridico di assegnare un certo significato alle parole o ai fatti, ma sarà invece possibile ideare un sistema che sappia porre le giuste domande o far sorgere dubbi legittimi, per recuperare la visione simbiotica di J.C.R. Licklider³¹. Al contempo, le procedure valgono come regole di legittimazione e convalida in un dato ambiente. La fondazione della normatività resta così ancorata ad una certa "forma di vita" [*Lebensform*] ed ai giudizi di convalida che poggiano sul modo tipicamente umano di stare al mondo³².

²⁹ Mauss 2002. L'accostamento tra la trasformazione digitale e la categoria maussiana è di Garapon, Lassègue 2021, 79.

³⁰ Il riferimento è all'articolo di Lessig 2000.

³¹ Licklider 1960, 4-11.

³² Wittgenstein 1995, part. 25. Sul rapporto tra Turing e Wittgenstein cfr. Floyd 2016, 7-89 ed ancora Floyd 2018, 59-92. Spunti interessanti si rinvengono anche in De Anna 2019, 125-142.

Non si può dunque escludere la necessità che sia l'uomo a conferire carattere normativo all'algoritmo e ai dati di cui si serve. Il diritto e il processo hanno, invero, natura *androritmica* e cioè sono e funzionano diversamente dalla scrittura di un algoritmo e dalla logica del calcolo³³. Per questa ragione i rischi possono valutarsi nel processo e non fuori di esso, attraverso una stretta inter-azione uomo-macchina che sappia cogliere il senso ed il significato delle varie *situazioni* percepite nel giudizio. Non è comunque riducibile ad un mero controllo (è il tratto tipico di una visione *human in the loop*), perché questo si rende impossibile per la stessa capacità umana di comprendere pienamente i processi di decisione robotica. L'inter-azione richiama, invece, un'idea associativa tra l'uomo e la macchina che consenta alla decisione resa con il supporto di modelli simbolici e computabili di essere con-validata in uno spazio giuridico umano. La cooperazione uomo-macchina avviene ancora sul piano della metodologia. Con l'associazione si combina l'euristica umana con processi computazionali, affidando un ruolo decisivo all'argomentazione giuridica³⁴. D'altra parte, i dati non possono spiegare altri dati³⁵. L'argomentazione diventa da questo punto di vista fondamentale, così come è centrale il diritto delle parti di ricevere una spiegazione giuridicamente valida. Una sentenza immotivata, frutto soltanto del capriccio o della correlazione statistica indifferente al caso specifico resta un atto privo di senso. Anche la decisione presa con l'ausilio del sistema di IA non può trascurare tale condizione. Questo perché la sentenza del giudice, pur adoperando meccanismi informativi derivanti dai ragionamenti stocastici profondi, non può fare a meno di formalizzarsi. E si formalizza osservando "condizioni di felicità" [*felicity conditions*] procedurali in un *habitat* necessariamente umano in cui sono assicurati stabili ed effettivi canali di comunicazione. D'altronde, solo una buona argomentazione si mostra equa, opportuna e socialmente utile³⁶.

³³ Androritmo è il neologismo coniato da Leonhard 2016. Secondo Leonhard "Androrithms must take center stage with technology if we are to remain a society that is concerned with human flourishing".

³⁴ Leonhard 2016, 6: "[...] men are noisy, narrow-band devices, but their nervous systems have very many parallel and simultaneously active channels. Relative to men, computing machines are very fast and very accurate, but they are constrained to perform only one or a few elementary operations at a time. Men are flexible, capable of 'programming themselves contingently' on the basis of newly received information. Computing machines are single-minded, constrained by their 'pre-programming.' Men naturally speak redundant languages organized around unitary objects and coherent actions and employing 20 to 60 elementary symbols. Computers 'naturally' speak nonredundant languages, usually with only two elementary symbols and no inherent appreciation either of unitary objects or of coherent actions".

³⁵ D'Acquisto 2022, 12, 23.

