

ANTICIPAZIONI

CRISTIANA VALENTINI

La ‘scoperta’ dei *bias cognitivi* nel processo penale**

Alcune riflessioni sull'intelligenza artificiale forniscono il destro per esplorare un dato ormai non più agibile, ovvero gli errori cognitivi contenuti nei dati di addestramento e dunque, prima ancora, gli errori cognitivi degli esseri umani decisori nei procedimenti penali, che finiscono per alimentare e moltiplicare i *bias* cognitivi della macchina.

The ‘discovery’ of cognitive biases in criminal proceedings

Some reflections on artificial intelligence provide the opportunity to explore an unavoidable fact: the cognitive errors contained in training data and therefore, even before that, the cognitive errors of human decision-makers in criminal proceedings, which end up fueling and multiplying the machine’s cognitive biases.

SOMMARIO: 1. Realtà e fantascienza – 2. IA. Limiti strutturali – 3. *Bias* nei dati di addestramento – 4. Ragionamento o ‘pappagallo stocastico’? – 5. Il giudice e gli algoritmi – 6. Il *bias* cognitivo dell’essere umano: quasi uno sconosciuto per il giurista italiano – 7. Due parole sul *bias* cognitivo, ovvero un primo approccio del processualista – 8. *Bias* e sistema penale – 9. I decisori nel procedimento penale – 10. Regole per i processi decisionali.

1. *Realtà e fantascienza.* A leggere molteplici titoli di recenti studi in materia di rapporti tra sistema penale e intelligenza artificiale, parrebbe di trovarsi sull’orlo di un nuovo mondo, quasi sbilanciati verso le vertiginose prospettive dischiuse da scienziati/scrittori come Isaac Asimov, il cui ultimo volume del notissimo ciclo della Fondazione¹ immagina una realtà nella quale un’intelligenza diversa e superiore in quanto algoritmica ha saggiamente diretto millenni di storia dell’umanità verso il superamento dei suoi intrinseci difetti.

Non è proprio così, nei fatti; si tratta di una percezione parzialmente errata – forse prodotta da certi tratti mediatici con cui è stato divulgato il fenomeno

¹ Un ringraziamento speciale va a mio nipote, Alessandro Di Matteo, imminente laureando in Ingegneria informatica del Politecnico di Torino, grande amante dell’IA e lucido indagatore dei suoi enigmi. Senza di lui, i suoi entusiasmi e le sue brillanti prese di posizione, non avrei mai potuto scrivere questo articolo.

“ L’articolo costituisce un’anticipazione rispetto al Il volume “Giudizio penale e intelligenza artificiale. Una riflessione sistematica”, a cura di Triggiani e Incampo, che sarà pubblicato nel 2026 per i tipi dell’Editore Cacucci (Bari) nella collana “Unità del sapere giuridico. Quaderni di scienze penalistiche e filosofico-giuridiche”.

² Ci riferiamo chiaramente al suggestivo epilogo del ciclo della Fondazione, scritto da I. Asimov nel 1986 e pubblicato in Italia con il nome di *Fondazione e Terra*.

dell'IA - e l'uso dell'intelligenza artificiale per scopi decisionali nel processo penale² parrebbe una realtà in lento divenire, di cui val la pena di curarsi solo a scopo di tempestiva prevenzione o anche a fini elegantemente speculativi, se non fosse per un particolare interessante: la minaccia 'robotica' ci consente di estrarre dall'oggi alcuni dati di realtà nei cui confronti siamo rimasti singolarmente ciechi, ad onta degli ampi studi coltivati altrove da molti anni; ci riferiamo ai *bias* cognitivi, ovvero agli errori che la nostra intelligenza 'non artificiale' compie continuamente, nella vita privata come in quella professionale, spesso con costi sociali notevoli.

Ma su ciò si tornerà a breve. Per il momento, merita un breve intermezzo, integralmente dedicato alle vivaci sorti dell'IA.

2. *IA. Limiti strutturali.* Il termine IA - *Intelligenza Artificiale* - è classicamente usato con ampia approssimazione, in ossequio all'originaria rappresentazione dell'idea di intelligenza non umana suggerita dalla celebre frase di *Marvin Minsky*, considerato uno dei padri fondatori dell'intelligenza artificiale moderna: "AI is The Science of Making Machines do Things That Would Require Intelligence if Done by Humans"³.

Oggi, il riferimento forse più corretto è quello che conferisce la definizione di IA a tutta una serie di modelli matematici basati su antenati comuni, le cosiddette *Neural Networks*, ovvero modelli che si basano su strutture complesse e modulari ispirate dalla biologia.

Precisiamo: allo stato i modelli statistici più tradizionali, tutt'ora usati al di fuori dello scenario noto al grande pubblico, venivano (e vengono) progettati per modellare entro una certa approssimazione una distribuzione reale partendo da formule note, e l'addestramento - *training* - serve semplicemente a trovare i parametri migliori, ovvero quelli per cui l'approssimazione è nel miglior caso trascurabile.

Con lo sviluppo delle reti neurali, invece, in particolare delle *DNN - Deep Neural Networks*, si è passati a modelli dove la ricerca non è focalizzata sul reperimento diretto di una formula di *input-output* complessiva, quanto

² La prospettiva cambia di molto allorquando ci si confronti, invece, sull'uso degli LLM in sede di interpretazione. Si leggano, in questa prospettiva, le brillanti riflessioni di DI GIOVINE, *Interpretazione umana e robotica (breve riflessione sul possibile ingresso dei Large Language Models nella giurisdizione)*, in www.sistemapenale.it, 4 settembre 2025. Merita leggere, sul punto, in prospettiva processuale, MAZZA, *Distopia del processo artificiale*, in Arch. pen. (web), 7 gennaio 2025.

³ MINSKY, *Semantic information processing*, The MIT Press, Cambridge, Massachussets, 1969.

piuttosto sullo sviluppo di elementi più piccoli detti *perceptrons*, ispirati al funzionamento biologico dei neuroni e adoperati come mattoncino di base per creare reti estremamente complesse che stratificano milioni di tali elementi al fine di ottenere una plasticità enorme.

L'esito immediatamente intuibile è che la funzione complessiva di *input-output* dipende da tutti i parametri dei milioni di neuroni di cui la rete è composta: mentre nei modelli statistici tradizionali i parametri sono relativamente pochi e possiedono unicamente lo scopo di compiere il cd. *fitting* del modello sui dati reali (adattamento ai dati), qui invece i parametri sono in numero tale che essi stessi vanno a definire il modello, rendendolo quindi estremamente flessibile, ma al contempo privo di una formulazione matematica globale significativa.

È proprio per questo motivo che i modelli moderni di IA vengono spesso definiti '*Black Box*'; non certo a causa di un fantomatico velo oscuro lato re di presagi foschi, ma piuttosto quale semplice conseguenza di formule matematiche estremamente complesse e dense di coefficienti fissati a valori apparentemente privi di significato, certo spogliandoli dell'eleganza matematica che possiedono i modelli statistici tradizionali, ma al contempo permettendo una flessibilità enorme, tale da adattarsi ai casi più disparati.

In buona sostanza, questi modelli assumono decisioni senza però essere in grado di fornire accesso al percorso causale che collega *input* e *output*: la decisione è una funzione distribuita su milioni di parametri ad oggi non tracciabili, men che meno in termini, ad esempio, logico-giuridici.

Il significato è chiaro; la potenza predittiva si è impennata, ma al costo della trasparenza della decisione.

3. Bias *nei dati di addestramento*. A partire dalla non-interpretabilità *in primis* logico/matematica degli *output* di tali algoritmi, l'esito è quello di una più agevole propagazione di *bias* ed errori sistematici, portando avanti o addirittura amplificando fenomeni di *disparate impact*; com'è stato efficacemente notato⁴, la neutralità degli algoritmi non basta da sola a prevenire effetti (ad esempio) di discriminazione, così spostando un (ulteriore) *focus* sul problema dell'individuazione dei dati e nella loro rappresentazione.

