

ATTUALITÀ

ENRICO MAESTRI

Verità giudiziaria e intelligenza artificiale. Dalla prova baconiana alla meta-prova algoritmica

Questo saggio indaga il rapporto tra epistemologia giudiziaria e intelligenza artificiale, concentrandosi sulla *quaestio facti* e sulla trasformazione della prova legale nell'era digitale. A partire dall'opposizione tra probabilità bayesiana e plausibilità baconiana, valuta criticamente i limiti del formalismo statistico nel ragionamento giudiziario e il ruolo marginale della regola di Bayes nei contesti processuali. Lo studio esamina l'IA – in particolare il *deep learning* – come sistema inferenziale non narrativo e opaco, capace di classificare, correlare e prevedere, ma privo di spiegabilità in senso giuridico. Viene tracciata la frizione epistemica tra giudici umani e algoritmi, esplorando ipotesi relative a *bias*, ragionamento causale e *accountability* delle macchine. Particolare attenzione è riservata all'algoritmo come «prova della prova», alla sfida delle tracce digitali come evidenze e alla domanda di trasparenza, contestabilità e interpretabilità. Il lavoro si conclude proponendo un quadro di digisprudenza responsabile, capace di integrare la potenza inferenziale della macchina entro procedure giudiziarie epistemicamente solide.

Judicial Truth and Artificial Intelligence: From Baconian Evidence to Algorithmic Meta-Evidence

This paper investigates the relationship between judicial epistemology and artificial intelligence, focusing on the quaestio facti and the transformation of legal evidence in the digital age. Drawing from the opposition between Bayesian probability and Baconian plausibility, it critically assesses the limitations of statistical formalism in judicial reasoning and the marginal role of Bayes' rule in court settings. The study examines AI—especially deep learning—as a non-narrative, opaque inferential system capable of classifying, correlating, and predicting, yet lacking explainability in the legal sense. It maps the epistemic friction between human judges and algorithms, exploring hypotheses around bias, causal reasoning, and machine accountability. Particular attention is paid to the algorithm as «evidence of evidence», the challenge of evidentiary digital traces, and the demand for transparency, contestability, and interpretability. The work concludes by proposing a framework of responsible digisprudence capable of integrating machine inferential power within epistemically robust judicial procedures.

SOMMARIO: 1. Introduzione: la *quaestio facti* nell'epoca computazionale. 2. Epistemologia giudiziaria contemporanea 3. Intelligenza artificiale e prova: un paradigma computazionale. 3.1. Verso una epistemologia algoritmica del fatto? 3.2. L'algoritmo come esperto computazionale: deferenza, perizia, *accountability*. 3.2.1 Deferenza epistemica vs deferenza semantica. 3.2.2 Opacità strutturale e spostamento della responsabilità. 3.2.3 Condizioni di ammissibilità e *accountability*. 3.2.4 Dalla cooperazione alla supplenza: linea di demarcazione. 3.3. Meta-prova algoritmica e *gatekeeping* dell'evidenza. 3.4. Oltre Bayes? Per una epistemologia computazionale non riduzionista del giudizio probatorio. 4. Frizioni epistemiche e paradosso della spiegabilità. 5. Prospettive: ibridazione o incompatibilità? 5.1 Epistemologie ibride: la «co-prova» tra giudice e IA. 6. Verità computate. Le evidenze digitali tra tecnica forense e crisi del giudizio. 7. Scardinamenti positivi. Verso una computazione emancipante della prova. 8. L'algoritmo e il diritto della prova: oltre l'antitesi. 9. Conclusione: l'algoritmo come istituzione razionale.

1. *Introduzione: la quaestio facti nell'epoca computazionale.* L'accertamento del fatto nel processo giuridico è sempre stato un banco di prova epistemologico, uno spazio teorico in cui il diritto incontra la conoscenza e ne misura le

condizioni di legittimità¹. Nella distinzione classica (per quanto epistemologicamente opaca²) tra *quaestio iuris* e *quaestio facti*, è la seconda a porre le sfide più profonde: è qui che il giudice, oltrepassando la dimensione normativa e interpretativa, si confronta con l'indecidibile del reale, con la materia refrattaria alla codificazione. Il diritto, in quanto pratica discorsiva regolata, si trova così a dover costituire una verità che non gli è immanente, ma che dev'essere evocata, costruita e giustificata attraverso procedure probatorie³.

In tale contesto, l'epistemologia giudiziaria – intesa come riflessione sui modi attraverso cui si forma il convincimento del giudice – assume una rilevanza sempre maggiore, specialmente nell'epoca della *datafied justice*, in cui il diritto si confronta con nuove forme di conoscenza mediata, algoritmica, computazionale. L'ingresso dell'intelligenza artificiale nei meccanismi dell'accertamento del fatto non costituisce soltanto un problema tecnico, bensì una provocazione teorica: che tipo di inferenza è legittima in un processo? Quale statuto epistemico può essere riconosciuto a una decisione che emerge da correlazioni statistiche e non da un processo argomentativo? Il giudizio sul fatto, in quanto esempio paradigmatico di decisione in condizioni di incertezza⁴, è compatibile con una logica predittiva e automatizzata, oppure si radica in una forma di razionalità pratica non formalizzabile?

L'obiettivo di questo saggio è esplorare le tensioni teoriche e le frizioni epistemologiche che emergono dal confronto tra la razionalità probatoria tradizionale – fondata su principi quali la coerenza narrativa, la plausibilità logica e l'analisi cumulativa della prova – e i modelli inferenziali propri dell'intelligenza artificiale, in particolare quelli sviluppati nel paradigma del *deep learning*. La riflessione si colloca dunque all'intersezione fra filosofia del diritto, teoria della prova e scienza dei dati, cercando di chiarire non solo se l'IA possa supportare il giudice, ma se possa davvero partecipare, a pieno titolo, alla costruzione della verità giudiziaria.

La *quaestio facti* non è semplicemente un problema di corrispondenza tra proposizioni e fatti, bensì una forma di costruzione istituzionale della verità in

¹ POLI, *Logica e razionalità nella ricostruzione giudiziale dei fatti*, Modena, 2021.

² TARUFFO, «Note in tema di giudizio di fatto», in *Rivista di diritto civile*, vol. I, 1971, 33-51.

³ HRUSHKA, *La costituzione del caso giuridico. Il rapporto tra accertamento fattuale e applicazione giuridica*, Bologna, 2009.

⁴ LUPARIA DONATI, *Il concetto di prova alla luce dell'intelligenza artificiale*, Milano, 2005.

un contesto normativamente orientato⁵. In tale quadro, l'adozione di strumenti epistemici provenienti dal dominio della statistica computazionale pone una questione decisiva: il diritto è disposto a sacrificare il proprio paradigma veritativo (contestuale, prudenziale, qualitativo) in favore di un criterio predittivo (quantitativo, atomistico, correlazionale)? Oppure è possibile pensare un'ibridazione, un nuovo spazio epistemico dove la razionalità algoritmica e la razionalità giudiziaria possano convivere in una forma di co-produzione epistemica?

Il saggio si apre dunque con una ricognizione sull'epistemologia giudiziaria contemporanea (§2), concentrandosi sulla dicotomia tra probabilità logica (o baconiana) e probabilità bayesiana (o pascaliana), con attenzione ai criteri delineati nelle sentenze Franzese e Daubert. Segue un'analisi dei modelli epistemici sottesi all'intelligenza artificiale (§3), e in particolare del paradigma del *deep learning* come nuova forma di inferenza automatizzata. Il paragrafo 4 discuterà le frizioni tra le due razionalità (giudiziaria e algoritmica), mentre il paragrafo 5 cercherà di delineare alcune prospettive di integrazione critica. L'obiettivo finale (§8-9) non è solo verificare se l'IA possa contribuire all'accertamento del fatto, ma chiarire a quali condizioni ciò avvenga senza erodere i fondamenti epistemici del processo; prima, il saggio attraversa le frizioni (§4) e le vie d'integrazione critica (§5), affronta la specificità delle verità computate (§6) e propone al §7 una tesi costruttiva: una *computazione forense emancipante* come co-prova giudice-IA, subordinata a trasparenza, contestabilità e responsabilità.

2. Epistemologia giudiziaria contemporanea. Negli ultimi decenni la teoria della prova ha progressivamente abbandonato il paradigma veritativo-formalista per assumere un modello narrativo-coerente⁶, che valorizza la capacità ermeneutica del giudice, intesa come ricomposizione di elementi eterogenei in una trama significativa⁷.

Ne consegue che la «verità processuale» non è più l'esito di una corrispondenza tra proposizioni e fatti, ma il prodotto di una costruzione inferenziale

⁵ FARANO, *Il giudice e la buona scienza. L'istituzione del fatto scientifico in giudizio*, Torino, 2024,

⁶ TISSONE, *Breviario di logica giuridica. Gli strumenti per un ragionamento giuridico chiaro e persuasivo*, Bologna, 2024

⁷ PASTORE, *Giudizio, prova, ragion pratica. Un approccio ermeneutico*, Milano, 1996.

in condizioni di incertezza⁸ che integra fonti eterogenee secondo criteri di plausibilità tutto considerato⁹, coerenza interna e congruenza esterna con ciò che sappiamo del mondo¹⁰. La motivazione non è mero adempimento formale¹¹: è il luogo in cui si rende pubblicamente visibile il tragitto cognitivo che sorregge la decisione e si offre alla contestazione intersoggettiva¹².