³⁶ Perelman 1961, 137, 165.

Guardando proprio alla motivazione quale condizione necessaria di una decisione valida, essa sarà nei contenuti consistente e cioè l'argomentazione adottata e la giustificazione resa dal giudicante deve risultare persuasiva in tutti gli elementi in fatto e in diritto. Si tratta di una costruzione positiva della motivazione resa con l'ausilio dell'IA che deve però precedere la decisione e non seguire essa. L'argomentazione della sentenza, dunque, deve aversi simultaneamente con metodo problematico e deve fare emergere la consonanza tra i vari elementi della decisione. L'impegno del giudice dinanzi ai propri dubbi è quello di pensare sin da subito il problema e razionalizzare la logica con il peso emotivo e morale di dover produrre una giustificazione valida ed efficace che tenga conto del destinatario. Al giudice spetta il compito, non semplice, di esporre tutti gli elementi necessari per una comprensione effettiva e razionale del *reasoning*³⁷. Dovrà quantomeno far riferimento agli *input* inseriti nella macchina, il punteggio di ciascuno, i rispettivi criteri di valutazione e, infine, la logica intrinseca di funzionamento, assicurando ai destinatari l'affidabilità tecnica della macchina. E dovrà operare con un linguaggio che sia chiaro mediante passaggi argomentativi significativi per garantire la piena interiorizzazione della decisione finale da parte del destinatario. Perché solo assicurando una chiara comprensione del *decisum* è ipotizzabile la partecipazione dialettica del destinatario alla decisione finale e persino fornendo utili elementi di contestazione. Se il giudice dialogasse adeguatamente con le parti in causa, queste sarebbero messe nelle condizioni di contestare sin da subito l'uso di alcune tecnologie. Ad esempio, potrebbero convincere il giudice dell'inaffidabilità della tecnologia processuale e della logica sottesa, escludendo nel processo l'uso di tecnologie in sé oscure come quelle di *machine learning* e di *deep learning* secondo una prospettiva *undesign*³⁸. Ovvero, facendo prevalere le ragioni processuali e la tutela dei diritti sul segreto industriale, sino ad ottenere informazioni rilevanti del codice sorgente e la logica specifica degli algoritmi utilizzati quale ineludibile condizione per l'uso di tali tecniche nel processo, al di là di specifiche istanze della parte³⁹. Se l'obiettivo dei sistemi di IA è quello di garantire la certezza degli esiti e in tempi rapidi, non potrà allora negarsi un'informazione significativa sul sistema impiegato, sui codici sorgenti e sulle logiche sottese⁴⁰. Non basta, pertanto, che la motivazione rinvii semplicemente agli esiti algoritmici. Si tratterebbe di un approccio fondato solo sui risultati. Per il diritto, invece, è

³⁷ Sul ruolo dell'argomentazione giuridica, tra gli altri, MacCormick 1978, MacCormick, 2007 e MacCormick 2005.

³⁸ E cioè "exclusionary action as design" attraverso le strategie della *inhibition*, *displacement* ed *erasure* come ha evidenziato Pierce 2014, 36-39.

³⁹ De Minico 2024, 92.

⁴⁰ Ibid.

determinante il modo di procedere per avere una decisione legittima, valida ed effettiva. La metodologia dei mezzi primeggia su quella dei risultati, pur senza negare l'importanza di questi ultimi.

Insomma, più che di procedure di valutazione *ex ante* sull'affidabilità, comprensibilità, solidità e sicurezza del sistema (secondo la formula adoperata nelle indicazioni del Consiglio dell'Unione europea del 14 ottobre 2020), i diritti hanno bisogno di procedure che regolino l'uso giudiziale delle tecnologie e che sappiano trasferire la normatività agli algoritmi che per loro natura non la possiedono.