⁴ BAROCAS, SELBST, "Big data's disparate impact", in *Calif. L. Rev.*, 104 (2016), n. 3, 671 ss.

La porzione dei dati utilizzati può essere essa stessa inquinata da *bias* ed errori sistemici dovuti (ad esempio) da discriminazioni avvenute in passato, o comunque non essere rappresentativa della realtà.

Ma non solo. La scelta delle variabili utilizzate è fondamentale, in quanto queste possono essere non rappresentative per il problema trattato o, ancor peggio, anche se apparentemente neutre (ad esempio: codice postale del quartiere di residenza) possono fungere da *proxies* (surrogati) per caratteristiche protette come etnia, genere, orientamento politico, *et coetera*, rischiando quindi di mascherare autentiche disparità di trattamento dietro semplici categorizzazioni.

Paradossalmente problemi siffatti all'interno dei dati rischiano addirittura di 'legittimare matematicamente' le ingiustizie, non perché la matematica sbagli, ma perché proprio la sicurezza affidata alla matematica algoritmica, se non correttamente indirizzata, rischia di cristallizzare gli errori presenti nei dati storici.

In ambito penale, questo significa rischio concreto di errori sistematici replicati e amplificati senza meccanismi intrinseci di controllo: il fatto che una decisione sia stata presa in modo sbagliato in passato diventa un problema per l'apprendimento delle macchine, che possono imitare *pattern* decisionali potenzialmente pericolosi, senza la capacità di rivalutarli in maniera logica.

L'AI, insomma, apprende passivamente dai dati senza vera comprensione o capacità critica. Come è stato ben rilevato⁵, strumenti sofisticati come *COMPAS*⁶ possono egualizzare le prestazioni dell'essere umano proprio perché lavorano su base statistica: si tratta, nei fatti, di inferenze empiriche che ricordano le scorciatoie cognitive e i luoghi comuni degli umani, seppur con la maggiore potenza determinata dalla matematica, piuttosto che una forma di ragionamento logico-normativo capace di spiegare e giustificare le decisioni.

D'altra parte, i decisori umani hanno la capacità di riflettere consapevolmente sui propri pregiudizi, applicare principi giuridici astratti e intervenire per correggere errori sistematici; inoltre, anche qualora le decisioni di un giudice non fossero neutrali, questo può motivare la propria opzione con

⁵ DRESSEL, FARID, *The accuracy, fairness, and limits of predicting recidivism*, in *Science Advances*, 4 (2018), n. 1, 5580.

⁶ L'acronimo COMPAS è il nome con cui è conosciuto un noto *software* utilizzato per la valutazione del rischio di recidiva negli Stati Uniti.

ragionamenti esplicativi e contestabili, contrariamente alle possibilità degli algoritmi *black box*.

Infine, anche un intervento mirato sui dati di addestramento appare difficilmente in grado di soddisfare simultaneamente tutte le principali definizioni di equità elaborate in letteratura per i sistemi di classificazione algoritmica. Come è stato dimostrato⁷, esiste un'incompatibilità matematica tra alcune metriche di equità comunemente adottate (ad esempio *calibration* e *balance for the positive/negative class*), tale per cui non possono coesistere se non in scenari rari e fortemente vincolati, lasciando quindi il problema aperto.

4. *Ragionamento o ‘pappagallo stocastico’?* Da qualche anno, come anticipato, abbiamo assistito ad una rivoluzione tecnologica nell’AI, portata dall’innovazione di architetture neurali capaci di assorbire, manipolare e trasformare i dati a vari scopi.

Uno dei prodotti di maggior successo di siffatta evoluzione è configurato dagli *LLM - Large Language Models*, modelli linguistici enormi come ChatGPT e simili – che lavorano con dati testuali (se non anche multimodali), che hanno portato una gigantesca innovazione nel campo del *NLP - Natural Language Processing*, tale da permettere agli utenti un’interazione informale e accessibile da un pubblico non specializzato.

Tuttavia, non alligna in questo fenomeno alcun parto di coscienza virtuale, né tantomeno una parvenza di ragionamento: i modelli linguistici lavorano sulle statistiche testuali, non per logica simbolica o sussunzione normativa; non distinguono tra validità giuridica e plausibilità linguistica; generano *output* verosimili, non giuridicamente fondati.

In pratica, concatenano le parole sulla base di quanto il discorso mantiene la sua coerenza, non sulla base di norme, precedenti o men che meno principi; anche quando sembrano ‘ragionare’, stanno solo riproducendo schemi linguistici appresi, senza alcuna comprensione semantica e men che meno quella capacità di applicare regole astratte a casi concreti che va sotto il nome di sussunzione.

In letteratura numerosi approcci hanno tentato di addentrarsi nel tema cercando di spiegare le differenze e le vicinanze tra questo lavoro linguistico e

⁷ KLEINBERG, MULLAINATHAN, RAGHAVAN, *Inherent Trade-Offs in the Fair Determination of Risk Scores*, in *arXiv:1609.05807*, 17 novembre 2016.

un ‘vero’ ragionamento, spesso anche in maniera interdisciplinare. Risulta ad esempio interessante l’approccio⁸ con cui si è tentato di definire un formalismo lessicale adeguato a parlare di *LLM* in modo da non creare fraintendimenti, aiutandosi con le definizioni linguistiche di ‘forma’, ‘significato convenzionale’ e ‘intento comunicativo’: ciò che viene fatto da noi umani per comunicare risulta infatti il prodotto di una relazione di questi tre elementi, mentre gli *LLM* lavorano solamente sul primo.

In pratica, non solo non ha senso attendersi dalle macchine un intento comunicativo, ma è anche impossibile comprendere quello presente nei dati di *input*, in quanto non gestito come obiettivo dell’addestramento. È solo in situazioni ‘fuori dall’ordinario’ che si notano le criticità di tale approccio, in quanto sarebbe necessaria una vera comprensione alla base per effettuare deduzioni verso direzioni che non rientravano nel dominio di addestramento del modello; senza questo, si rischia di finire nelle cosiddette allucinazioni, dove la coerenza lessicale rimane ma si perde ciò che nei dati di addestramento possedeva un vero intento comunicativo.

Intrigante la prospettiva⁹ che continua ad esplorare l’ambiguità di questi sistemi, definendoli ‘pappagalli stocastici’, che appunto imitano *pattern* linguistici, certo mantenendo coerenza su larga scala, senza però un intento comunicativo sottostante diverso da quello attribuito dall’interlocutore.

Tutto ciò di fatto rende ancora più difficile identificare eventuali problemi, in quanto mediati dai nostri stessi occhi con una visione paradossalmente distorta su quello che è la nostra comprensione; quello che sintatticamente e linguisticamente può apparire corretto, è in grado però di distorcere i messaggi trasportati nel sottotesto, che non vengono colti dall’intelligenza artificiale proprio per limiti logici di implementazione.

Il discorso sin qui condotto - diciamolo nel concludere queste brevi note sul fenomeno dell’Intelligenza Artificiale - nulla toglie all’enorme utilità dell’IA quale supporto alle decisioni da assumersi in ambito giudiziario tramite modelli statistici a carattere tradizionale, come sopra descritti, capaci, ad esempio, di fornire ed elaborare dati essenziali per ragionamenti sulle

⁸ BENDER, KOLLER, *Climbing towards NLU: On meaning, form, and understanding in the age of data*, in *Proceedings of the 58th annual meeting of the association for computational linguistics*, <https://aclanthology.org/2020.acl-main.463/>, luglio 2020, 5185 ss.

⁹ BENDER, GEBRU, McMILLAN-MAJOR, SHMITCHELL, *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?*, in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, <https://dl.acm.org/doi/10.1145/3442188.3445922>, 1 marzo 2021, 610 ss.

probabilità *a priori* di un dato accadimento, come pure nel contesto delle valutazioni del peso delle prove consentite dalla cd. formula di Bayes, su cui si tornerà in seguito.