La svolta è ben espressa da Franzese (Cass., Sez. Un., 2002, n. 30328)¹³, che riformula il nesso causale in chiave di probabilità logica «tutto considerato»: niente automatismi probabilistici, ma una sintesi ragionata della convergenza indiziaria e delle regolarità empiriche. In parallelo, il diritto della prova scientifica si è ancorato a *standard* di affidabilità (validità della metodologia, controllabilità, tasso di errore, comunità scientifica), lungo una traiettoria che mette in dialogo i criteri Daubert (U.S. Supreme Court, 1993)¹⁴ con l'orientamento italiano inaugurato da Cozzini (Cass., 2010)¹⁵, soprattutto nei casi di eziologia complessa: non si chiede la certezza matematica, ma metodologie buone e ragioni controllabili per accreditare un risultato tecnico come base cognitiva del processo.

La cornice teorica più feconda per comprendere questo passaggio è la proposta di Susan Haack, che difende un fondherentismo della prova: né puro fondazionalismo delle «pietre angolari», né puro coerentismo della «storia che regge», ma un reticolo in cui elementi empirici, regole di inferenza e pratiche argomentative si sostengono reciprocamente (*crossword-model*). In chiave giuridica ciò implica: (i) attenzione alla qualità delle singole evidenze, (ii) alla coerenza sistemica dell'insieme, (iii) alla giustificabilità pubblica dei passaggi. Ne discende un diritto probatorio pragmatista e antiriduzionista: aperto agli apporti della scienza, ma refrattario sia al decisionismo sia allo scientismo; disponibile a utilizzare strumenti statistici o computazionali, purché integrati nella motivazione e soggetti a contraddittorio.

⁸ TARUFFO, *La prova dei fatti giuridici. Nozioni generali*, Milano, 1992.

⁹ TUZET, *La prova ragionata*, Milano, 2023,

¹⁰ HAACK, *Legalizzare l'epistemologia. Prova, probabilità e causa nel diritto*, Milano, 2015.

¹¹ BARBERIS, *Cosa resta del sillogismo giudiziale? Riflessioni a partire da Beccaria*, in *Materiali per una storia della cultura giuridica*, 1, 2015, 163-176.

¹² FERRER BELTRÁN, *La valoración racional de la prueba*, Madrid, 2016.

¹³ Cass., Sez. Un., 11.9.2002, n. 30328 (Franzese).

¹⁴ Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993).

¹⁵ Cass., Sez. IV, 2010 (Cozzini).

Il lessico della probabilità baconiana¹⁶ – induzione cumulativa, incremento di forza attraverso convergenza e robustezza – non esclude in principio gli strumenti bayesiani; semplicemente ricusa ogni riduzione del giudizio a «peso numerico» di evidenze atomistiche. La soglia processuale non è una percentuale astratta, ma un equilibrio ragionato tra prove, alternative e rischi di errore, reso persuasivo da buone ragioni e *standard* metodologici¹⁷. È su questo terreno che, nel § 3, si discuterà se e come integrare strumenti algoritmici come fonti ausiliarie di ordine e controllo, senza convertirli in sostituti della deliberazione¹⁸.

3. *Intelligenza artificiale e prova: un paradigma computazionale.* L'ingresso dell'IA nei meccanismi probatori rappresenta una provocazione epistemica: l'adozione di strumenti algoritmici nel processo segna una frattura nell'epistemologia giudiziaria tradizionale, fondata su una razionalità dialogica e su procedure argomentative¹⁹.

L'intelligenza artificiale, nella sua declinazione più avanzata e diffusa, ossia quella fondata su architetture di deep learning, introduce un mutamento radicale nel paradigma epistemico con cui si affronta il problema dell'accertamento del fatto. Come osserva Barberis, ci troviamo davanti a una transizione dal ragionamento probatorio alla computazione performativa²⁰.

Essa non si fonda su premesse assiomatiche, né su inferenze logico-deduttive, bensì su un processo di apprendimento statistico volto a ottimizzare funzioni predittive a partire da grandi moli di dati eterogenei²¹. Nelle tecnologie neurali, la regola è rimpiazzata dalla regolarità: il principio di legalità è messo sotto pressione dalla prevalenza di correlazioni statistiche prive di fondamento giustificativo²².

¹⁶ COHEN, *The Probable and the Provable*, Oxford, 1977.

¹⁷ TUZET, *Assessment criteria or standards of proof?* In *Artificial Intelligence and Law*, vol. 28, n. 1 (2018/2020), 91-109.

¹⁸ BARBERIS, *Separazione dei poteri e giustizia digitale*, Milano-Udine, 2023.

¹⁹ BATTELLI, *Giustizia predittiva, decisione robotica e ruolo del giudice*, in *Giustizia Civile*, n. 2, 2020, 281-319.

²⁰ BARBERIS, *Giustizia predittiva: ausiliare e sostitutiva. Un approccio evolutivo*, in *Milan Law Review*, 3(2), 2022, 1-18.

²¹ GARAPON, LASSEGUE, *La giustizia digitale. Determinismo tecnologico e libertà*, Bologna, 2021.

²² PUNZI, *Difettività e giustizia aumentata. L'esperienza giuridica e la sfida dell'umanesimo digitale*, in *Ars Interpretandi. Rivista di ermeneutica giuridica*, 1, 2021, 113-128.

L'IA, in tal senso, si configura come un sistema epistemico non narrativo: essa non costruisce spiegazioni, non articola narrazioni causali, non ricerca coerenze ermeneutiche²³. Piuttosto, individua *pattern* ricorrenti, correlazioni latenti e classificazioni emergenti. L'algoritmo non partecipa al discorso giuridico, bensì opera come attore muto, privo di intenzionalità e di comprensione della rilevanza normativa dei dati²⁴.

Nel dominio giudiziario, tale impostazione si traduce in un approccio alla prova che privilegia la previsione dell'esito rispetto alla comprensione della struttura causale dei fatti²⁵. Nel diritto, il valore probatorio di un elemento non è mai soltanto statistico, ma è inscritto in una rete di significati normativi, cosa che la predizione algoritmica tende a oscurare.

L'algoritmo, addestrato su *dataset* giudiziari, può apprendere quali caratteristiche ricorrono nei casi di condanna, quali fattori predicono la recidiva, quali *pattern* testuali sono associati a sentenze favorevoli o sfavorevoli. Ma in nessun caso esso spiega perché quei dati abbiano rilievo: si limita a correlare per inferire, in una logica performativa piuttosto che giustificativa. Barberis sottolinea che «la sostituzione di giudici umani con giudici macchina aumenterebbe la certezza del diritto non per ragioni logiche ma solo statistico-probabilistiche»²⁶.

La capacità del *deep learning* di operare su dati non strutturati – testi, immagini, video, segnali biometrici – ne fa uno strumento formidabile di classificazione e selezione probatoria²⁷. Ma tale forza computazionale ha un costo: la rinuncia alla trasparenza inferenziale, la perdita della giustificazione epistemica esplicita che caratterizza il giudizio umano²⁸. Il modello neurale apprende, ma non spiega; decide, ma non giustifica. Questo è il cuore dell'asimmetria epistemica tra giudice e algoritmo²⁹.

²³ MARCUS, *Deep learning: A critical appraisal*, in *arXiv preprint arXiv:1801.00631*, 2018.

²⁴ ERCOLE, *Il giudice-macchina: da bouche de la loi a slave to the algorithm?*, in *SudEuropa*, 2-3/2024, in spec. 15-18.

²⁵ DOMINGOS, *A few useful things to know about machine learning*, in *Communications of the ACM*, 55(10), 2012, 78-87.

²⁶ BARBERIS, *Separazione dei poteri e giustizia digitale*, cit., 106.

²⁷ HILDEBRANDT, *Smart technologies and the end(s) of law*. Cheltenham, 2015.

²⁸ CITRON, PASQUALE, *The Scored Society: Due Process for Automated Predictions*, in *Washington Law Review*, vol. 89, n. 1, 2014, 1-33.

²⁹ PUNZI, *Difettività e giustizia aumentata. L'esperienza giuridica e la sfida dell'umanesimo digitale*, cit.

3.1. *Verso una epistemologia algoritmica del fatto?* L'uso dell'intelligenza artificiale nei contesti giudiziari impone una riflessione profonda su ciò che potremmo chiamare una epistemologia algoritmica del fatto. A differenza del giudice, che motiva, argomenta, rende conto del proprio percorso inferenziale, l'algoritmo opera in modalità *black box*: la decisione è il risultato di una catena di trasformazioni matematiche opache, spesso non interpretabili nemmeno dai progettisti del sistema. L'opacità è, in tal senso, strutturale: non si tratta solo di una mancata comunicazione, ma di una non giustificabilità epistemica³⁰.

Questo scarto si riflette nel modo in cui l'IA «vede» le prove. Non come elementi dotati di valore semantico e contesto, ma come vettori numerici su cui calcolare distanze, similarità, pesi. Gli strumenti computazionali producono *score*, *ranking*, *cluster*, *embeddings*: rappresentazioni latenti che organizzano l'informazione secondo criteri di ottimizzazione statistica, ma che restano muti quanto al loro significato giuridico³¹.

Nel passaggio dalla semantica umana alla sintassi computazionale, si perde la dimensione deliberativa del giudizio, sostituita da una logica di ordinamento e correlazione. L'epistemologia algoritmica non costruisce verità, ma produce *output* ottimali, cioè funzionali allo scopo predittivo del sistema³². Questa logica strumentale mal si concilia con il carattere prudentiale, discorsivo e pubblicamente giustificabile dell'accertamento giudiziario dei fatti.

Eppure, la tesi che riduce l'algoritmo a un classificatore cieco — «L'algoritmo non vede l'imputato, ma lo inquadra in una classe»³³ — presume che ogni classificazione algoritmica sia inevitabilmente deindividualizzante, trascurando il fatto che le tecniche più avanzate (quali *embeddings*, modelli di *attention*, *few-shot learning*³⁴) consentono oggi una notevole granularità e sensibilità al

³⁰ BURRELL, *How the machine 'thinks': Understanding opacity in machine learning algorithms*, in *Big Data & Society*, vol. 3, n. 1, 2016, 1-12.