5. Visioni *undesign* dell'intelligenza artificiale

Si parla spesso di controllo umano significativo [*meaningful human control* – MHC] o di controllo umano supervisionato [*supervisory human control* – SHC], con una terminologia ispirata agli usi dei sistemi autonomi militari. Per l'esperienza giuridica, tuttavia, tali approcci non sono sufficienti, in quanto sacrificano la "consapevolezza situazionale" da parte dell'uomo che in tale ambito è fondamentale⁴¹. In un contesto associativo, invece, la cooperazione uomo-macchina [*human-machine teaming* – HMT] richiede la partecipazione attiva degli agenti (artificiali e umani) in sistemi multi-verso. Con questo approccio, il controllo umano non è basilare quanto il principio associativo⁴². Tuttavia, per essere effettiva la cooperazione non può che fondarsi sulle capacità comunicative non riducibili alla semplice trasmissione di contenuti, del tutto inutile se non vengono processati ed effettivamente compresi dai destinatari⁴³. Perciò, affidare l'interazione uomo-macchina unicamente alle interfacce (e cioè agli algoritmi e alla combinazione di dati), e tramite queste veicolare gli *input* per le decisioni digitali, non basta a rendere queste ultime neutrali, certe ed effettive; migliori, in poche parole, delle decisioni umane. Sono, infatti, le stesse interfacce a filtrare, condizionare, manipolare, predeterminare la volontà del soggetto/utente inconsapevole. È, invece, a livello della giustificazione che può rendersi compiuta l'interazione uomo-macchina in dinamiche procedurali⁴⁴. Costruendo un adeguato contesto di giustificazione il giudice si allena a controllare i condizionamenti esterni, impegnandosi in un coinvolgimento significativo nell'intero proces-

⁴¹ È il motivo da cui si muove lo studio di Tsamados, Floridi, Taddeo 2024, 1-14, part. 7-8.

⁴² O'Neill, McNeese, Schelble 2020, 904-938. Qui si legge: "at least one human working cooperatively with at least one autonomous agent, where an autonomous agent is a computer entity with a partial or high degree of self-governance with respect to decision-making, adaptation, and communication".

⁴³ Loddo 2024, 367-368, la quale richiama la teoria dei sistemi di Niklas Luhmann.

⁴⁴ Thagard 2021, 34.

so decisionale (e non solo come impotente controllore) per produrre una motivazione rafforzata, opposta ad una motivazione apparente che si appropri sbrigativamente degli esiti del calcolo. E può anche optare per l'esclusione del processo decisionale macchinico quando risulta opaco, incomprensibile, poco trasparente o verificabile e, in generale, iniquo⁴⁵. La scelta sul non uso giuridico dell'IA, in una prospettiva opposta all'etica *by design*, diventa indispensabile per il diritto giurisdizionale. Il processo, infatti, a differenza di qualsiasi altro campo di decisione e azione non può soggiacere *sic et simpliciter* a forme di determinismo tecnologico e ad una visione manageriale dell'esperienza giuridica, ma costruisce il senso della normatività nella pratica del diritto in azione [*law in action*]. Ed anche a questo livello la scelta umana rimane indispensabile.

Ebbene, nell'individuare il diritto, nel valutare il fatto, nell'incontrare la parte e nell'interpretare le regole, il giudice non può che interrogare problematicamente la propria coscienza, con libertà e responsabilità completa. Ma non basta un'attività solipsistica, perché la decisione giudiziale è essenzialmente il frutto di procedure alla base dell'esperienza dialogica nel processo. Certamente i sistemi di IA possono fornire un utile supporto per scegliere la soluzione comparativamente migliore tra quelle possibili, ma soffrirà maggiori problemi nella ricerca di una soluzione ideale da comparare con una possibile. Invero, in molti casi, il giudice cerca la soluzione ideale confrontandola con quelle logico-giuridicamente possibili. In questo modo si comporta da legislatore. È noto l'art. 1 del codice civile svizzero elaborato nel 1912 da Eugen Huber che stabilisce: “[...] Nei casi non previsti dalla legge il giudice decide secondo la consuetudine e, in difetto di questa, secondo la regola che egli adotterebbe come legislatore”. Potremmo richiamare l'algoritmo di Huber, ma anche questo è impotente dinanzi alla necessità che l'interprete si confronti con l'evidenza⁴⁶. L'evidenza non si analizza, né si simula. È semmai il frutto di un *Hin und Her*, un engischiano “avanti ed indietro” fino alla decisione finale. Ben più articolato sarebbe il caso di decisioni dedotte da un sistema di valori. Non ho volutamente affrontato questa ipotesi, ma si può ben pensare che anche in questo ambito il riconoscimento del diritto vigente è comunque frutto della scelta tra tanti valori possibili. Insomma, sia che si voglia trovare la soluzione ideale (cioè, la soluzione assolutamente migliore), sia quando si voglia fornire la soluzione comparativamente migliore tra quelle possibili, la scelta umana è una condizione ineliminabile per