5. Il giudice e gli algoritmi. In verità, agli appassionati dell'argomento appare chiaro già oggi – ad anni di distanza da ogni uso effettivo dell'IA nel processo di cognizione – come il cd. diritto vivente abbia scarsa simpatia per le funzioni matematiche decisionali e per la logica computazionale.

Anzi, invero, esiste chi – da esperto – ha già avuto modo di operare una diagnosi di questo fenomeno di rifiuto per l'ausilio degli algoritmi, che è di tale vastità da oltrepassare i confini del sistema penale, andando ad improntare di sé fattispecie di ‘decisori’ tra i più disparati: si tratta di una «ostilità verso gli algoritmi»¹⁰, che include medici e psicologi, giustappunto giudici ed economisti, tutti uniti in una «profonda resistenza alla demistificazione della competenza umana»¹¹ intrisa, in verità, di venature più morali che razionali.

Un buon esempio di questa avversione si trova per avventura cristallizzato da una rarissima giurisprudenza che prende in esame l'uso difensivo della cd. formula di Bayes, algoritmo decisionale il quale – pur adoperato con una certa frequenza da consulenti e periti nella valutazione esperta della prova scientifica e pur modernissimo e foriero di dense riflessioni in termini di autentica capacità di tutela della presunzione d'innocenza¹² – compare con la rarità degli animali in estinzione tra le sentenze edite, per lo più in modo nascosto e sotteso alla motivazione, in quanto adoperata dal ragionamento peritale che il giudice adotta o respinge.

Sintetizziamo così questa nota formula decisionale, riducendo al minimo le accorte riflessioni di chi ne ha studiato l'essenza in applicazione al giudizio penale¹³: secondo la formula bayesiana applicata alle funzioni decisorie penali, il rapporto fra le probabilità *a posteriori* di due ipotesi logicamente incompatibili (l'ipotesi dell'Accusa e quella della Difesa che si pongono a spiegazioni opposte del medesimo risultato di una prova, ad esempio: lo stato del cadavere costituisce prova di azione omicidiaria oppure dimostra

¹⁰ KAHNEMAN, SIBONY, SUNSTEIN, *Rumore. Un difetto del ragionamento umano*, Torino, 2024, 641.

¹¹ *Ibidem*.

¹² GARBOLINO, *Probabilità e logica della prova*, Milano, 2014, 85 ss.

¹³ Il riferimento è sempre dovuto allo studio di GARBOLINO, *op. cit., passim*.

piuttosto un atto suicida?), è una funzione del rapporto tra le probabilità *a priori* delle ipotesi medesime (quanto è probabile, sulla scorta dei dati statistici rilevanti nello specifico, che si tratti di un omicidio piuttosto che di un suicidio?) e le probabilità condizionate dal risultato della prova medesima (quanto è probabile che si tratti di un omicidio o piuttosto di un suicidio tenuto conto delle prove A-B-C?).

Si tratta di una formula matematica che in realtà, come vedremo meglio nelle pagine successive, rappresenta bene alcuni assiomi fondamentali per un processo decisionale esente da errori cognitivi, ovvero la completezza della piattaforma informativa su cui devono basarsi le decisioni, il sequenziamento delle informazioni medesime e l'attitudine al dubbio sulla reale diagnosticità delle prove; anzi, in effetti, è molto interessante anticipare qui come una delle critiche più ricorrenti della psicologia cognitiva nello studio dei processi decisionali è la rilevante tendenza dei decisorи ad evitare la considerazione delle probabilità *a priori*, per una innata propensione della psiche umana a preferire le spiegazioni causali rispetto a quelle meramente statistiche¹⁴.

Eppero, come sanno bene proprio gli statistici, che amano distinguersi sul punto in statistici ‘frequentisti’ e statistici ‘bayesiani’, la formula di Bayes è utile non solo ad evidenziare le fallacie della mera statistica frequentista, ma anche ad evitare gli scarti logici che si osservano allorquando si seguano tentativi di prescindere dalla statistica frequentista senza il supporto di strumenti altrettanto affidabili.

Il dato interessante è che la diffidenza giudiziale nei confronti dell’algoritmo bayesiano risiede paradossalmente o in una scarsa comprensione dello strumento (certo piuttosto complesso) ovvero nella sua considerazione di arnese logico in grado di condizionare eccessivamente il libero convincimento giudiziale.

Cogliamo qui due esempi: nel primo, l’uso della formula bayesiana proposto dalla difesa a supporto dell’insostenibilità logica dell’ipotesi accusatoria recepita dalla sentenza di condanna, è respinto sulla scorta di un argomento che dimostra un’effettiva assenza di cognizioni in merito alla formula bayesiana, che viene testualmente confusa con un’applicazione statistica ‘frequentista’, ovvero esattamente all’opposto della realtà del teorema di Bayes; si scrive infatti che

¹⁴ KAHNEMANN, *Pensieri lenti e veloci*, Milano, 2012, 250 e *passim*, qui citato nella versione *e-book*.

«Appare del tutto infondata la prospettazione di applicare il teorema di Bayes ai fini di valutare la probabilità che un determinato soggetto abbia commesso una determinata azione. Infatti, tale teorema, ampiamente applicato in ambito statistico, inerisce alla valutazione della probabilità che eventi ripetibili e la cui frequenza sia nota possano avvenire o essere avvenuti. La ricostruzione della condotta umana in un dato momento storico implica invece la valutazione di un evento non necessariamente ripetibile e che può rispondere a logiche devianti, irrazionali, al di fuori di logiche ripetitive con la conseguenza che la bassa percentuale di probabilità di un dato esito non implica che lo stesso non possa essere avvenuto»¹⁵.

Com'è piuttosto evidente dal testo, qui si confonde la statistica frequentista (probabilità *a priori* di un dato evento) con la statistica bayesiana (che include la valutazione congiunta delle probabilità *a priori* e delle probabilità condizionate dalle prove relative al singolo evento), così come pare piuttosto chiaro che la valutazione della sentenza sia ignara degli studi da tempo compiuti per propiziare l'uso della formula matematica bayesiana proprio come supporto ai decisorи penali nel contesto della valutazione di singole prove o multiple.

Nella seconda sentenza che si adduce ad esempio¹⁶, stavolta estraendola da quelle emesse in fase di merito, la difesa aveva criticato la condanna per omicidio volontario rilevando come la consulenza tecnica anatomico-patologica disposta dal pubblico ministero avesse dedotto la causa di morte della vittima applicando la formula di Bayes in modo esattamente opposto a quanto predicato dal teorema, per così dire stravolgendo la formula in termini tali che anche uno statistico alle prime armi avrebbe potuto avvedersi dell'errore: qui la Corte, nonostante la gravità della censura e la grossolanità del vizio, ha preferito trincerarsi nella completa *ignoratio* della censura stessa, omettendola dalla valutazione ad onta del peso della specifica prova, decisiva addirittura al fine primario di valutare se la vittima fosse morta di morte naturale o per violenza altrui.

¹⁵ Cass. Sez. II, 20 dicembre 2018, n. 19138.

¹⁶ Corte d'assise d'appello, L'Aquila, n. 6, 16 dicembre 2021, Santoleri, inedita.

Di non poco interesse, ancora in questa prospettiva, appare la disamina che compie altra sentenza¹⁷ in cui, invece, il giudice adopera direttamente la formula di Bayes per valutare il peso della prova, giungendo ad un esito assolutorio tutt’altro che scontato, ma inevitabile con una corretta applicazione del teorema, il quale – in verità – dimostra nei fatti di essere un serio baluardo dell’oltre ogni ragionevole dubbio.

Le vicissitudini della formula di Bayes dimostrano come – con un fenomeno tutt’altro che nuovo a livello scientifico¹⁸ – il giudice penale, al pari di tanti altri decisor, spesso rifugge dalle regole stringenti, privilegia le fattispecie attributive di poteri discrezionali e mal sopporta, sulla stessa lunghezza d’onda, gli strumenti logico-scientifici che pongano limiti all’esercizio del libero convincimento.