³¹ PASQUALE, *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard, 2015.

³² KROLL et al., *Accountable Algorithms*, in *University of Pennsylvania Law Review*, 165(3), 2017, 633-706.

³³ LETTIERI, *Against Prediction. Three Arguments for a Critique of Predictive Analytics and its Use in the Legal Field*, in *Ars Interpretandi. Rivista di ermeneutica giuridica*, n. 1/2021, 83-96.

³⁴ LEPRI, OLIVER, LETOUZÉ, PENTLAND, VINCK, *Fair, Transparent, and Accountable Algorithmic Decision-Making Processes*, in *Philosophy & Technology*, vol. 31, n. 4, 2018, 611-627.

contesto³⁵. L'obiettivo di delineare un'epistemologia algoritmica del fatto mira proprio a costruire un ponte tra il contestualismo del giudizio e l'adattività delle architetture neurali, non a contrapporli.

3.2. *L'algoritmo come esperto computazionale: deferenza, perizia, accountability*. Nel giudizio probatorio contemporaneo il ruolo degli esperti — medici, genetisti, chimici, psicologi forensi — è diventato essenziale per la costruzione della *quaestio facti*. In questo orizzonte, l'intelligenza artificiale (IA) si presenta come un nuovo perito computazionale: capace di classificare, correlare, riconoscere *pattern* e svolgere analisi forensi (riconoscimento facciale, *voice-print*, analisi del DNA, *document review*) con una potenza elaborativa inedita. Proprio per questo, il suo statuto epistemico esige una chiarificazione, pena lo scivolamento dalla cooperazione tecnica alla supplenza decisionale³⁶.

3.2.1 *Deferenza epistemica vs deferenza semantica*. Sulla scorta della distinzione proposta in dottrina³⁷, occorre separare:

Deferenza epistemica: il giudice utilizza l'*output* dell'IA come base conoscitiva rilevante (es.: uno *score* di rischio o una similarità tra tracce), mantenendo il controllo critico sulla sua pertinenza e sul suo peso probatorio;

Deferenza semantica: il giudice adotta il lessico computazionale (punteggi, soglie, similarità vettoriali) come se fosse isomorfo a categorie giuridiche (responsabilità, pericolosità, recidiva). È qui che nasce il rischio maggiore: lo *score* diventa sinonimo di colpevolezza, la soglia di classificazione rimpiazza il ragionevole dubbio, la misura vettoriale sostituisce la *ratio decidendi*³⁸.

Questa seconda forma di deferenza ha effetti deformanti sul contraddittorio: ciò che è calcolato tende a presentarsi come ciò che è; lo spazio per la critica

³⁵ VASWANI, A. et al., *Attention Is All You Need*, in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017, 5998-6008.

³⁶ HAACK, Legalizzare l'epistemologia. Prova, probabilità e causa nel diritto, cit.

³⁷ UBERTONE, CANALE, *Il giudice e l'esperto. Deferenza epistemica e deferenza semantica nel processo*, Torino, 2022.

³⁸ ZARSKY, *The Trouble with Algorithmic Decisions*, in *Science, Technology, & Human Values*, vol. 41, n. 1, 2016, 118-132; PASQUALE, *New Laws of Robotics: Defending Human Expertise in the Age of AI*, Harvard, 2020.

si restringe perché il linguaggio della statistica viene preso per linguaggio del diritto. Ne consegue il passaggio dall'argomentazione alla ratifica tecnica.

3.2.2 Opacità strutturale e spostamento della responsabilità. La deferenza (specie semantica) si aggrava quando il sistema è opaco (*black box*): l'IA non fornisce *reason-giving explanations*, ma soltanto esiti³⁹. L'autorità si trasferisce dalla giustificazione umana alla procedura computazionale. In questa traslazione, la responsabilità epistemica rischia di essere dislocata: il giudice valida l'output, ma non può discutere la catena inferenziale che lo ha prodotto; la motivazione scivola verso una *ex post* conferma di un risultato già normalizzato dallo strumento⁴⁰.

In più, quando l'IA seleziona e ordina le evidenze (*ranking* di documenti, *matching* di immagini o di profili genetici, *topic modeling*), essa agisce come meta-prova: determina cosa sia visibile e come debba essere valutato. Il perito computazionale diventa così un meta-valutatore che pre-struttura il campo dell'argomentazione giudiziale⁴¹.

3.2.3 Condizioni di ammissibilità e accountability. Per evitare che il perito computazionale degeneri in supplenza, il suo impiego deve essere vincolato a presidi di *accountability* proporzionati all'impatto probatorio:

1. *Reason-giving* minimo e *model facts*: tasso d'errore empirico, limiti d'uso, ipotesi operative, alternative scartate⁴².
2. Documentazione tecnica (*model cards*, *data sheets*, *versioning*): dati di addestramento, *pre-processing*, metriche d'ottimizzazione, *drift* e aggiornamenti⁴³.
3. *Audit trail* forense: tracciabilità non alterabile di parametri, soglie, passaggi di *pipeline*, possibilità di *replay* controllato⁴⁴.

³⁹ BURRELL, *How the machine 'thinks': Understanding opacity in machine learning algorithms*, cit.

⁴⁰ KROLL et al., *Accountable Algorithms*, cit.; CITRON, PASQUALE, *The scored society: Due process for automated predictions*, cit.

⁴¹ KROLL ET AL., *Accountable Algorithms*, cit.

⁴² Ibidem.

⁴³ MITCHELL, *Model Cards for Model Reporting*, in *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAccT)*, 2019, 220-229; GEBRU, *Datasheets for Datasets*, in *Communications of the ACM*, vol. 64, n. 12, 2021, 86-92.

4. Simmetria difensiva: controperizia effettiva (accesso ai materiali o *sandbox* certificata con la stessa versione del modello)⁴⁵.
5. Qualificazione processuale: l'IA come ausilio cognitivo (non prova piena), salvo i casi in cui la legge disciplini espressamente lo status e gli oneri di prova⁴⁶.

Queste condizioni non chiedono all'IA una «spiegabilità totale», ma una spiegabilità situata sufficiente a sostenere confronto, verifica e *reason-giving* giudiziale⁴⁷.

3.2.4 *Dalla cooperazione alla supplenza: linea di demarcazione.* Nel dibattito sull'epistemologia giudiziaria contemporanea, la figura del perito occupa una posizione ambivalente: da un lato è garanzia di competenza scientifica in ambiti che eccedono la formazione del giudice; dall'altro rappresenta un potenziale fattore di squilibrio epistemico nel processo, in quanto portatore di un sapere tecnico non sempre contestabile secondo le categorie del diritto. La letteratura giuridica ha indagato tale tensione attraverso la categoria di «deferenza epistemica», ovvero quel meccanismo di attribuzione fiduciaria mediante il quale il giudice recepisce, talora in modo acritico, l'opinione dell'esperto come elemento probante. Il rischio, come hanno sottolineato Ubertone e Canale⁴⁸, è che la deferenza possa scivolare dalla sua forma epistemicamente giustificata a una deferenza semantica, in cui il giudice rinuncia alla comprensione effettiva del contenuto tecnico e si limita ad accettarne l'autorità formale. Tale rischio si amplifica esponenzialmente nel caso dell'intelligenza artificiale, in particolare quando essa agisce come sistema autonomo di supporto alla decisione probatoria, producendo *score*, *cluster* o raccomandazioni sulla base di reti neurali profonde o modelli probabilistici non interpretabili. L'IA, in questo scenario, tende ad assumere il ruolo di un «super-esperto» computa-

⁴⁴ RAJI, *Closing the AI Accountability Gap: Defining Auditing and Its Role in AI Governance*, in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAccT)*, 2020, 33-44.

⁴⁵ KROLL, et al., *Accountable Algorithms*, cit.

⁴⁶ HILDEBRANDT, *Algorithmic Regulation and the Rule of Law*, in *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 376, n. 2128, 2018, 1-15.

⁴⁷ SELBST, BAROCAS, *The Intuitive Appeal of Explainable Machines*, in *Fordham Law Review*, vol. 91, 2022, 1085-1139.

⁴⁸ UBERTONE, CANALE, *Il giudice e l'esperto. Deferenza epistemica e deferenza semantica nel processo*, cit.

zionale, la cui *auctoritas* non deriva da un sapere umano esplicito ma da una capacità predittiva emergente e statistico-distributiva, priva di trasparenza logico-argomentativa.

La deferenza epistemica nei confronti di tali sistemi non è solo una questione di sbilanciamento cognitivo, ma anche un nodo teorico che coinvolge la filosofia della responsabilità giudiziaria. Se il giudice delega la valutazione del fatto a un sistema di apprendimento automatico opaco, egli abdica al suo ruolo di garante ermeneutico della razionalità pubblica, trasformando il processo in un protocollo tecnico-esecutivo, anziché in un luogo di giustificazione razionale intersoggettiva⁴⁹.

L'«automatismo argomentativo» che può derivare dall'uso non riflessivo delle prove scientifiche rischia di innescare una nuova forma di deferenza cieca, resa ancora più insidiosa dall'illusione di oggettività che accompagna l'*output* algoritmico. Non si tratterebbe più di un giudice che ascolta l'esperto per poi decidere, ma di un giudice che accetta l'*output* perché non può né spiegarlo né contraddirlo.

Un simile scenario rende urgente una riflessione critica sui modelli epistemici della prova giudiziaria, in grado di distinguere tra assistenza e supplenza, tra uso ausiliario e delega. Come ha scritto Susan Haack⁵⁰, il diritto non può rinunciare a una propria epistemologia interna, pena la sua dissoluzione in una pratica tecnocratica. Questo vale oggi, a maggior ragione, nell'era dell'intelligenza artificiale, dove la deferenza non si esercita più verso una persona, ma verso un sistema, e dove il problema dell'affidabilità si confonde con quello dell'inaccessibilità semantica.