⁴⁵ Secondo le conclusioni del Consiglio dell'Unione europea “Accesso alla giustizia – Cogliere le opportunità della digitalizzazione” (2020/C 342 I/01).

⁴⁶ Il riferimento all'algoritmo di Huber è in Lombardi 1982, 309.

la decisione giudiziale. Così come è necessaria la scelta anche per decisioni algoritmiche.

È pur vero che nulla impedisce di considerare gli esiti del calcolo come verità infallibili per giustificare un incondizionato affidamento e la delega di ogni attività umana, inclusa quella giudiziale. È una rappresentazione possibile ma, come visto, non senza rischi. Il rischio maggiore è forse quello di sacrificare lo Stato di diritto, le sue forme e le procedure interne indispensabili per renderlo effettivo. Il diritto è, difatti, costruito su procedure che assicurano un ruolo proattivo di partecipazione dialettica del soggetto ai processi decisionali che lo riguardano. Alcune forme rafforzano questa partecipazione e con essa i diritti tutelati nel processo. Si pensi alla motivazione di una decisione giudiziale. Lo Stato di diritto non può, d'altronde, ridursi all'efficientismo dei processi computazionali propri dell'IA debole (o riproduttiva) e dell'IA forte (o produttiva). Quest'ultima non ha ancora realizzato tutte le proprie potenzialità, ma si mostra da subito altrettanto insufficiente. Dall'altro lato, come si è mostrato, ciò non vuol dire negare la rilevanza anche regolativa degli esiti tecnologici per la realtà sociale. Ma non basta. Per la decisione giudiziale occorre, in ogni caso, stabilire le condizioni per costruire con metodo e intenzione la decisione e realizzare le finalità del diritto in un contesto responsabile (*accountable*). I sistemi di IA possono avere funzioni documentarie e decisionale-consulente, ma non danno esiti certi ed anzi moltiplicano l'ambito umano delle possibilità di scelta. E prima di ogni scelta, resta la coscienza del soggetto messa ancora davanti alla conoscenza dispersa nelle tante informazioni.

Riferimenti

- Agrusti, Stefania. 2013. "Incompletezza gödeliana dei sistemi giuridici." In: *Annali del Dipartimento Jonico in Sistemi Giuridici ed Economici del Mediterraneo: Società, Ambiente, Culture*, vol. I, 1-17. Bari: Cacucci.
- Anscombe, Gertrude Elizabeth Margaret. 1958. "On Brute Facts." *Analysis* 18(3): 69-72.
- Boden, Margaret Ann. 2019. *L'Intelligenza Artificiale*. Bologna: il Mulino.
- D'Acquisto, Giuseppe. 2022. *Decisioni algoritmiche. Equità, causalità, trasparenza*. Torino: Giappichelli.
- De Anna, Gabriele. 2019. "Automi, responsabilità e diritto." *Rivista di filosofia del diritto* VIII(1): 125-142.
- De Minico, Giovanna. 2024. "Giustizia e intelligenza artificiale: un equilibrio mutevole?" *AIC* 2: 85-108.