Torneremo a breve su questo tema, per notare come le acquisizioni della psicologia cognitiva spieghino quest’amore per le fattispecie attributive di poteri discrezionali, dimostrato da fenomeni processuali noti, come il caso dell’istituto dell’esperimento giudiziale, strumento che in teoria si vorrebbe privilegiato da decisor moderni, consapevoli di quanto le necessità cognitive del processo penale possano giovarsi di un mezzo vocato a ricostruire se un *quid* storicamente accaduto trovi riscontro empirico nella sua concreta modalità di accadimento; strumento prezioso in astratto, ma relegato da una prassi annosa al ruolo di mezzo di prova giustappunto ceduto alla discrezionalità giudiziale, e di poi ridotto sulle soglie della non esistenza da una nota prassi di rifiuto dell’ammissione per avventura richiesta dalla difesa¹⁹.

6. *Il bias cognitivo dell’essere umano: quasi uno sconosciuto per il giurista italiano.* Torniamo per un attimo al tema dell’IA, per svilupparne quelle che ci appaiono ad oggi le implicazioni più interessanti.

L’avvento dell’IA nel nostro universo di pensiero ci consente, ancora, e anzi a dire il vero ci costringe – proprio nella prospettiva dianzi indicata, ovvero la necessità di addestrare l’IA all’uso dei dati indispensabili al ragionamento giuridico – a riflettere su quello che abbiamo indicato come uno dei grandi

¹⁷ GIP Milano, 18 giugno 2015, Giud. Gennari, con bella nota di DELLA TORRE, *Il ‘teorema di Bayes’ fa capolino al Tribunale di Milano*, in <https://archiviodpc.dirittopenaleuomo.org>, 21 ottobre 2015.

¹⁸ L’argomento sarà oggetto di specifica trattazione nel paragrafo successivo.

¹⁹ GIUNCHEIDI, *Gli esperimenti giudiziari. L’imitazione e le potenzialità di un istituto negletto*, Milano, 2023, *passim*.

problemi che ostano all'uso dell'intelligenza artificiale in ambito decisionale giuridico, ovvero i *bias* nei dati adoperabili in addestramento, tra cui rientrano, senza meno, i precedenti giurisprudenziali.

La vera preoccupazione che deve nutrire le nostre riflessioni sull'IA è in effetti proprio questa: l'IA si nutre di dati storici, proprio come un cervello umano, con l'evidente conseguenza che i *bias* cognitivi contenuti nei dati di addestramento andranno ad inquinare la decisione dell'IA già in partenza.

Riflettiamo, per cominciare, sulle innumerevoli decisioni foriere di esegesi delle norme codistiche irrelate rispetto ai principi, come alle innumerevoli violazioni della legalità processuale che costellano la giurisprudenza di merito e di legittimità, come pure, al contempo, sul fatto che l'addestramento dell'IA costituisce per definizione il regno dei grandi numeri; sembra agevole immaginare la sorte di un addestramento che si svolga su materiale della qualità di quello che allo stato conosciamo; l'esito prevedibile sarebbe la replicazione *ad infinitum* dell'errore umano, per di più legittimato – come sopra si notava – dalla saldezza della matematica algoritmica.

Qui, banalmente, la sfida che propone un futuro uso dell'IA consiste in ciò che comunque dovremmo fare ad oggi, ovvero riuscire nell'arduo compito di risanare la legalità processuale, non foss'altro che ad impedire che l'IA ne impari le deviazioni.

Ma c'è un aspetto ulteriore ed anzi più importante che risulta disvelato dall'approccio al tema dell'intelligenza artificiale: paradossalmente il diffuso timore dei *bias* cognitivi della macchina ci costringe a rivolgere la nostra attenzione al fenomeno dei *bias* cognitivi umani, concetto evidentemente assai diverso da quello di corretta applicazione delle norme giuridiche vigenti, volgendosi direttamente agli errori commessi nella valutazione della prova e di quello che chiamiamo il fatto *sub iudice*.

Prima di addentrarsi nel tema, peraltro, occorre precisare; la riflessione giuridica italiana che comincia ad avviarsi²⁰, arriva ora a mostrare interesse per

²⁰ Alcuni scritti mostrano un nascente interesse. Volentieri segnaliamo i lavori di BLAIOTTA, *Giustizia, errore, intelligenza artificiale*, in www.sistemapenale.it, 23 ottobre 2023; DE PAOLIS, *A carte scoperte su intelligenza artificiale e processo penale. Prospettive di indagine*”, in *Arch. pen. (web)*, (2025), n. 2; FALATO, *Il giudice come utente euristico nel ragionamento induttivo basato sul senso comune. Soluzioni di debiasing per assicurare l'igiene decisionale*, in *Arch. pen. (web)*, (2025), n. 2. Segnaliamo anche interessantissime applicazioni concrete in seno a precedenti giurisprudenziali di Cassazione; cfr. in particolare Cass., Sez. VI, 3 dicembre 2024, n. 09906, sul cd. *bias* del risultato o del senso di poi, che costituisce uno dei temi più importanti trattati nel volume di Kahnemann già citato, proprio con riferi-

un tema d'indagine che, nato attorno agli anni '70 nello specifico settore della psicologia cognitiva e, in particolare, nello studio dei processi decisionali umani²¹, si va espandendo a macchia d'olio nell'universo ordinamentale d'oltreoceano, con la complicità delle neuroscienze, imponendosi all'attenzione degli utenti più disparati²² - medici, psicologi, economisti, giuristi, uffici governativi di vario ordine e grado - con la potenza di tutto ciò che s'impone non solo per l'autorevolezza della voce, ma prima ancora per l'autoevidenza dei concetti espressi e la loro percettibile centralità nell'esperienza cognitiva umana.

Data la sua attuale alienità rispetto al nostro armamentario concettuale, merita spendere qualche riga in più sul concetto di *bias* cognitivo, prima di verificare se effettivamente (ed in quale misura) esso sia destinato ad esercitare influenza anche nel sistema di giustizia penale.

7. *Due parole sul bias cognitivo, ovvero un primo approccio del processualpenalista.* Sintetizziamo, di seguito, in maniera ovviamente sommaria, le acquisizioni della psicologia cognitiva sul punto, avvalendoci fondamentalmente di alcuni studi²³, che sono lettura ineliminabile per chiunque voglia disvelare le modalità del pensiero umano nella fase decisionale, le sue capacità e le sue mistificazioni, tanto diffuse, quanto

mento ai processi giudiziari. La sentenza in questione è dedicata al processo per il disastro di Rigopiano e trovasi efficacemente annotata in senso adesivo da QUARANTA, *Il problema non è la caduta (della valanga), ma l'atterraggio (del sistema di prevenzione): la Cassazione sul disastro di Rigopiano, tra gestione (anticipata) del rischio ed esigibilità della condotta doverosa*, in www.sistemapenale.it, 22 maggio 2025. Elegantissima, poi, la disamina della sentenza in parola realizzata da CONSULICH, *Disastri evitabili e forme della colpa nell'organizzazione pubblica*, in *Cass. pen.*, 2025, 2715 e ss.

²¹ Il primo articolo di Tversky e Kahneman è stato pubblicato nella rivista *Science* nel 1974, con il titolo *Il giudizio in condizioni di incertezza: euristiche e bias*, ed è stato nuovamente pubblicato in appendice a KAHNEMAN, *op. cit.*

²² È appena il caso di ricordare che il defunto Daniel Kahneman ha vinto il premio Nobel per l'economia nell'anno 2002 e che i suoi studi sulle decisioni in condizioni d'incertezza hanno ispirato interi insegnamenti universitari nei Dipartimenti di Economia di molteplici nazioni. Nel mondo nostrano della ricerca, il pensiero dell'illustre autore è penetrato nelle riflessioni degli economisti e dei ricerchatori di scienze sociali, mentre non si rilevano approfondimenti nel mondo giuridico, men che meno con riferimento al sistema penale.