In sintesi, l'IA diventa un super-esperto indebito quando:

- impone il proprio lessico come criterio normativo (deferenza semantica),
- rende inaccessibile la catena inferenziale rilevante per la decisione,
- sposta di fatto la motivazione dal foro esterno della giurisdizione al foro interno della macchina.

La linea di demarcazione è semplice: *experta docet, iudex motivat*. L'IA istruisce e potenzia la raccolta/valutazione delle evidenze; il giudice rimane ti-

⁴⁹ KAUFMANN, *Filosofia del diritto ed ermeneutica*, Milano, 2003.

⁵⁰ HAACK, *Legalizzare l'epistemologia. Prova, probabilità e causa nel diritto*, cit.

tolare della giustificazione pubblica della decisione⁵¹. Senza questa gerarchia, la computazione non assiste la giurisdizione: la sostituisce.

3.3. *Meta-prova algoritmica e gatekeeping dell'evidenza.* Nel quadro della progressiva digitalizzazione dei processi cognitivi e decisionali, l'IA non si limita a produrre nuove prove: svolge una funzione *meta-probatatoria* che organizza, seleziona, ordina e pesa l'insieme delle evidenze disponibili, incidendo strutturalmente sulla percezione giudiziale della verità. In questa veste, l'algoritmo non è solo un ausilio alla ricostruzione del fatto, ma una architettura epistemica che determina quali prove diventano accessibili, rilevanti e in quale forma sono presentate⁵².

Tale mediazione computazionale si manifesta in molteplici ambiti: *risk assessment* (recidiva, pericolosità), *forensics* automatizzata (riconoscimento facciale, *image matching*, analisi del DNA con *probabilistic genotyping*), *document review* ed *e-discovery*, *topic modeling* e *anomaly detection* in ambito contabile. L'output dell'IA viene spesso presentato come elemento «oggettivo»; in realtà è il prodotto di una modellizzazione statistica basata su ipotesi, metriche e soglie talora opache e non immediatamente contestabili da difesa e giudice⁵³.

L'IA agisce così come «prova della prova»: una macchina euristica che struttura l'evidenza secondo criteri di similarità, prossimità e correlazione, talvolta in sostituzione dell'argomentazione narrativa. Come osservano Zerilli et al.⁵⁴, molti sistemi operano tramite euristiche di prossimità che forniscono classificazioni operative, non spiegazioni, comprimendo la complessità del ragionamento giuridico. Ne deriva un problema non solo epistemico ma procedurale: che status attribuire a tali strumenti? Sono mezzi di prova o ausili cognitivi soggetti a *gatekeeping* metodologico (sul modello Daubert) e a obblighi di *disclosure*?

⁵¹ PASQUALE, *New Laws of Robotics: Defending Human Expertise in the Age of AI*, cit.

⁵² DIAKOPOULOS, *Algorithmic Accountability: Journalistic investigation of computational power structures*, in *Digital Journalism*, 3(3), 2015, 398-415.

⁵³ GROTH, NGAN, HANAOKA, *Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects*, in *NISTIR 8280*, 2019.

⁵⁴ ZERILLI, et al. *Transparency in Algorithmic and Human Decision-Making: Is There a Right to Explanation?*, in *Philosophy & Technology*, 32, 2019, 661-677.

Nel silenzio di molti ordinamenti, la domanda cruciale diventa: chi seleziona il selezionatore? Se l'algoritmo pesa, combina e filtra senza traccia della propria catena inferenziale, il contraddittorio rischia di essere svuotato a monte. Gillespie descrive gli algoritmi di classificazione come curatori invisibili⁵⁵: decidono cosa vedere (e cosa non vedere), producendo un effetto di normalizzazione dell'universo probatorio. Per questo occorrono presidi minimi: *audit trail* completo della *pipeline*⁵⁶, accesso simmetrico per la controperizia, spiegabilità situata e oneri di motivazione tecnica proporzionati all'impatto⁵⁷. In sintesi, la meta-prova algoritmica introduce un potere di *gatekeeping* che va normato: l'IA deve essere qualificata come strumento cognitivo sottoposto a regole di *accountability* e contestabilità; solo così il giudice può integrare l'output nella propria motivazione senza scivolare nella ratifica tecnica.

3.4. *Oltre Bayes? Per una epistemologia computazionale non riduzionista del giudizio probatorio.* Nonostante la seduzione esercitata dai modelli matematici sulla cultura giuridica contemporanea, la regola di Bayes rimane largamente inapplicata nella prassi giudiziaria, specie nei sistemi di *civil law*. Le ragioni sono note: la *Bayes Rule* presuppone un aggiornamento quantitativo delle credenze atomistico e sequenziale, difficilmente compatibile con la valutazione olistica e integrativa tipica del giudizio probatorio⁵⁸. Questa diffidenza non equivale a un rifiuto della matematica, ma a una critica della decontestualizzazione: modelli che riducono la prova a frequenza statistica rischiano di dissolvere relazioni semantiche, concatenazioni causali e *standard* pubblici di ragionevolezza⁵⁹. Il giudice non lavora con meri pesi numerici, bensì con soglie di plausibilità argomentativa che devono essere motivate e contestabili tutto considerato⁶⁰.

⁵⁵ RASO, HILLIGOSS, KRISHNAMURTHY, BAVITZ, KIM, *Artificial Intelligence & the Law of Public Administration*, Harvard JOLT Special Symposium, 2018.

⁵⁶ GILLESPIE, *The Relevance of Algorithms*. In GILLESPIE, BOCZKOWSKI, FOOT (Eds.), *Media Technologies*, 2014, Cambridge (MA), 167-194.

⁵⁷ SELBST, BAROCAS, *The intuitive appeal of explainable machines*, cit.

⁵⁸ FENTON, NEIL, BERGER, LAGNADO, TARONI, (Eds.), *Bayes and the Law*. Oxford, 2016.

⁵⁹ TARONI, AITKEN, GARBOLINO, BIEDERMANN, *Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science* (2nd ed.), Weinheim, 2014.

⁶⁰ AMAYA, *The Tapestry of Reason: An Inquiry into the Nature of Coherence and its Role in Legal Argument*, London, 2015.

Come avverte Susan Haack, l'applicazione cieca della *Bayes Rule* favorisce una mitologizzazione del calcolo, che occulta presupposti epistemici sotto una patina di neutralità formale: i modelli probabilistici possono irrigidire il giudizio se non si accompagnano a trasparenza sugli a priori e a condizioni effettive di contraddittorio e rivedibilità⁶¹. In questa scia, il garantismo epistemico segnala il rischio di sostituire la giustificazione pubblica con la sua simulazione computazionale⁶².

Occorre però evitare il luddismo metodologico: non tutte le inferenze computazionali sono riduzioniste. La linea di frattura non è *Bayes sì/Bayes no*, ma tra modelli che pretendono di sostituire la razionalità discorsiva e modelli che la supportano rendendo esplicito l'aggiornamento delle ipotesi e auditabile la pesatura degli indizi⁶³.

Sul versante dell'IA generativa, il *deep learning* non opera come aggiornamento bayesiano esplicito: è un sistema adattivo che apprende pattern da dati grezzi ottimizzando pesi via retro-propagazione; è quindi sub-simbolico e spesso opaco⁶⁴.

L'intelligenza artificiale generativa non ragiona per sillogismi, né applica regole logiche nel senso formale del termine. La sua forza risiede in un'altra modalità cognitiva: l'euristica. Non si tratta di una semplificazione, ma di un paradigma alternativo. L'euristica è un'arte dell'orientamento, una capacità di trovare soluzioni operative in condizioni incerte, incomplete, ambigue. In questo senso, la generatività linguistica dei modelli come GPT non è una logica dimostrativa, ma una *strategia pragmatica di senso*.

La macchina non deduce, ma costruisce. E costruisce in modo situato, adattivo, relazionale. Ogni risposta non è la conseguenza di una premessa formale, ma il risultato di una traiettoria inferenziale attivata da un contesto linguistico e pragmatica. La sua «intelligenza» si misura nella coerenza dinamica e nell'efficacia comunicativa, non nella verità o nella correttezza sintattica.

⁶¹ HAACK, *Legalizzare l'epistemologia*, cit.

⁶² IASEVOLI, *Giudizio e metodo epistemologico garantista*, in *Sistema Penale*, 11/2022, 5-20.

⁶³ FENTON, NEIL, *Risk Assessment and Decision Analysis with Bayesian Networks*, Boca Raton (FL), 2013.

⁶⁴ GOODFELLOW, BENGIO, COURVILLE, *Deep Learning*, Cambridge (MA), 2016.

In questo, si avvicina paradossalmente a certe intuizioni dell'ermeneutica filosofica: l'interpretazione non è applicazione di regole, ma circolo dinamico tra parte e tutto, tra frase e contesto. Il modello, pur senza comprendere, simula un processo interpretativo. Non per introspezione, ma per calcolo strutturato. Non per coscienza, ma per inferenza.

Ciò che colpisce, in questi sistemi, è proprio la coerenza emergente: il modo in cui producono significati plausibili anche in assenza di conoscenza simbolica del mondo. È un'ermeneutica cieca, ma efficace. E questo è, forse, l'aspetto più filosoficamente inquietante della questione.

La sua filosofia resta tuttavia statistica: inferire comportamenti attesi da distribuzioni empiriche, più vicina alla correlazione che alla causalità⁶⁵. Esiste, è vero, un filone di *Bayesian Deep Learning* che tenta di coniugare reti neurali e trattamenti espliciti dell'incertezza⁶⁶, ma la pratica giudiziaria incontra qui limiti di spiegabilità e oneri di validazione.