- Finocchiaro, Giusella. 2024. *Intelligenza artificiale. Quali regole?* Bologna: il Mulino.
- Floridi, Luciano. 2021. “Agere sine Intelligere. L’intelligenza artificiale come nuova forma di agire e i suoi problemi etici.” In: *L’intelligenza artificiale. L’uso delle nuove macchine*, curato da Luciano Floridi, Federico Cabitza, 115-183. Milano: Bompiani, Milano.
- Floridi, Luciano. 2020. *Pensare l’infosfera. La filosofia come design concettuale*. Milano: Raffaello Cortina Editore.
- Floridi, Luciano. 2023. *The Ethics of Artificial Intelligence. Principles, Challenges, and Opportunities*. Oxford: Oxford University Press.
- Floridi, Luciano. 2014. *The Fourth Revolution. How the Infosphere Is Reshaping the Human Reality*. Oxford: Oxford University Press.
- Floyd, Juliet. 2016. “Chains of Life: Turing, Lebensform, and the Emergence of Wittgenstein’s Later Style.” *Nordic Wittgenstein Review* 5(2): 7-89.
- Floyd, Juliet. 2018. “Lebensformen: Living Logic.” In: *Language, Form(s) of Life and Logic. Investigations after Wittgenstein*, a cura di Martin Christian, 59-92. Berlin: De Gruyter.
- Garapon, Antoine, Lassègue, Jean. 2021. *La giustizia digitale. Determinismo tecnologico e libertà*. Bologna: il Mulino.
- Gerd, Leonhard. 2016. *Technology vs. Humanity: The coming clash between man and machine*. Independently published.
- Gödel, Kurt. 1931. “Über formal unentscheidbare Sätze der «Principia Mathematica» und verwandter Systeme I.” *Monatshefte für Mathematik* 38(1): 173-198.
- Hilbert, David. 1923. “Die logischen Grundlagen der Mathematik.” *Mathematische Annalen* 88(1-2): 151-157.
- Hilbert, David. 1926. “Über das Unendliche.” *Mathematische Annalen*, 95(1): 161-190.
- Incampo, Antonio. 2017. “Propter iustos. Probabilità e ragioni del dubbio nella giustizia penale.” In: *Giudizio penale e ragionevole dubbio*, a cura di Incampo Antonio, Scalfati Adolfo, 69-72. Bari: Cacucci.
- Incampo, Antonio. 2016. *Metafisica del processo. Idee per una critica della ragione giuridica*. Bari: Cacucci.
- Incampo, Antonio. 1996. *Sul fondamento della validità deontica. Identità non-contraddizione*. Roma-Bari: Laterza.

- Lessig, Lawrence. 2000. "Code Is Law: On Liberty in Cyberspace." *Harvard Magazine*. <https://www.harvardmagazine.com/2000/01/code-is-law-html>
- Licklider, Joseph Carl Robnett. 1960. "Man-computer symbiosis." *IRE Transactions on Human Factors in Electronics* 1: 4-11.
- Loddo, Olimpia Giuliana. 2024. "L'agire sociale ai tempi dell'intelligenza artificiale. Il concetto di 'istituzione sociale-digitale'." *L'Ircocervo – Prima rivista elettronica italiana di Metodologia giuridica, Teoria generale del diritto e Dottrina dello Stato* 23(1): 359-371.
- Lombardi Vallauri, Luigi. 1982. "Informatica e criteri «politici» o valutativi della decisione giuridica." *Jus* 3: 303-315.
- Lombardi Vallauri, Luigi. 1998. "Norme vaghe e teoria generale del diritto." *Ars Interpretandi*: 155-163.
- Lombardi Vallauri, Luigi. 1981. *Corso di filosofia del diritto*. Padova: Cedam.
- Luzzati, Claudio. 1990. *La vaghezza delle norme. Un'analisi del linguaggio giuridico*: Milano: Giuffrè.
- MacCormick, Neil. 2007. *Institutions of Law: an Essay in Legal Theory*. Oxford: Oxford University Press.
- MacCormick, Neil. 1978. *Legal Reasoning and Legal Theory*. Oxford: Oxford University Press.
- MacCormick, Neil. 2005. *Rhetoric and The Rule of Law: A Theory of Legal Reasoning*. Oxford: Oxford University Press.
- Mauss, Marcel. 2002. *Saggio sul dono. Forma e motivo dello scambio nelle società arcaiche*. Torino: Einaudi.
- McCarthy, John, Minsky, Marvin L., Rochester, Nathaniel, Shannon, Claude E. 1955. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence." <http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf>
- Nobili, Massimo. 2003. "Storie d'una illustre formula: il 'libero convincimento' negli ultimi trent'anni." *Rivista italiana di diritto e procedura penale* 46(1/2): 71-92.
- O'Neill, Thomas, McNeese, Nathan, Schelble, Beau. 2020. "Human-autonomy teaming: a review and analysis of the empirical literature." *Human Factors* 64(5): 904-938. <https://doi.org/10.1177/0018720820960865>.
- OECD. 2019. *Artificial Intelligence in Society*. Parigi: OECD Publication.
- Palmirani, Monica, Sapienza Salvatore, Bompreszi Chantal. 2022. "Il ruolo dell'Intelligenza Artificiale nel sistema giustizia: funzionalità,