²³ Il riferimento è naturalmente al già pluricitato volume di KAHNEMAN, *Pensieri lenti e veloci*, come pure al precedente volume di KAHNEMAN, SLOVIC, TVERSKY, *Decidere nell'incertezza*, 1982, ed. italiana 2024, per i tipi di Mondadori, come pure al più recente volume di KAHNEMAN, SIBONY, SUNSTEIN, *op. cit., passim*.

percettibili solo tramite un'adeguata conoscenza del fenomeno del *bias* cognitivo.

Precisiamo qui - in abbrivio - una personale convinzione: è tempo che iniziamo a lasciarci alle spalle il solipsismo che ha indotto la dottrina processualpenalistica ad elidere troppo spesso il ricorso a scienze ‘altre’ per arricchire di contenuti necessari la ‘forma’ su cui lavora il processualpenalista, anzi, per contribuire ad individuare proprio la forma più corretta a tutelare gli interessi sottostanti.

Ma, senza ulteriori divagazioni, eccoci al punto.

Secondo gli autorevoli e rivoluzionari approdi della psicologia cognitiva cui abbiamo fatto sopra riferimento, le modalità di pensiero umano sono convenzionalmente suddivise in due ‘sistemi’, denominati sistema N. 1 e N. 2, etichetta consuetudinaria che lo stesso Kahneman riferisce ad ulteriori studiosi che l’hanno preceduto, non senza celare come essa sia fornita di tale utilità esplicativa da cancellare il rischio che realmente si pensi alla mente umana come ad un’entità scissa²⁴.

Il sistema N. 1 è forgiato per fornire risposte veloci, utilissime nella quotidianità personale e professionale; esso opera “in fretta e automaticamente, con poco o nessuno sforzo, e nessun senso di controllo volontario”²⁵ e ciò anche allorquando compie attività astrattamente complesse, ma magari usuali per il singolo, come una mossa di scacchi. Al contrario il sistema mentale N. 2 è chiamato a compiere attività lente ed impegnative, corrispondenti all’esperienza soggettiva della concentrazione, che esige, per la sua persistenza, assenza di distrazioni.

Il sistema N. 1 adopera, nel corso del suo funzionamento, alcune scorciatoie cognitive, dette ‘euristiche’, ossia procedimenti associativi che consentono di raggiungere un risultato provvisorio, che esigerebbe verifica e infatti l’interazione tra i due sistemi è ovviamente fondamentale, atteso che l’uno compensa l’altro nelle attività cerebrali: il sistema 1 è il più attivo, e produce continuamente materiali d’analisi per il sistema 2, costituiti da impressioni, intuizioni, sensazioni, ragionamenti semplici basati su scorciatoie cognitive, mentre il sistema 2 entra in opera solo allorquando emerge un dato che il sistema 1 non è in grado di elaborare efficacemente, ultimando operazioni complesse sostanzialmente prive di scorciatoie cognitive.

²⁴ KAHNEMAN, *op. cit.*, p. 50.

²⁵ *Ivi*, pp. 51 ss.

Un dato fondamentale è che il riflessivo sistema N. 2 compie operazioni che esigono sforzo, ed è un'altra delle nozioni che ci derivano dalla psicologia cognitiva, come dalle neuroscienze, il fatto che la stessa chimica del cervello lo induce a ‘risparmiare energie’, accontentandosi delle strategie cognitive intuitive del sistema N. 1, in base ad una «sofisticata allocazione dell’attenzione...affinata attraverso una lunga storia evolutiva»²⁶.

Impossibile, qui, non citare ampiamente l'autore, allorquando descrive cosa accade quando l'essere umano si trova dinanzi ad una situazione reale da analizzare a fini decisionali:

«Per il sistema 1 la misura del successo è la coerenza della storia che riesce a costruire. La quantità e la qualità dei dati su cui si basa la storia sono in gran parte irrilevanti. Quando le informazioni sono scarse, cosa che accade spesso, il sistema 1 funziona come una macchina per saltare alle conclusioni... La combinazione di un sistema 1 in cerca di coerenza con un sistema 2 affetto da pigrizia, fa sì che il sistema 2 avalli molte credenze intuitive, le quali rispecchiano fedelmente le impressioni generate dal sistema 1»²⁷.

I *bias* cognitivi nascono proprio dai genetici difetti di un sistema 1 sempre impegnato a reagire con velocità, tramite le scorciatoie cognitive che ne costituiscono ineliminabile caratteristica, provocando il meno possibile l'attivazione dell'oneroso sistema 2, in una inarrestabile compulsione a comprendere (credere di comprendere), scegliere, decidere e agire; il sistema 1 – si scrive²⁸, suscitando echi che dovrebbero colpire il processualpenalista – non è affatto incline al dubbio, tende anzi, piuttosto, a reprimere ogni ambiguità, lottando per mantenere la coerenza con la prima idea che ha ricevuto al momento del contatto con la situazione.

Queste teorie – accennate già nel prima articolo sul tema, datato al 1974²⁹ – sono state sviluppate da Kahneman, dal suo gruppo di lavoro e dai colleghi ed allievi nei quarant'anni successivi, scoprendo una quantità di *bias* cognitivi, validandoli *in vitro* attraverso decine di esperimenti e raccogliendo migliaia di dati su decisioni e anche previsioni dell'uomo della strada, come di economisti, finanziari, politologi, medici, psichiatri, avvocati, giudici ed esperti

²⁶ *Ivi*, 93.

²⁷ *Ivi*, 245.

²⁸ *Ivi*, 316.

²⁹ TVERSKY, KAHNEMAN, *op. cit.*, *passim*.

vari, tra cui quelli di scienze forensi, come la medicina legale o la dattiloscopia.

La correttezza delle conclusioni di questi studi può essere sperimentata in proprio da chiunque li legga e possieda l'ineludibile coraggio che occorre per verificare quanto spesso le decisioni che assumiamo, private e professionali, sono il frutto di scorciatoie cognitive foriere di errori.

Tra i principali *bias* analizzati dalla psicologia cognitiva nello studio dei processi decisionali umani, ci limiteremo ad evocare, in virtù della loro importanza per il processo penale, sin d'ora intuibile, il *bias* cagionato dall'euristica della disponibilità, il *bias* della conferma e, infine, il cd. *hindsight bias*, reso in italiano come *bias del senno di poi* o *bias del risultato*.

8. *Bias e sistema penale.* È interessante osservare come in questo complesso sistema che si è cercato di tracciare con mano veloce, la psicologia cognitiva individua, per così dire, una genesi essenziale degli errori cognitivi nel fenomeno cerebrale battezzato con l'acronimo di WYSIATI, siglatura della frase “*What You See Is All There Is*”, traducibile in lingua italiana come “quello che vedi è tutto quello che c’è”.

Si tratta della scoperta della tendenza ineliminabile del cd. sistema 1 a decidere ignorando la potenziale incompletezza delle informazioni disponibili, in sostanza accontentandosi di quel che trova e sottostimando, anzi, la rilevanza delle informazioni potenzialmente mancanti (perfino laddove esse siano magari immaginabili), al contempo sovrastimando la qualità delle informazioni già esistenti.

In sintesi, ciascun decisore - inclusi quelli altamente specializzati della materia da decidere, anzi per certi versi proprio quelli altamente specializzati³⁰ - tende, in maniera del tutto inconsapevole, ad affidarsi ai risultati delle valutazioni espresse dal sistema 1 sulla base di quelle che banalmente chiameremmo le prime impressioni, ad onta del fatto che l'*imprinting* ottenuto nell'esaminare la materia della decisione sia stato fondato su materiale informativo parziale o erroneo. Per converso, sempre il sistema 1 tenderà inesorabilmente (e per quanto se ne sia inconsapevoli) a cancellare i

³⁰ In realtà, i decisori esperti risultano, all'analisi della psicologia cognitiva, affetti da molteplici *bias*, tra cui alcuni decisamente caratteristici del *genus*, come il *bias* chiamato “illusione di validità”. Sul tema di tornerà appresso, ma sin d'ora può dirsi che esso è diffusamente trattato nel volume di KAHNEMAN, SIBONY, SUNSTEIN, *op. cit., passim*.

dubbi ricostruttivi che emergerebbero al controllo del sistema 2, cercando conferme piuttosto che smentite rispetto alla valutazione iniziale.