Dunque, non è Bayes a essere inaccettabile per sé, bensì il riduzionismo epistemico. L'epistemologia della prova deve rimanere critica, pragmatista e contestuale: ogni evidenza è parte di una narrazione giustificabile in pubblico; gli strumenti computazionali vi entrano come ausilio soggetto a trasparenza, contestabilità e responsabilità. In concreto: Bayes come *scaffolding* per esplicitare aggiornamenti; Deep Learning come motore di scoperta di pattern; il giudice come garante della motivazione⁶⁷.

La domanda, allora, non è «Bayes va bene o no?», ma: come usare strumenti inferenziali computazionali senza rinunciare al principio di giurisdizione e all'etica della motivazione giudiziaria? La risposta richiede un disegno istituzionale che integri spiegabilità situata, audit trail e controperizia simmetrica, preservando la natura pubblica e contestabile del giudizio di fatto.

4. *Frizioni epistemiche e paradosso della spiegabilità.* L'introduzione dell'intelligenza artificiale nella valutazione della prova produce una tensione profonda tra due modelli epistemici. Il giudizio umano si struttura entro una

⁶⁵ PEARL, MACKENZIE, *The Book of Why*, New York (NY), 2018.

⁶⁶ BLUNDELL, CORNEISE, KAVUKCUOGLU, WIERSTRA, *Weight Uncertainty*, in *Neural Networks*, in *ICML*, 2015.

⁶⁷ HILDEBRANDT, *Smart Technologies and the End(s) of Law*; cit.

cornice narrativa, contestuale e motivata, fondata su coerenza, cumulatività e verosimiglianza ragionata⁶⁸. Il modello algoritmico, invece, elabora correlazioni ad alta dimensionalità, ottimizza funzioni-obiettivo e restituisce classificazioni/predizioni spesso opache sul piano giustificativo⁶⁹. Da qui tre linee di frizione e un paradosso normativo.

Il giudizio probatorio, nella sua struttura profonda, si fonda su una logica della plausibilità, che valuta le prove secondo coerenza narrativa, continuità semantica e coesione inferenziale tra fonti eterogenee. È un processo di sintesi che si sviluppa in un quadro pragmatico, fallibilista, dove il giudice costruisce uno scenario verosimile «tutto considerato». In questo senso, la prova è *una costruzione epistemica argomentata*, non una semplice somma di evidenze.

L'intelligenza artificiale, in particolare nei modelli di *deep learning*, elabora le informazioni secondo una logica connessionista e bayesiana implicita, che individua pattern nei dati, classifica *input*, produce score di rischio o di probabilità. Ma questa classificazione è spesso atomistica: non tiene conto della coerenza semantica del contesto, ma aggrega correlazioni statistiche anche deboli, frammentarie, ridondanti. L'IA massimizza metriche (accuratezza, AUROC, F1-score) rispetto a una *loss function* definita a monte.

Il risultato è una decisione computazionale che può funzionare predittivamente, ma è incapace di rendere conto della struttura logica della prova in senso giuridico.

È una asimmetria di *telos*: la giurisdizione chiede giustificazione; l'algoritmo offre prestazione⁷⁰. Senza una mediazione istituzionale, la metrica rischia di spiazzare gli *standard* motivazionali e di ri-semantizzare categorie giuridiche in punteggi o soglie⁷¹.

Nel diritto, la coerenza è un principio inferenziale e pubblico: la tesi è tanto più credibile quanto più lega le evidenze e resiste a obiezioni⁷². L'ottimizzazione predittiva è, invece, strumentale e tecnica: ciò che conta è l'errore atteso. Questa divergenza non rende l'IA «illegale», ma impone vin-

⁶⁸ TUZET, *Assessment criteria or standards of proof?*, cit.

⁶⁹ BURRELL, *How the machine "thinks"*, cit.

⁷⁰ WELLER, *Transparency: Motivations and Challenges*, in *Explainable AI*, Berlin, , 2019.

⁷¹ ZARSKY, *The trouble with algorithmic decisions*, cit.

⁷² AMAYA, *The Tapestry of Reason*, cit.

coli: la prestazione deve essere ri-ancorata a criteri di rilevanza, affidabilità e oneri di ragione compatibili con la struttura discorsiva del processo⁷³.

Più il modello è performante, meno è spiegabile: è il paradosso XAI (*Explainable Artificial Intelligence*). L'opacità non dipende soltanto da mancanze comunicative, ma da rappresentazioni interne che resistono all'interpretazione⁷⁴. Ciò non implica illegittimità: la legittimazione deriva da auditabilità, contestabilità e *reason-giving* proporzionato allo scopo⁷⁵. Il giudice stesso non è una «lavagna trasparente»: la motivazione spesso razionalizza *ex post*⁷⁶. Il rimedio non è la trasparenza «totale», ma una spiegabilità situata: dati usati, limiti, tassi d'errore, alternative scartate, condizioni d'impiego⁷⁷.

Un argomento radicale sostiene l'equivalenza tra *black box* e illegittimità epistemica⁷⁸. Tale tesi assolutizza la spiegabilità *ex ante*. Ma la giustizia non si esaurisce nella trasparenza dei mezzi: si fonda sulla ragionevolezza degli esiti e sulla loro contestabilità pubblica⁷⁹. La misura non è «capire tutto del modello», ma poterlo criticare efficacemente.

Se l'IA operasse solo a livello di sintassi o statistica, i risultati sarebbero banalmente spiegabili. In realtà, nel profondo neurale agisce una forma di inferenza sub-simbolica non riducibile alla logica proposizionale, capace di produrre esiti intelligibili negli effetti ma non sempre tracciabili nei passaggi⁸⁰.

Il *refrain* «l'IA è solo sintassi» non regge più come confine concettuale. La tradizione della semantica distribuzionale⁸¹ e, più in generale, la linguistica dell'uso e delle *construction grammars*⁸² mostrano che le regolarità formali dell'uso veicolano significati: forme ricorrenti sono accoppiamenti forma-funzione. Nel dominio formale, la *Montague grammar* ha reso computabile

⁷³ PASQUALE, *New Laws of Robotics*, cit.

⁷⁴ LIPTON, *The mythos of model interpretability*, in *Queue*, 16(3), 2017.

⁷⁵ SELBST, BAROCAS, *The intuitive appeal of explainable machines*, cit.

⁷⁶ MILLER, *Explanation in AI: Insights from the social sciences*, in *Artificial Intelligence*, 2019, 267.

⁷⁷ DOSHI-VELEZ, KIM, *Towards a rigorous science of interpretable ML*, in *arXiv:1702.08608*, 2017.

⁷⁸ ERCOLE, *Contro la giustizia predittiva*, cit., 114.

⁷⁹ SELBST, BAROCAS, *The intuitive appeal of explainable machines*, cit.

⁸⁰ GOODFELLOW, I., BENGIO, Y., & COURVILLE, A., *Deep Learning*, cit.

⁸¹ FIRTH, *Papers in Linguistics 1934-1951*, Oxford, 1957.

⁸² FILLMORE, *Frame Semantics. Linguistics in the Morning Calm*, Seoul, 1982.

la composizione semantica, mostrando che la semantica non è un «altrove» ineffabile rispetto alla struttura⁸³.

I modelli *Transformer* radicalizzano questa intuizione: l'attenzione multi-testa seleziona *pattern* di dipendenza a lungo raggio, apprendendo rappresentazioni latenti dove sintassi e pragmatica si co-strutturano. La finestra referenziale (Marr)⁸⁴ – che Schank riteneva inaccessibile ai codici binari – è oggi operata come finestra di attenzione dinamica: non un singolo varco sul reale, ma un fascio di pesature contestuali in grado di approssimare ruoli, *frames* e *script*⁸⁵ senza ricorrere a simboli espliciti.

Ne segue una tesi più sobria e, insieme, più impegnativa: la semantica non è negata dal calcolo, bensì estratta dall'uso e compressa in geometrie latenti. Dire che «è solo sintassi» equivale a ignorare che, in linguaggio naturale, la sintassi è già semantica perché è uso socialmente investito di funzioni.

Se spostiamo il baricentro dalla comprensione presunta alla giustificazione pubblica (coerenza, pertinenza, contestabilità), IA ed ermeneutica giudiziaria risultano due tecniche di produzione di ragioni. La macchina non «capisce» come un soggetto, ma fornisce giustificazioni-candidate- sotto forma di pattern, analogie e predizioni; l'ermeneuta non «trascende» il calcolo, ma lo integra in una prassi di validazione istituzionale. La differenza normativa non è metafisica (chi capisce davvero?), bensì procedurale: come rendiamo contestabili, spiegabili e rivedibili gli esiti⁸⁶.

In questa prospettiva, la riduzione dell'IA a una competenza puramente sintattica non equivale a una totale cecità semantica, ma va letta come una forma di razionalità operativa alternativa, capace di produrre inferenze formalmente fondate, seppur non motivate secondo criteri intenzionali o narrativi. A conferma di ciò, l'assenza di una sintassi trasparente nel linguaggio giuridico tradizionale – come osservato da Cordero nella sua critica al «lessico opaco, ermetico, criptico e insinuante» di molte sentenze⁸⁷ – costituisce una delle cause principali dell'indeterminatezza argomentativa del giudizio umano e

⁸³ MONTAGUE, *Formal Philosophy*, Yale, 1974.

⁸⁴ MARR, *Vision*, New York (NY), , 1982.

⁸⁵ SCHANK, ABELSON, *Scripts, Plans, Goals, and Understanding*, Hillsdale (NJ), 1977.

⁸⁶ DIVER, *Digisprudence: Code as Law Rebooted*, Edinburgh 2022.

⁸⁷ CORDERO, *Stilus curiae (analisi della sentenza penale)*, In AA.VV. *La sentenza in Europa. Metodo, tecnica e stile*, Padova, 1988, 293–312.

della sua resistenza alla razionalizzazione esplicita. L'opacità sintattica, in tal senso, alimenta un effetto di chiusura epistemica che può degenerare in arbitrarietà e autoreferenzialità.

