

LLM e ragionamento giuridico. Perché la fluidità non è giudizio

di *Walter Quattrocioni*

Il contributo propone un'analisi critica dell'impatto dei modelli linguistici di grandi dimensioni sul ragionamento giuridico e sull'esercizio della funzione giudiziaria. La diffusione di strumenti in grado di generare testi giuridicamente plausibili introduce un rischio di natura non tecnica, ma cognitiva e istituzionale: la progressiva sostituzione della verifica argomentativa con la mera plausibilità linguistica. A partire dall'esame del funzionamento degli LLM come sistemi di generazione probabilistica del linguaggio, l'articolo evidenzia la loro strutturale mancanza di criteri interni di verità, rilevanza giuridica e responsabilità. La coerenza formale dell'*output* può alimentare un'illusione di comprensione idonea a incidere sulle condizioni di formazione del giudizio. Viene pertanto messa in luce l'asimmetria qualitativa tra giudizio umano e *output* artificiale, delimitando ambiti di impiego compatibili con la giurisdizione, circoscritti a funzioni meramente ausiliarie. Ne deriva l'esigenza di una competenza critica volta a preservare l'autonomia e la responsabilità del giudicare.

1. Perché il tema riguarda direttamente chi giudica / **2.** Cosa fanno davvero gli LLM (e cosa no) / **3.** Linguaggio giuridico e illusione di comprensione / **4.** L'asimmetria decisiva: giudizio umano *vs.* *output* artificiale / **5.** Dove l'LLM può essere utile senza essere pericoloso / **6.** implicazioni per la funzione giudiziaria / **7.** Conclusione: una competenza nuova, non uno strumento nuovo

1. Perché il tema riguarda direttamente chi giudica

L'ingresso dei sistemi di intelligenza artificiale generativa nel discorso giuridico non è avvenuto attraverso riforme, protocolli o sperimentazioni istituzionali, ma per via laterale: tramite strumenti di uso comune, facilmente accessibili, capaci di produrre testi giuridicamente plausibili in pochi secondi. Questa modalità di penetrazione è già di per sé significativa. Non si tratta dell'introduzione di una nuova tecnologia nel processo, ma dell'emersione di un nuovo *ambiente cognitivo* in cui il lavoro giuridico rischia di essere progressivamente riformattato.

Per chi esercita la funzione giudiziaria, il problema non è l'accuratezza media di questi sistemi, né la loro capacità di imitare lo stile della motivazione o della dottrina. Il nodo centrale è un altro: la *tentazione delegante* che nasce quando la qualità formale dell'*output* anticipa e, talvolta, sostituisce l'atto del giudicare. La fluidità del linguaggio, la coerenza argomentativa e la sicurezza espositiva producono una soglia di accettazione precoce, che può indurre a scambiare una risposta ben scritta per una risposta fondata.

Questo rischio non riguarda singoli comportamenti impropri, ma ha una portata istituzionale. Il giudizio giuridico è un'attività strutturalmente vincolata: ai fatti accertati, alle regole del procedimento, al contraddittorio, alla responsabilità della decisione. L'adozione, anche informale, di strumenti che producono soluzioni linguisticamente convincenti, ma prive di criteri interni di verifica introduce una

frattura silenziosa tra forma e fondamento, tra apparenza di razionalità e razionalità giuridica effettiva.

Per questo motivo la questione non può essere ridotta a un problema tecnico, né affrontata solo in termini di affidabilità o di percentuali di errore. Anche un sistema altamente performante, se utilizzato senza una chiara comprensione del suo funzionamento, può alterare le condizioni cognitive del giudizio. Il punto non è stabilire se l'AI “funzioni”, ma chiarire *in che modo il suo funzionamento interagisce con l'esercizio della funzione giudiziaria*, e quali spazi di decisione rischiano di essere occupati, di fatto, da un automatismo linguistico privo di responsabilità.

In questa prospettiva, comprendere cosa siano davvero i modelli linguistici e cosa non possano essere non è un'esigenza di alfabetizzazione tecnologica, ma una condizione preliminare per preservare l'autonomia del giudizio.