L'esito di questo fenomeno è, giustappunto, l'apertura del decisore verso i *bias* cognitivi.

Come si avverte, il fenomeno cognitivo denominato WYSIATI è stato fondamentale in termini di evoluzione della specie, poiché esso facilita «la realizzazione della coerenza e della fluidità cognitiva che ci inducono ad accettare un'affermazione come vera»³¹ e dunque consente di decidere e reagire con rapidità; esso ci rende capaci di pensare e decidere in fretta, sì, ma a costo della perdita di affidabilità nelle scelte che ne derivano, poiché «né la quantità né la qualità delle prove» contano nella prospettiva del sistema 1 e del decisore che ne segue gli *input*.

E' difficile sottrarsi all'immediata connessione di queste note con una serie di problematiche tipiche del processo penale, in particolare laddove si consideri quelli che Kahneman individua come 'antidotì' allo stato di cose determinato dal WYSIATI: la ricerca di informazioni quantitativamente superiori prima di valutare l'opzione decisoria, il vaglio di qualità sulle informazioni, la diversificazione e l'autonomia delle fonti informative, il sequenziamento delle informazioni, l'accettazione del dubbio, come fonte di scoperta di opzioni decisorie meno intuitivamente disponibili e dunque meno gradite al sistema 1.

Siamo in presenza, in effetti, qui, di un insegnamento che costruisce un fondamento scientifico saldo e ineludibile alle regole del contraddittorio, a quelle in materia di utilizzabilità della prova e, prima di ogni altra, in ordine logico e cronologico, alla negletta regola della completezza delle indagini.

Soffermiamoci su quest'ultimo punto, per una ragione in particolare, ovvero che la carenza informativa e la tendenza del sistema 1 a trarre conclusioni immediate sia pure sulla scorta di dati incompleti, è individuata dalla psicologia cognitiva appunto come luogo d'origine di molteplici *bias* di cui gli attori sono del tutto ignari.

Il principio di completezza delle indagini - lo sappiamo - è stato a sua volta indicato dalla Corte costituzionale³² come pre-condizione di un vaglio effettivo tanto sulla decisione di inazione quanto sull'azione esercitata dal pubblico ministero.

³¹ KAHNEMAN, *op. cit.*, p. 248.

³² Corte cost., sent. n. 88 del 1991.

Siamo egualmente ben consapevoli di quanto la regola della completezza declinata dalla Consulta e per di più costruita dal legislatore ordinario sul monito codicistico di cui all'art. 358 c.p.p. (e dunque, esplicitamente, sull'obbligo di indagare anche su elementi di prova che ridondino a discarico) sia stata di fatto cancellata dalla prassi quotidiana del sistema giustizia.

Leggendo gli studi sui processi decisionali, riusciamo ora a comprendere come questo tipo di regola schiettamente giuridica (la completezza delle indagini, il dovere di indagare anche su spunti informativi non collimanti con l'ipotesi investigativa) non costituisce solo il frutto dell'applicazione di principi, ma sia una vera e propria regola di «igiene decisionale»³³, con riferimento a tutte le risoluzioni che, nel *sequitur* procedimentale, vengono assunte con riferimento alla data *notitia criminis*: dall'informativa di PG che magari predilige una certa ipotesi ricostruttiva piuttosto che un'altra, all'adozione di un dato strumento investigativo, sino alle richieste del pubblico ministero (cautelari personali o reali, di azione/inazione, etc.), a vedere bene, par chiaro che ogni singola decisione, che costella la fase delle indagini preliminari sino alla sua conclusione, è fitta di possibilità di errori cognitivi, cagionati da euristiche concretamente inevitabili se non si seguono dati precetti, i quali esigono, giustappunto quale pre-condizione di igiene decisionale, la completezza del quadro informativo. E ciò non (solo) perché lo comanda una precisa norma di legge e lo impone un principio di rango costituzionale, ma anche perché, senza il rispetto di quella pre-condizione, è enorme il rischio che l'indagine preliminare si tramuti in una sequela di *bias* cognitivi, che si replicano e si moltiplicano esponenzialmente.

Perché qui si noti: un vaglio non sorvegliato dei primi dati probatori emergenti all'abbrivio delle indagini, genera convinzioni forgiate dal *bias* cd. della disponibilità (scelta veloce sui dati immediatamente disponibili), la quale, a sua volta, genera agevolmente *bias* di conferma, ovvero l'inesorabile propensione del procedimento valutativo verso dati di riprova della convinzione originaria, come pure *bias* di effetto alone (o coerenza emozionale esagerata), che inducono gli inquirenti ad essere letteralmente ciechi³⁴ rispetto ad elementi conoscitivi incoerenti con l'ipotesi prescelta.

³³ Così pure recepisce per il processo penale la terminologia restituita dalla psicologia cognitiva, FALATO, *op. cit.*, p. 10.

³⁴ Esempi multipli di cecità cognitive evidenti sono riportati dagli studiosi in tutti e tre i volumi citati in abbrivio.

In altre parole, lo scienziato cognitivo spiega al giurista un dato con cui sembra indispensabile fare i conti: la natura unilaterale delle indagini preliminari è capace di generare e, per così dire, incistare all'interno del processo, una serie di *bias* originari, che tendono a trovare conferma all'interno del prosieguo decisionale, ivi compreso quello affidato al giudice e anche in presenza di contraddittorio, non solo perché - come si notava - il *bias* della conferma si autorigenera, provocando ulteriori *bias* di conferma nei decisori successivi, ma anche perché l'unilaterale direzione dell'indagine provoca la dispersione delle potenziali prove a discarico, sempre inesorabilmente compromesse, per definizione, dal decorso del tempo³⁵. Non è casuale che lo stesso autore degli studi su cui fondiamo le presenti riflessioni, analizzi questo tipo di *bias* proprio con riferimento a processi penali reali, leggendo nella situazione un precipizio cognitivo verso l'errore giudiziario³⁶.

9. *I decisori nel procedimento penale.* Il discorso viene qui appena accennato, nell'intento di cogliere spunti per successivi studi e andrebbe fondato su approfondite ricerche empiriche.

Consideriamo, in effetti, che con il termine 'decisore' la psicologia cognitiva indica il soggetto cui compete la scelta in situazioni incerte, con riferimento al futuro (previsione) o anche al passato, come quando occorre valutare quale sia la causa di un dato evento, tra le varie immaginabili (ad esempio: morte naturale o per violenza altrui).

Il sistema penale include importanti decisioni proiettive, *de futuro*, come nel caso in cui il giudice della cautela valuta la possibilità di reiterazione del reato *ex art. 274, lett. c), c.p.p.* o in sede di concessione della sospensione condizionale della pena; ma - come si accennava poc'anzi - l'intero contesto normativo strutturato dal codice di procedura penale è costellato di decisioni in situazione incerta: se il giudice è il decisore per antonomasia, in sede cautelare e tanto più in sede di cognizione, è indispensabile sottolineare come 'decisore' sia anche l'ufficiale di polizia giudiziaria che sceglie quale pista investigativa seguire prioritariamente e di conseguenza seleziona il materiale probatorio pertinente al suo approfondimento, non meno del pubblico ministero che delega una determinata attività d'indagine piuttosto che altra, ai

³⁵ Lapidaria sul punto Corte EDU, Sezione I, sent. 2 marzo 2017, Talpis c. Italia.