Ne consegue che, paradossalmente, l'impiego di architetture computazionali dotate di coerenza sintattica interna — per quanto prive di una semantica intenzionale — può rappresentare un vettore di trasparenza inferenziale e di ricomposizione formale del giudizio, in contesti in cui la sintassi giuridica tradizionale si è resa opaca, stereotipata o tautologica. L'intelligenza artificiale, dunque, pur restando su un piano non narrativo, può contribuire alla chiarificazione strutturale del ragionamento giudiziale, servendo da dispositivo di controllo e da catalizzatore di razionalità argomentativa.

In quest'ambiguità epistemica si inserisce il valore positivo del *deep learning*: non come modello esplicativo, ma come *pensiero statistico della possibilità*. Come ha notato Judea Pearl, il problema non è solo l'assenza di causalità nei modelli connessionisti, ma la loro incapacità di articolare perché un evento si è verificato⁸⁸. Ma, è proprio nel non dire *perché*, ma nell'indicare il *quanto*, che il DL eccelle. E questa eccellenza sfida ogni epistemologia giuridica fondata sulla *ratio* e sul logos.

Nel confronto tra giudizio umano e inferenza computazionale, il ruolo della sintassi emerge come punto di snodo epistemologico. Se l'intelligenza artificiale, in particolare nei modelli neurali profondi come i *Transformer*, opera sulla base di correlazioni statistiche tra sequenze linguistiche, è pur vero che tali correlazioni si fondano su una struttura sintattica latente, appresa attraverso meccanismi di auto-attenzione e modellazione del contesto. La «finestra di attenzione», che sostituisce l'analisi simbolica tradizionale, consente di cogliere regolarità grammaticali e strutture sintattiche implicite, senza che queste vengano formalizzate esplicitamente. Si tratta dunque di una semantica computazionale emergente, che non postula significati stabili ma genera inferenze a partire da *pattern* distribuiti e contesti dinamici.

Questa forma di razionalità, pur non accessibile all'introspezione umana né riconducibile ai canoni argomentativi classici, si rivela tuttavia dotata di coerenza inferenziale e capacità predittiva. Al contrario, nel linguaggio giuridico tradizionale, l'opacità sintattica descritta da Cordero — fatta di anacoluti, formule stereotipate, lessico ermetico e clausole tautologiche — ostacola la possi-

⁸⁸ PEARL, *Causality* (2nd ed.), cit.

bilità stessa di una ricostruzione logico-inferenziale trasparente. Ciò produce un paradosso normativo: l'elaborazione sintattica artificiale, per quanto priva di intenzionalità e di semantica forte, può risultare epistemicamente più accessibile e verificabile della retorica giudiziaria convenzionale, spesso refrattaria alla formalizzazione.

Ne deriva che la presunta inferiorità epistemica dell'IA, fondata sull'assenza di comprensione semantica, va contestualizzata. Non è la coscienza intenzionale o la narrazione a garantire la validità del giudizio, ma la tracciabilità e la contestabilità delle sue premesse inferenziali. In tal senso, l'IA può fungere da sintassi ausiliaria della giustizia, offrendo una griglia strutturale contro l'indeterminatezza e contribuendo a superare quella «opacità paranoica» del giudizio umano che, secondo Cordero, affonda le sue radici proprio nella mancanza di una sintassi razionale condivisa.

Da qui il paradosso: una sintassi artificiale coerente — pur senza semantica intenzionale — può talvolta risultare più verificabile della retorica giudiziaria ordinaria. Come ricorda Jackendoff⁸⁹, la sintassi è un'interfaccia che abilita inferenze: un'organizzazione formale che può chiarire relazioni evidenziali anche senza «comprendere» come un umano.

In sintesi, la frizione tra giudice e algoritmo non è solo normativa o politica: è una frizione tra *ontologie cognitive*. Il giudice chiede coerenza narrativa e spiegabilità simbolica. L'algoritmo offre correlazione, ottimizzazione e risultati. Nel mezzo, la crisi del diritto come *istituzione del senso*.

Per ricomporre le frizioni occorrono regole di compatibilità:

1. *Qualificazione*: l'IA è ausilio cognitivo (meta-prova), non prova piena, salvo disciplina espressa.
2. *Accountability tecnica*: *model cards*, *datasheets*, *audit trail* riproducibili, controllo difensivo⁹⁰.
3. *Spiegabilità situata*: ragioni operative, limiti, tasso d'errore, ambiti di non-uso⁹¹.
4. *Causalità e contesto*: quando gli effetti giuridici dipendono da nessi causali, si richiedono modelli espliciti⁹² o protocolli che leghino l'*output* alle ipotesi rivali e agli *standard* di prova⁹³.

⁸⁹ JACKENDOFF, *Foundations of Language*, Oxford, 2002.

⁹⁰ MITCHELL, *Model Cards for Model Reporting*, cit.

⁹¹ SELBST, BAROCAS, *The intuitive appeal of explainable machines*, cit.

5. *Gerarchia di ruoli: experta docet, iudex motivat*: l'IA istruisce, il giudice giustifica e si assume la responsabilità pubblica⁹⁴.

5. *Prospettive: ibridazione o incompatibilità?* Le tensioni fin qui analizzate non impongono esiti apocalittici. L'innesto dell'IA nella dinamica probatoria introduce criticità epistemiche e democratiche, ma non dissolve per ciò solo la razionalità giudiziaria: chiede piuttosto di ridefinire i confini tra computazione e giurisdizione e di progettare modelli ibridi in cui l'IA svolga un ruolo ausiliario, controllato e contestabile⁹⁵.

In tale configurazione l'IA opera come strumento di supporto alla *quaestio facti*: analisi preliminare dei dati, emersione di *pattern* ricorrenti, *document review*, comparazione di precedenti, verifica della coerenza statistica di ipotesi alternative. Rimane invece non delegabile la funzione deliberativa e *reason-giving*, nucleo irriducibile della giurisdizione⁹⁶.

Ne discende l'esigenza di una epistemologia ibrida della prova, dove l'algoritmo fornisce mappe probabilistiche e il giudice rimane «cartografo ermeneutico»: interroga gli esiti, li integra nel contesto, li espone al contraddittorio e ne risponde in motivazione. È la traiettoria di una digisprudenza responsabile⁹⁷, in dialogo con il *rule-of-law by design*⁹⁸: tecnologie progettate per la contestabilità, la trasparenza proporzionata e la legittimazione reciproca tra umano e artificiale.

Perché ciò sia possibile, l'IA va ricontestualizzata: non giudice sostitutivo, ma «attante tecnico»⁹⁹ sottoposto a vincoli epistemiche e giuridici¹⁰⁰. Ciò implica, da un lato, formazione epistemica dei giuristi; dall'altro, progettazione responsa-

⁹² PEARL, *Causality* (2nd ed.), Cambridge 2009.

⁹³ TUZET, *Assessment criteria or standards of proof?*, cit.

⁹⁴ PASQUALE, *New Laws of Robotics*, cit.

⁹⁵ HILDEBRANDT, *Algorithmic Regulation and the Rule of Law*, in *Phil. Trans. Royal Society A*, 376, 2018, 2128.

⁹⁶ AMAYA, *The Tapestry of Reason*, cit.

⁹⁷ DIVER, *Digisprudence: Code as Law Rebooted*, cit.

⁹⁸ HILDEBRANDT, *Law for Computer Scientists and Other Folk*, cit.

⁹⁹ TEUBNER, *Ibridi ed attanti. Attori collettivi ed enti non umani nella società e nel diritto*, Milano-Udine, 2015.

¹⁰⁰ KROLL et al., *Accountable Algorithms*, cit.

bile dei sistemi (spiegabilità situata, auditabilità, diritto alla motivazione, proporzionalità d'uso). Solo a queste condizioni prende forma una *razionalità giudiziaria aumentata*: l'IA non surroga la coscienza ermeneutica del giudice, ma accompagna la ricerca della verità senza sostituirla¹⁰¹.

5.1 *Epistemologie ibride: la «co-prova» tra giudice e IA*. L'integrazione dell'IA nella valutazione della prova sembra opporre giudizio contestuale e classificazione statistica, coerenza narrativa e ottimizzazione predittiva. L'alternativa secca, però, è fuorviante: si apre lo spazio per una co-prova in cui giudice e IA co-operano in modo asimmetrico ma non subalterno.

Per «co-prova» intendiamo una costruzione probatoria distribuita: il modello neurale produce ipotesi inferenziali (*pattern, ranking*, anomalie) che il giudice traduce, verifica e sottopone a contraddittorio, inserendole in un percorso pubblicamente giustificabile. L'IA non è testimone né giudice, ma co-attore epistemico la cui voce necessita di traduzione e controllo.

Questa co-prova richiede una doppia condizione: (i) competenza epistemica del giudice sui principi e limiti dell'IA; (ii) *design* dei sistemi conforme a *explainability, auditability* e contestabilità difensiva¹⁰². La giusta distanza – né prossimità tecnica che assorbe la motivazione, né separazione che rifiuta l'ausilio – evita l'*udienza algoritmica* e preserva la vocazione deliberativa del processo.

Operativamente, un *protocollo minimo di co-prova* richiede i seguenti passaggi:

1. *qualificazione* dello strumento come meta-prova (non prova piena) e indicazione dello scopo consentito.
2. *disclosure* tecnica proporzionata: dati, metriche, errori, *non-use cases*, *model card* e *datasheet*.
3. *audit trail* riproducibile e controperizia simmetrica.
4. *ancoraggio* agli *standard* di prova: come l'*output* incide su oltre ogni ragionevole dubbio / preponderanza ecc.
5. *reason-giving* situato nella motivazione: perché l'*output* è rilevante/affidabile nel caso concreto.