2. Cosa fanno davvero gli LLM (e cosa no)

Per valutare in modo corretto l'impatto dei modelli linguistici di grandi dimensioni (“*Large Language Models*”, LLM) sul ragionamento giuridico è necessario chiarire, senza ricorrere a metafore fuorvianti, quale sia il loro principio di funzionamento. Gli LLM non interpretano testi, non comprendono fatti, non valutano norme. Generano linguaggio.

Ogni *output* prodotto da un LLM è il risultato di un processo di *generazione probabilistica*: il modello stima, passo dopo passo, quale sequenza linguistica sia più plausibile sulla base delle regolarità statistiche apprese durante l'addestramento. Il testo non è il prodotto di un ragionamento che muove da premesse verso conclusioni, ma di una traiettoria che attraversa regioni di alta densità nello spazio del linguaggio. In questo senso, la coerenza dell'*output* non è una garanzia di correttezza, ma un effetto strutturale del modello.

Da questo deriva una seconda conseguenza fondamentale: gli LLM *non possiedono criteri interni di verità, correttezza o rilevanza giuridica*. Non sono in grado di distinguere tra una fonte autorevole e una inesistente, tra un principio consolidato e una formulazione solo apparentemente plausibile, tra un fatto provato e una ricostruzione verosimile. Tutti questi elementi, per il modello, sono semplicemente *pattern* linguistici con diversa frequenza di occorrenza. Il sistema non verifica, non confronta con il mondo, non riconosce i propri limiti conoscitivi.

In questo quadro, il termine comunemente utilizzato di “allucinazione” risulta fuorviante. Esso suggerisce l'idea di un malfunzionamento episodico, di una deviazione rispetto a un comportamento normalmente corretto. In realtà, ciò che viene etichettato come allucinazione è l'esito coerente di un sistema che genera linguaggio senza accesso a criteri epistemici. Quando il modello produce una citazione inesistente o una ricostruzione giuridica infondata, non sta “sbagliando” nel senso umano del termine: sta operando esattamente secondo la logica che lo governa.

Per il giurista, questa distinzione è cruciale. Non ci si trova di fronte a uno strumento fallibile ma migliorabile con l'aumento dei dati o della potenza di calcolo, bensì a un sistema che, per costruzione, non può sapere se ciò che afferma è vero, pertinente o

giuridicamente valido. Comprendere questo punto significa spostare il problema dal piano dell'affidabilità tecnica a quello, ben più rilevante, della *compatibilità tra generazione linguistica automatica e attività di giudizio*.

3. Linguaggio giuridico e illusione di comprensione

Il linguaggio giuridico possiede una forza persuasiva intrinseca. La forma argomentativa, la struttura della motivazione, il richiamo a fonti e precedenti non svolgono solo una funzione espressiva, ma contribuiscono a costruire l'idea stessa di razionalità della decisione. In questo contesto, la capacità degli LLM di produrre testi giuridicamente ben formati rappresenta un fattore di particolare criticità. Non perché tali testi siano necessariamente errati, ma perché *appaiono immediatamente legittimi*.

La coerenza testuale, tuttavia, non-coincide con la fondatezza giuridica. Un testo può essere internamente coerente e al tempo stesso infondato, costruito su presupposti non verificati, su ricostruzioni fattuali solo verosimili, o su riferimenti normativi inesistenti o fuori contesto. Nel ragionamento umano, la forma linguistica è il risultato finale di un percorso di verifica; nei modelli linguistici, è invece il punto di partenza e il criterio di successo del processo generativo. Questa inversione introduce un rischio sottile ma profondo: la sostituzione della verifica con la plausibilità.

È in questo scarto che si colloca il fenomeno che può essere descritto come *epistemia*: l'illusione di comprensione prodotta da un linguaggio fluido, articolato e privo di frizioni, che induce chi legge a ritenere concluso il processo di valutazione prima che esso abbia effettivamente avuto luogo. L'epistemia non è semplice disinformazione né ignoranza. È una forma di falsa competenza, in cui la qualità espositiva maschera l'assenza di un ancoraggio epistemico.