³⁶ KAHNEMAN, SIBONY, SUNSTEIN, *op. cit., passim*, ma in particolare p. 600.

sensi dell'art. 370, comma 1, c.p.p. o decide di chiedere al GIP l'autorizzazione a disporre intercettazioni su questa o quell'utenza o l'acquisizione dei dati provenienti dai tabulati; ma decisore è anche il consulente tecnico del pubblico ministero che identifica una certa causa di morte o l'esperto di scienze forensi che decide di poter identificare (oppure no) una corrispondenza con il DNA reperito sulla scena del crimine.

Ciascuna di queste decisioni è potenzialmente gravida di conseguenze, conducendo all'acquisizione di una prova, piuttosto che alla sua dispersione, oppure conduce ad un risultato probatorio falsato da *bias*, sulla cui base vengono assunte iniziative cautelari, magari di carattere restrittivo della libertà personale.

Gli esempi potrebbero continuare e sono tanti quanti gli snodi che costellano un'indagine, la sua conclusione in un senso o nell'altro e, infine, l'intera attività di cognizione di cui consta il giudizio.

Ciascuno dei decisori di questi snodi produce decisioni capaci di interagire con quelle successive, talora influenzandole *funditus*.

Per misurare l'entità del problema dei *bias* e delle loro interazioni con il prosieguo procedimentale è sufficiente porre mente ad esempi notori, come la sgargiante attualità del caso di Garlasco, oppure a casi analizzati dalla dottrina in modo dettagliato, come il processo di Perugia per la morte di Meredith Kercher³⁷.

Consideriamo, in particolare, i rilievi spesi dal cognitivista su un problema di sistema penale ben noto anche all'interno del nostro ordinamento, ossia la prassi – non codificata, ma costante e ben rilevabile dagli atti – di fornire agli esperti forensi (ad esempio: al consulente tecnico del PM piuttosto che alle sezioni scientifiche delle forze dell'ordine) porzioni intere del fascicolo delle indagini o anche il fascicolo nella sua integralità, unitamente al materiale su cui viene demandato il vaglio.

Anni di studi della psicologia cognitiva dedicati esattamente a questa specifica abitudine, invalsa in molteplici sistemi ordinamentali, anche di *common law*, dimostrano in termini tanto difficilmente confutabili quanto inquietanti, che le decisioni degli esperti forensi sullo specifico tema di prova ad essi affidato mutano in termini drastici in relazione alle informazioni aggiuntive sul caso che vengono ad essi fornite, informazioni pertinenti rispetto al caso

³⁷ V. sul punto il volume *L'assassinio di Meredith Kercher. Anatomia del processo di Perugia*, a cura di Montagna, Roma, 2012, *passim*.

medesimo, ma non rilevanti quanto allo specifico tema sottoposto ad *expertise*³⁸: nelle centinaia di situazioni esaminate - le più varie: dal confronto delle impronte digitali, all'identificazione di immagini, sino all'analisi anatomo-patologica sulle cause di morte - gli esiti diventano esiti di conferma dell'ipotesi investigativa perseguita anche laddove sarebbero stati negativi qualora l'esperto non avesse ricevuto alcuna informazione aggiuntiva oltre a quelle strettamente pertinenti al *thema probandum*.

Si colgano le implicazioni di questo fondamentale passaggio:

«L'effetto delle informazioni condizionanti non è circoscritto alle conclusioni degli esaminatori (identificazione, esclusione o inconcludenza), ma incide *su ciò che percepiscono*, oltre che su *come* interpretano tale percezioni... gli esaminatori posti in un contesto affetto da *bias* non vedono le stesse cose - in senso letterale - di chi non viene esposto ad informazioni condizionanti: quando l'impronta latente è accostata alla stampa di un'impronta archiviata *target*, gli esaminatori osservano un numero decisamente inferiore di dettagli (chiamati *minutiae*) rispetto a quanti ne individuano se gli viene mostrata la sola impronta latente... La suscettibilità degli esperti forensi al *bias* di conferma non è un semplice problema teorico, perché sul piano pratico non vengono attuate misure precauzionali sistematiche di nessun tipo per garantire che gli esaminatori non siano esposti a informazioni condizionanti»³⁹.

Come detto, si tratta di rilievi letteralmente inquietanti, soprattutto perché fondati su dati sperimentali ed empirici solidissimi; rilievi che sottendono come le regole processuali siano ad oggi del tutto inadeguate ad impedire fenomeni capaci di condurre direttamente all'errore giudiziario.

Ma spigoliamo ancora, sempre per cogliere spunti dalla psicologia cognitiva con riferimento a temi cari al processualpenalista, come il rapporto tra *media* e processo penale, e notiamo questo aspetto: per il neuroscienziato e per il cognitivista i *media* causano le cd. cascate di disponibilità, determinando l'importanza o al contrario la collaterаль di dati eventi, causando illusioni di diffusione di un certo fenomeno magari di fatto episodico, determinando - in modo quasi incancellabile per un normale processo cognitivo - l'immagine (positiva o negativa) dei soggetti di cui si occupano, con conseguente *bias*

³⁸ Usiamo, qui, i termini di pertinenza e rilevanza nello specifico significato ben esplicato da TRAPELLA, *Indagini preliminari e parità delle armi*, Milano, 2025, 235 ss.

³⁹ KAHNEMAN, SIBONY, SUNSTEIN, *op. cit.*, p. 598 ss.

denominato ‘effetto alone’, e se è facile notare che accurati studi⁴⁰ denunziano la rischiosità delle orge mediatiche⁴¹ scatenate in occasione di taluni (troppi) processi penali, purtroppo è ancora diffusa l’illusione che il Giudice professionale, a differenza delle giurie popolari, sia immune al contagio di *bias* cognitivi provenienti dai *media*⁴².

Non è così, e sono proprio gli studi dei processi decisionali a dimostrarlo in termini letteralmente inconfutabili, anzi: l’(illusoria) convinzione dei decisori esperti in merito alla propria igiene decisionale è un altro *bias* cognitivo, provvisto di nome proprio, guarda caso come «illusione di validità»⁴³.

10. *Regole per i processi decisionali.* In effetti, ‘illusione di validità’ è un sintagma alquanto adatto a descrivere quel che rimane delle nostre idee sulla capacità del sistema penale di approdare a decisioni da considerarsi giuste in senso anzitutto etico.

L’insegnamento di Kahnemann *et alia* è imperioso, epperò curiosamente consonante con la mentalità del processualpenalista, quando si rifletta che,

⁴⁰ Per doviziosi riferimenti v. *Informazione e giustizia penale. Dalla cronaca giudiziaria al “processo mediatico”*, a cura di Triggiani, Bari, 2022, *passim*.

⁴¹ Il termine è nuovamente di KAHNEMAN, *op. cit.*, p. 24.

⁴² Impossibile non ricordare, qui, la decisione emessa dalla Corte EDU proprio con riferimento ad un caso italiano notissimo, ovvero il processo contro Bettino Craxi (Corte EDU, 5 dicembre 2002, n. 34896/97): con assoluta insensibilità per le argomentazioni scientifiche della psicologia cognitiva, pure all’epoca ben note nell’universo di *common law*, la Corte rileva quanto segue: “Secondo il ricorrente, il suo processo si sarebbe svolto in un clima di ostilità la cui responsabilità ricadrebbe sulle autorità giudiziarie italiane e sui media. La stampa scritta, parlata e televisiva avrebbe stabilito la sua colpevolezza prima della decisione dei giudici. Questa campagna avrebbe influenzato sia l’opinione pubblica, sia i giudici chiamati a pronunciarsi nei vari processi. Peraltro, la procura avrebbe sistematicamente fornito alla stampa ed ai media informazioni coperte dal segreto istruttorio... La Corte rileva che nella fattispecie l’interesse dei media italiani per il caso e l’importanza che questo rivestiva agli occhi dell’opinione pubblica derivava dalla posizione eminentemente occupata dal ricorrente, dal contesto politico nel quale i fatti incriminati avevano avuto luogo, così come dalla natura e dalla gravità di questi ultimi. La Corte ritiene che sia inevitabile, in una società democratica, che la stampa esprima dei commenti a volte severi su un caso sensibile che, come quello del ricorrente, metteva in discussione la moralità di alti funzionari ed il rapporto tra il mondo della politica e quello degli affari. Inoltre, è il caso di notare che gli organi giudiziari che hanno trattato la causa erano interamente composti da giudici professionali. Contrariamente ai membri di un giurì, questi ultimi godono di un’esperienza e di una formazione che permette loro di fugare qualsiasi suggestione esterna al processo. Peraltro, la condanna del ricorrente è stata pronunciata alla fine di un procedimento in contraddittorio, nel corso del quale l’interessato ha avuto la possibilità di sottoporre alle autorità giudiziarie competenti gli argomenti che riteneva utili alla sua difesa”. Argomenti insostenibili, quando si rifletta sulle centinaia di pagine scritte sul punto dalla psicologia cognitiva.