¹⁰¹ PASQUALE, *New Laws of Robotics*; cit.

¹⁰² SELBST, BAROCAS, *The intuitive appeal of explainable machines*; cit.

6. *Valvole di sicurezza: off-switch* procedurale, *human-in-the-loop*, limitazioni d'uso in ambiti causalmente sensibili.

6. *Verità computate. Le evidenze digitali tra tecnica forense e crisi del giudizio.* Le evidenze digitali non sono meri residui informatici, ma atti computazionali prodotti in ambienti mediati da architetture tecniche che ne condizionano forma, contenuto e accessibilità. Come ricorda Casey, la *digital evidence* «non è auto-rivelante»: va interpretata e validata nel contesto¹⁰³. In questa prospettiva, l'unità epistemica non è più il «fatto» in senso ingenuo, bensì la sequenza algoritmica che lo rende inferibile.

Ne deriva una de-contestualizzazione ontologica della traccia: a differenza della prova materiale, quella digitale non si dà quasi mai in forma originaria, ma come evento computato. Il passaggio dalla *chain of custody* alla *digital chain of inferences*¹⁰⁴ sposta l'asse dal mero trasferimento documentale alla mediazione tecnica dell'accesso, conservazione e interpretazione. L'*autenticità* non è un dato, ma l'esito di un percorso inferenziale e procedurale.

In questo ambiente, l'algoritmo diventa attore epistemico: seleziona, ordina, pesa e valorizza le evidenze. Sistemi di *scoring*, *anomaly detection* o *matching* operano come meta-prove: non aggiungono solo dati, ma strutturano l'universo di ciò che conta come prova. La prova diventa così indiretta due volte: traccia mediata + filtro computazionale. L'opacità che ne consegue — *delegated seeing* — è un limite non solo tecnico ma epistemico e democratico¹⁰⁵.

Di qui il rischio di un contraddittorio svuotato: quando il percorso computazionale è indecifrabile, la dialettica difensiva scivola da critica sul contenuto a critica sullo strumento, spesso senza mezzi effettivi per contestarlo. Le note riserve di Pearl verso il connessionismo — assenza di ragionamento causale e inadeguatezza per domande controfattuali — rendono evidente come, senza

¹⁰³ CASEY, *Digital Evidence and Computer Crime* (3rd ed.), San Diego (CA), 2011.

¹⁰⁴ NATH, SUMMERS, BAEK, AHN, *Digital Evidence Chain of Custody: Navigating New Realities of Digital Forensics*, 2024 IEEE 6th International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (TPS-ISA), Washington (DC), 2024, 11-20.

¹⁰⁵ HILDEBRANDT, *Law for Computer Scientists and Other Folk*; cit.

modelli di causalità esplicita, si finisca per giudicare configurazioni statistiche più che fatti¹⁰⁶.

L'aspirazione alla verità *by design* promette rimedi (*log, hash, secure provenance*, tracciabilità). Ma, se assolutizzata, rischia di erodere ciò che il processo tutela: comprensibilità, fallibilità, contestabilità. Serve allora una *epistemologia forense computazionale* che riconosca la non neutralità della traccia digitale e domesticchi la mediazione algoritmica: non oracolo, non servo muto, bensì co-attore trasparente e rendicontabile.

In sintesi, la prova digitale è strutturalmente mediata e la meta-prova algoritmica ne condiziona l'emersione. Senza strumenti di *audit*, spiegabilità situata e garanzie procedurali effettive, il rischio è duplice: crisi della contestualità e svuotamento del contraddittorio. L'onere del diritto è ricondurre l'*evento computato* entro forme di giustificazione pubblica.

7. Scardinamenti positivi. Verso una computazione emancipante della prova.

Se l'epistemologia computazionale applicata alla prova penale solleva interrogativi radicali in termini di opacità, giustificabilità e contestualizzazione, sarebbe miope ridurre il suo impatto a una semplice minaccia epistemonica. Al contrario, una riflessione teoricamente più sofisticata può riconoscere come l'intelligenza artificiale — e in particolare il *deep learning* forense — possa talvolta costituire una risorsa epistemonica, garantista e persino emancipante.

In primo luogo, l'adozione di strumenti computazionali nella *digital forensics* ha già trasformato il modo in cui la prova viene identificata, preservata, organizzata e analizzata.

Come ha mostrato Eoghan Casey¹⁰⁷, la natura digitale della prova consente operazioni un tempo impensabili: il recupero di frammenti cancellati, l'estrazione di *metadata* nascosti, la ricostruzione di catene temporali coerenti, la validazione tramite hash. Questa precisione ontologica della prova computata — se ben gestita — rafforza l'affidabilità del processo e riduce l'alea interpretativa.

Ma il salto qualitativo avviene quando si passa dalla computazione come strumento tecnico alla computazione come agente epistemonico. Laddove l'essere umano tende a operare su base narrativa e inferenziale (spesso con-

¹⁰⁶ PEARL, *Causality* (2nd ed.); cit.

¹⁰⁷ CASEY, *Digital Evidence and Computer Crime*; cit.

taminata da *bias* di conferma, effetto *framing*, ancoraggio iniziale), la macchina può offrire un processo di correlazione e valutazione dei dati meno suscettibile alle distorsioni cognitive. In casi come Garlasco, una modellizzazione bayesiana ausiliaria nelle fasi iniziali avrebbe potuto segnalare incoerenze e debolezze indiziarie, fungendo da *early-warning* epistemico — senza sostituire il giudizio. Ciò non implica una superiorità automatica del calcolo probabilistico, ma invita a riflettere sull'eventuale contributo di strumenti computazionali alla razionalità procedurale e alla prevenzione dell'errore giudiziario.

Inoltre, l'adozione di tecniche di *probabilistic reasoning* (*Bayesian Networks*) o di *explainable AI*, se orientata da criteri epistemici forti, può contribuire a rinforzare l'*accountability* del giudizio. Hildebrandt insiste sul concetto di *computational legal transparency*, ovvero la possibilità che la trasparenza sia costruita nel codice stesso, a livello di architettura e *governance*: *Model cards*, *datasheets*, *provenance logs*, *versioning* dei modelli, valutazioni *ex ante* del tasso d'errore, *stress-testing* su *edge cases*, e *audit* indipendenti rendono l'*output* contestabile e ripetibile¹⁰⁸.

Un altro aspetto spesso ignorato è la capacità della macchina di rendere conto di relazioni latenti, non visibili all'occhio umano. Integrare reti neurali con grafici causali e strumenti bayesiani calibrati consente verifiche *what-if* e spiegazioni pertinenti allo scopo — non la trasparenza totale del «codice», ma ragioni utili a chi deve deliberare¹⁰⁹.

Attraverso tecniche di *feature extraction* e *pattern detection*, l'IA può individuare connessioni indiziarie non esplicitamente formulate, aprendo così spazi di ipotesi nuovi, che il giudice potrà — e dovrà — verificare in sede motivazionale. In tal senso, l'opacità della *black box* può essere trasformata in un vincolo euristico per il giudice, che si ritrova nella posizione di dover giustificare (e non subire) il suggerimento algoritmico.

Infine, va riconosciuto che l'epistemologia computazionale, per quanto aliena ai canoni tradizionali della logica giuridica, non è affatto priva di razionalità. Come ha notato Judea Pearl, anche i modelli connessionisti operano inferenze implicite, spesso più sofisticate di quelle umane, soprattutto in relazione alla plausibilità causale e alla costruzione di modelli controfattuali. Se ben calibrata, attraverso regole processuali «*computer-aware*». Ammissione condizio-

¹⁰⁸ GEBRU, *Datasheets for Datasets*, cit.

¹⁰⁹ PEARL, *Causality* (2nd ed.); cit.

nata a: (i) descrizione tecnica del sistema e dati; (ii) tasso noto d'errore e *drift*; (iii) possibilità di contro-perizia; (iv) *logging* forense replicabile; (v) istruzioni al giudicante sullo *status* ausiliario dell'*output* (Daubert-style, adattato al digitale) la macchina può contribuire non a soppiantare ma a rafforzare la cultura della prova — una cultura che non abdica alla spiegabilità, ma che sa anche convivere con nuove forme di razionalità computabile.

In conclusione, l'IA non deve sostituire la motivazione: deve alimentarla. Se trattata come super-perito responsabile — soggetto a regole di spiegabilità proporzionata, *auditability* e contraddittorio effettivo — la computazione può aumentare la razionalità processuale, non comprimerla.

8. *L'algoritmo e il diritto della prova: oltre l'antitesi*. Il confronto tra epistemologia giudiziaria e intelligenza artificiale si rivela come una tensione strutturale tra due forme di razionalità: quella narrativa e contestuale del giudice, e quella computazionale e predittiva del sistema intelligente. Ma questa opposizione, se assunta come rigida, rischia di oscurare il vero problema: non è l'algoritmo in sé a essere incompatibile con il diritto, bensì l'assenza di un modello processuale che ne disciplini l'uso epistemico.

Il giudizio penale si regge su una razionalità della prova fondata su plausibilità, inferenza e coerenza contestuale, in vista di uno standard che, come noto, si colloca «oltre ogni ragionevole dubbio».

Il conflitto fra razionalità narrativa del giudice e razionalità predittiva dell'algoritmo rischia però di essere una falsa alternativa. Se lo *standard* oltre ogni ragionevole dubbio è letto come plausibilità epistemica e non come metrica numerica rigida, allora l'IA può concorrere a stimare quella plausibilità in modo tracciabile, falsificabile e contestabile, entro vincoli processuali. In tale quadro — come mostrano Amaya e Haack — l'algoritmo non sostituisce il giudice, ma ne potenzia la razionalità pubblica: propone ipotesi, segnala incoerenze, rende esplicite assunzioni e *trade-off*. La plausibilità può essere anche computazionale, purché incastonata in una procedura di motivazione e contraddittorio.