- metodologie, principi.” In: La trasformazione digitale della giustizia nel dialogo tra discipline. Diritto e Intelligenza artificiale, curato da Palmirani, Monica, Sapienza, Salvatore, 1-36. Milano: Giuffrè Francis Levebvre.
- Perelman, Chäim. 1961. “La distinction du fait et du droit. Le point de vue du logicine.” *Dialettica* 3/4: 268-278.
- Pierce, James. 2014. “Undesigning Interaction.” *Interactions* 21(4): 36-39.
- Ricci, Stefano, Rossetti, Andrea (a cura di). 2024. *Aspetti giuridici della società dell’informazione/Legal Issues in the Information Society*. Milano: Giuffrè Francis Levebvre.
- Romeo, Francesco. 2024. “La produzione culturale dell’artificiale: prodromi di uno studio metaepistemologico su diritto ed informazione.” *L’Ircocervo – Prima rivista elettronica italiana di Metodologia giuridica, Teoria generale del diritto e Dottrina dello Stato* 23(1): 51-69.
- Russell, Stuart, Norvig, Peter. 2010. *Artificial Intelligence: A Modern Approach*. Londra: Pearson.
- Scarpelli, Uberto. 1997. *Cos’è il positivismo giuridico*. Napoli: Esi.
- Schiavello, Aldo. 2023. *Conoscere il diritto*. Modena: Mucchi.
- Schiavello, Aldo, Velluzzi, Vito (a cura di). 2005. *Il positivismo giuridico contemporaneo. Una antologia*. Torino: Giappichelli.
- Searle, John Rogers. 1996. *La costruzione della realtà sociale*. Milano: Edizioni di Comunità.
- Thagard, Paul. 2021. *Cervelli a confronto. Perché l’intelligenza umana è diversa da quella degli animali e dei robot*. Milano: FrancoAngeli.
- Tsamados, Andreas, Floridi, Luciano, Taddeo, Mariarosaria. 2024. “Human control of AI systems: from supervision to teaming.” *AI and Ethics*: 1-14. <https://doi.org/10.1007/s43681-024-00489-4>.
- Turing, Alan Mathison. 1950. “Computer Machinery and Intelligence.” *Mind* 59 (236): 433-460.
- Turing, Alan Mathison. 1937. “On Computable Numbers with an Application to the «Entscheidungsproblem».” *Proceedings of the London Mathematica Society* 42(1): 230-265.
- Velluzzi, Vito. 2006. *Le clausole generali. Semantica e politica del diritto*. Milano: Giuffrè.
- Wittgenstein, Ludwig. 1995. *Ricerche filosofiche*. Torino: Einaudi.