⁴³ KAHNEMAN, *op. cit.*, p. 587.

per questi studiosi, la cura del male sta nell'adozione di regole per i processi decisionali che riducano il coefficiente di discrezionalità di ciascun decisore, regole che vengono appunto denominate di ‘igiene decisionale’; anzi, precisiamo meglio: per il cognitivista, la processualizzazione – si perdoni il gioco di parole – dei processi decisionali, è la chiave di volta per approdare ad una decisione esente da *bias* cognitivi e la discrezionalità dei decisorii corrisponde ad una apertura al *bias* tanto più ampia quanto più è vasta l’area libera da regole vincolanti.

Chi scrive ha dedicato anni di ricerca al tema della discrezionalità riservata dal codice di procedura penale a molteplici e importanti decisioni del giudice e del pubblico ministero, evidenziandone i vasti margini di contrasto con i principi costituzionali, l’intrinseca fragilità delle garanzie al confronto concreto con l’esercizio di poteri discrezionali, il valore infinitesimale delle forme di controllo previste. Sotto tale profilo, è un’enorme soddisfazione verificare che questa personale idiosincrasia per uno tra i molteplici snodi critici del processo penale attuale, ha un solidissimo fondamento non solo nei principi del diritto, ma – prima ancora – nella scienza dei processi cognitivi e decisionali.

Alla regola primaria appena illustrata, ne vengono aggiunte altre: mentre andrebbe garantita, mediante sanzioni processuali efficaci, la completezza informativa dell’investigazione nel suo complesso ed andrebbero potenziate le indagini della difesa, ad evitare i *bias* derivanti dalla parzialità della visione inquirente (WYSIATI), i singoli processi decisionali destinati ad incidere sulla decisione finale – tipicamente, le valutazioni degli esperti sulle prove a carattere scientifico – andrebbero sottratte *ab initio* alla possibilità del *bias* della conferma, sigillandole dall’influenza reciproca mediante comunicazione agli esperti dei soli dati direttamente rilevanti per il tema devoluto, evitando la trasmissione a ciascuno dell’intero compendio investigativo. Riflettiamo, qui, in effetti – quanto a questo specifico punto – sul fenomeno che si verifica tutte le volte in cui (ovvero, sempre) al perito nominato in incidente probatorio venga rilasciato il fascicolo delle indagini preliminari nella sua integralità: l’esito inesorabile, mai a sufficienza riflettuto in dottrina, è che addirittura il giudice chiamato a decidere sul merito della causa, viene esposto, tramite le cognizioni ricevute dal perito, ad un *bias* di conferma della convinzione espressa su elementi probatori raccolti durante le indagini

preliminari in assenza di contraddittorio, come se il principio di separazione delle fasi non esistesse affatto.

Più in generale, le nuove conoscenze derivanti, per il processualista, dalla psicologia cognitiva, devono necessariamente indurre a rivedere l'idea che il centro motore del processo penale sia l'istante cognitivo racchiuso nel dibattimento; il dato di realtà conforta un'idea ben diversa, per cui la centralità del giudizio non può e non deve occultare il fatto che, senza una severa igiene decisionale, la fase precedente, schiettamente esplorativa, configurata dalla indagini preliminari, è destinata a pesare come un macigno sulle sorti del giudizio.

Ed ecco un'altra regola che dovremmo attentamente meditare: apprezzare le fondamenta della formula decisionale bayesiana in sede di valutazione delle prove, sullo schema essenziale che lo stesso Kahneman propone.

«La prima è che le probabilità *a priori* contano anche in presenza di prove sul caso in questione: è un dato che spesso non è intuitivamente ovvio. La seconda è che l'impressione che abbiamo intuitivamente della diagnosticità delle prove è spesso esagerata. La combinazione del WYSIATI con la coerenza associativa tende a farci credere nelle storie che ci raccontiamo da soli... sono entrambe idee semplici, ma per me fu uno *shock* quando mi resi conto che non mi avevano mai insegnato ad applicarle e ancora oggi trovo innaturale farlo»⁴⁴.

Negli studi citati in abbrivio risultano esaminati numerosi casi giudiziari reali e numerosi esperimenti su casi giudiziari fintizi sottoposti ad esperti in cui l'applicazione della regola bayesiana nei semplici termini espressi dallo psicologo cognitivo, è capace (o sarebbe stata capace, nell'ipotesi di processi reali) di condurre ad esiti rispettosi di quelle che, a ben vedere, sono regole di igiene cognitiva che corrispondono, in termini addirittura rigorosi, alle regole del giusto processo, inclusa – ed anzi soprattutto considerata – la presunzione d'innocenza.

Si consideri – in via di mera esemplificazione di ben altre verifiche che il processualista dovrebbe compiere – quanto un principio di diritto come quello espresso dalla notissima sentenza Franzese⁴⁵ in materia di causalità, sia distante, nonostante l'apparenza, da un'autentica logica cognitiva giudiziale esente da *bias*.

⁴⁴ KAHNEMAN, *op. cit.*, p. 436 s.

⁴⁵ Cass. pen., SU, 10 luglio 2002, n. 30328, Franzese.

Secondo quella notissima pronunzia, anche coefficienti di probabilità statistica medio-bassi possono essere utilizzati per riconoscere un nesso di condizionamento tra condotta ed evento, purché vi sia positivo riscontro in merito alla sicura non incidenza, nel caso *sub iudice*, di ulteriori fattori causali interagenti in via alternativa.

Formulata con riferimento al reato omissivo improprio, questo principio giurisprudenziale ha avuto enorme fortuna, anche al di fuori dell'alveo, pure assai ampio, del reato colposo; si pensi, ad esempio, all'usatissima formula adoperata nelle decisioni di condanna per variegate tipologie di reati, ma classicamente per l'omicidio volontario, secondo cui la difesa non avrebbe addotto causali alternative idonee a spiegare l'evento concretamente verificatosi⁴⁶.

Eppure, nonostante ben si comprendano gli intenti esplicativi della formula/Franzese, una volta letta con lo spirito di chi ha apprezzato le teoriche sui processi decisionali, essa appare come il risultato di un *bias* del risultato o del senno di poi, quando si rifletta che conduce a trascurare le reali probabilità *a priori* di un dato evento (in quel caso: la frequenza statistica dell'evento morte nelle date condizioni), favorendo anche coefficienti statistici insufficienti a superare il ragionevole dubbio.

Queste brevi riflessioni intendono essere l'abbrivio di uno studio più ampio, che dovrebbe investire ogni snodo decisionale del procedimento penale, esaminandone i processi cognitivi sottintesi, al fine di disvelarne la fragilità, la soggezione a scorciatoie cognitive, la vulnerabilità all'uno o all'altro *bias*.

La realtà è che, prima di poter immaginare di nutrire e addestrare un sistema decisionale artificiale con i prodotti delle decisioni umane, dobbiamo compiere ancora un lungo cammino attraverso la legalità processuale, nel tentativo di rifondare l'idoneità cognitiva del procedimento penale.

⁴⁶ Inutili i riferimenti giurisprudenziali per questo formulario buono ad ogni uso; la teorica delle ipotesi alternative non dimostrate dalla difesa, è un *leit motiv* che costella sentenze di merito e di legittimità.