Ma se davvero, come è stato sostenuto in dottrina¹¹⁰, lo *standard* oltre ogni ragionevole dubbio viene inteso in chiave di plausibilità epistemica (non come

¹¹⁰ CECCHI, *Indizi, presunzioni, tesi alternative e dubbi ragionevoli al setaccio della certezza processuale. La lectio della Suprema corte nel caso Monachino e altri*, in *Archivio penale*, n. 2, 2025.

soglia numerica rigida), è legittimo chiedersi perché e come strumenti computazionali possano concorrere a stimare tale plausibilità in modo tracciabile e contestabile, senza sostituire il giudizio.

Come nota Amaya¹¹¹, la plausibilità probatoria è una categoria giuridica, ma non è necessariamente antropocentrica. Nulla vieta di pensare che la plausibilità possa essere anche *computazionale*, purché *integrata* in una cornice processuale garantita. In questa prospettiva, l'algoritmo probatorio diventa non un sostituto del giudice, ma un vettore di miglioramento della razionalità processuale, capace di offrire inferenze robuste, falsificabili e contestabili.

Dunque, se la giustizia non coincide con la verità assoluta ma con la migliore approssimazione epistemica possibile entro una procedura legittima, allora è plausibile che i sistemi intelligenti – debitamente regolati e controllati – possano contribuire non a *degiuridicizzare* la prova, ma al contrario a renderla più rigorosa, più riproducibile, e forse anche più giusta.

Come in ogni rivoluzione epistemica, il problema non è l'introduzione di nuovi strumenti, ma l'adattamento delle istituzioni alla loro gestione. L'algoritmo può essere pensato non come un oscuro oracolo, ma come una *protesi razionale* del sistema probatorio. E in questa direzione, una vera *giustizia computazionale* non è utopia, ma progetto.

Nel mondo classico della giurisdizione, il giudizio probatorio è stato inteso come un esercizio umano di equilibrio tra narrazione, logica, esperienza e responsabilità. Ma la comparsa dell'intelligenza artificiale nelle aule giudiziarie, nella prassi investigativa e nella *digital forensics* ha messo in crisi questo equilibrio, spingendo la riflessione epistemologica in un territorio nuovo.

La sfida che abbiamo analizzato in queste pagine non è semplicemente quella di una «nuova tecnologia». È una trasformazione ontologica della prova: da narrazione contestuale a pattern computazionale, da evidenza argomentata a indizio calcolato, da coerenza logica a plausibilità statistica. Il rischio maggiore è che tale trasformazione venga subita dal diritto, nel nome dell'efficienza e dell'automazione.

Il sospetto sistematico verso l'efficienza computazionale — «la giustizia predittiva valuta le decisioni solo per la loro efficienza statistica»¹¹² — semplifica e confonde i piani. La giustizia è anche una questione di probabilità epistemica,

¹¹¹ AMAYA, *The Tapestry of Reason*, cit.

¹¹² ERCOLE, *Contro la giustizia predittiva*, cit., 84.

e l'efficienza non è di per sé un disvalore: lo diventa solo quando viene impiegata per mascherare asimmetrie o per eludere la responsabilità istituzionale.

L'antitesi tra giudice e algoritmo, tra opacità e spiegabilità, tra *human reason* e *machine patterning*, è troppo rozza per rendere giustizia alla complessità in atto. Come abbiamo sostenuto, l'intelligenza artificiale non è (solo) un fattore di rischio, ma può essere uno strumento di riqualificazione epistemica, capace di mettere a nudo i limiti impliciti dell'intuizionismo giudiziario e dei *bias* cognitivi che lo attraversano. L'IA, se incardinata in una *digisprudence* responsabile¹¹³, può diventare non giudice, ma parte della prova: meta-prova, co-prova, vincolo euristico, oggetto contestabile e non oracolo infallibile.

Ma per questo occorre una nuova epistemologia giuridica, capace di integrare modelli inferenziali diversi: la probabilità logica (baconiana) e quella bayesiana, la narrativa e la computazione, la motivazione e la *black box*. Serve una *giustizia aumentata*, non automatizzata. Una giurisdizione capace di pensare l'improbabile, e dunque di valutare l'evidenza anche quando proviene da uno sguardo non umano – purché spiegabile, contestabile, e responsabile.

Nel dibattito contemporaneo sull'uso dell'intelligenza artificiale nel processo penale si è spesso reiterata una dicotomia strutturale tra giudizio umano e inferenza algoritmica, tra valutazione qualitativa e calcolo statistico. Eppure, tale opposizione rischia di rivelarsi fuorviante. Essa si fonda su un malinteso epistemologico: che la plausibilità probatoria sia un attributo esclusivamente umano, dipendente da facoltà narrative, sensibilità ermeneutica e capacità di cogliere il contesto. Ma se si assume, come fanno autori quali Susan Haack o Jordi Ferrer Beltrán¹¹⁴, che la prova si costituisce attraverso gradi di giustificazione razionale e non attraverso verità ontologiche, allora la questione diventa: quali strumenti inferenziali sono in grado di sostenere al meglio tale giustificazione, in modo riproducibile, falsificabile e trasparente?

In questa chiave, l'IA – e in particolare i modelli probabilistici avanzati, che integrano statistica bayesiana, reti neurali e metodi di *causal inference* – può essere intesa non come antagonista del giudizio, ma come potenziale agente epistemico secondario, in grado di cooperare alla costruzione razionale della decisione. La plausibilità computazionale, lungi dall'essere una riduzione

¹¹³ DIVER, *Digisprudence: Code as Law Rebooted*; cit.

¹¹⁴ FERRER BELTRÁN, *La valoración racional de la prueba*; cit.

meccanica del giudizio, può esprimere forme robuste di inferenza induttiva e abduttiva, capaci di valutare scenari, ipotesi, correlazioni e anomalie con una granularità preclusa all'operatore umano.

In effetti, come mostrano Galles e Pearl¹¹⁵, un sistema inferenziale dotato di modelli causali espliciti è capace non solo di descrivere *pattern* nei dati, ma anche di simulare *counterfactuals*, ossia alternative logiche coerenti con il sistema osservato. Questa capacità, cruciale nel diritto penale per costruire narrazioni controfattuali (es.: «cosa sarebbe accaduto se...»), suggerisce che l'intelligenza artificiale può contribuire non alla semplificazione del giudizio, ma al suo rafforzamento controfattuale e razionale, integrando dimensioni predittive e spiegazionali.

D'altra parte, se – come recita lo *standard* processuale anglosassone – la condanna può intervenire solo *beyond a reasonable doubt*, allora è necessario riconoscere che la ragionevolezza del dubbio non implica una soglia mistica o emotiva, ma una *plausibilità epistemica ragionevole*. E, in questo contesto, una stima probabilistica computata al 99% da un sistema trasparente, contestabile e supervisionato, può risultare più fondata del ragionamento intuitivo di un giudice o di una giuria.

Questo non significa affidarsi ciecamente alla “ragione algoritmica”, ma nemmeno persistere nella mitologia dell'irriducibilità del giudizio umano. Il compito attuale della teoria della prova è allora duplice: costruire un quadro epistemologico in cui la plausibilità computazionale sia concepibile, e insieme, istituire una *governance* giuridica che ne regoli l'uso, ne certifichi l'affidabilità, e ne assicuri la contestabilità procedurale.

La giustizia computazionale non è – o non dovrebbe essere – un'abiura del garantismo, ma la sua prosecuzione con altri mezzi. Una giustizia che non rinuncia alla coerenza, ma che la rende computabile.

9. Conclusione: l'algoritmo come istituzione razionale. Se la prova giudiziaria è, nella sua essenza epistemica, una procedura di giustificazione istituzionalmente regolata, allora l'intelligenza artificiale – nella misura in cui produce inferenze controllabili, replicabili e contestabili – può essere concepita non come un semplice strumento, ma come una nuova forma istituzionale di razionalità. Non un giudice alternativo, ma una *funzione epistemica* integrata

¹¹⁵ GALLES, PEARL, *An Axiomatic Characterization of Causal Counterfactuals*, in *Foundations of Science* 3, no. 1, 1998: 151-182.

nell'architettura del processo. Un attore silenzioso, ma normativamente sorvegliato, il cui apporto non è valutativo, bensì inferenziale, orientato non a decidere, ma a istruire la plausibilità delle decisioni.

La *black box* non è allora un oggetto opaco da respingere in blocco, ma un campo di tensione semantica e giuridica, che deve essere reso trasparente attraverso il diritto stesso, come nel paradigma della *contestable AI* proposto da Mireille Hildebrandt e Laurence Diver. L'algoritmo, da questo punto di vista, non può restare al di fuori dell'arena giuridica: deve entrare nella grammatica delle fonti, nei dispositivi della responsabilità e nelle procedure del contraddittorio. Solo così potrà assumere la forma di un'istituzione razionale emergente, capace di contribuire all'orizzonte regolativo e garantista della giustizia. Ma c'è di più. Come ha suggerito Luciano Floridi¹¹⁶, l'informazione è ormai l'ontologia di base delle nostre società: noi siamo attori, pazienti morali e soggetti epistemici *in un'infosfera*. Il giudizio, pertanto, non può sottrarsi alla logica computazionale che innerva la realtà. È nel confronto critico e costruttivo tra giudizio umano e inferenza algoritmica che può emergere una «nuova intelligenza giudiziaria», capace di unire rigore formale, sensibilità garantista e potenza inferenziale.

Il compito dell'epistemologia giudiziaria del XXI secolo sarà dunque duplice: proteggere il giudizio dalla cattiva automazione, ma anche difenderlo dalla nostalgia di un umanesimo cieco alla complessità tecnica del mondo contemporaneo.

¹¹⁶ FLORIDI, *Pensare l'infosfera. La filosofia come design concettuale*, Milano, 2020.