Nel contesto giudiziario, questo rischio assume una dimensione sistemica. Il giudizio non è solo il risultato di un atto cognitivo individuale, ma un passaggio istituzionale che produce effetti giuridici vincolanti. Quando l'apparenza di razionalità sostituisce la razionalità giuridica, il pericolo non è l'errore isolato, ma la normalizzazione di decisioni fondate su basi linguistiche anziché su criteri di prova, rilevanza e responsabilità.

Per questo motivo, l'uso non sorvegliato di strumenti capaci di simulare il linguaggio del diritto rischia di incidere non tanto sull'esito delle singole decisioni, quanto sulle condizioni stesse in cui il giudizio viene formato. L'epistemia non riguarda ciò che i modelli producono, ma *ciò che inducono chi decide a non fare*: verificare, dubitare, assumersi fino in fondo il peso della decisione.

4. L'asimmetria decisiva: giudizio umano vs. *output* artificiale

Il confronto tra giudizio umano e *output* artificiale non può essere impostato in termini di prestazioni, velocità o accuratezza comparata. L'asimmetria decisiva è di natura qualitativa e istituzionale. Il giudizio umano è un atto responsabile; l'*output* di un modello linguistico non lo è, né può diventarlo.

Il giudice non si limita a produrre una decisione, ma ne risponde. La motivazione non è un esercizio retorico, bensì il luogo in cui il percorso decisionale diventa verificabile, contestabile, rivedibile. Ogni passaggio è esposto al controllo delle parti, dei gradi successivi di giudizio, dell'ordinamento nel suo complesso. Questa struttura di responsabilità non è un accessorio del giudicare, ma il suo fondamento.

Un LLM, al contrario, non risponde di nulla. Non può giustificare le proprie affermazioni, non può rendere conto delle fonti utilizzate, non può distinguere tra ciò che è stato inferito e ciò che è stato semplicemente generato come sequenza plausibile. La verificabilità dell'*output* è sempre esterna al sistema e interamente a carico dell'utilizzatore umano. In assenza di un controllo attivo, l'*output* resta una proposizione senza garanzie.

A questa asimmetria si aggiunge un elemento cruciale: il problema del “non so di non sapere”. Il giudice è consapevole dei propri limiti cognitivi e istituzionali. Sa quando un caso è complesso, quando mancano elementi, quando è necessario approfondire o sospendere il giudizio. Questa consapevolezza dell'incertezza è parte integrante della funzione giudiziaria. Gli LLM, invece, non possiedono alcuna metacognizione. Non riconoscono l'assenza di informazione, non segnalano l'insufficienza dei dati, non sospendono il giudizio. Producono sempre una risposta.

Per questo motivo l'AI generativa non può essere considerata un soggetto valutativo, neppure in forma attenuata. Valutare implica attribuire peso, riconoscere rilevanza, assumersi una responsabilità rispetto alle conseguenze della decisione. Un sistema che opera esclusivamente sul piano della plausibilità linguistica, senza criteri interni di verità e senza consapevolezza dei propri limiti, non può svolgere questa funzione senza alterarne la natura.

L'errore concettuale più insidioso non consiste nel ritenere l'AI infallibile, ma nel trattarla come se fosse fallibile “come un umano”. Le due forme di errore non sono comparabili. L'errore umano è situato, contestualizzato e imputabile; l'errore artificiale è strutturale, opaco e non imputabile. Confondere queste due dimensioni significa spostare il rischio dal piano individuale a quello sistemico, con effetti potenzialmente rilevanti sulla tenuta del processo decisionale.

5. Dove l'LLM può essere utile senza essere pericoloso

Chiarire i limiti strutturali dei modelli linguistici non implica negarne ogni possibile utilità nel contesto giuridico. Al contrario, una comprensione realistica del loro funzionamento consente di individuare ambiti di impiego compatibili con la funzione giudiziaria, a condizione che il controllo umano resti effettivo e non meramente formale.

Gli LLM possono svolgere un ruolo legittimo come *strumenti di supporto redazionale ed esplorativo*. Possono assistere nella riorganizzazione di testi complessi, nella sintesi di materiali voluminosi, nella riformulazione di passaggi argomentativi già fondati, o nell'esplorazione preliminare di orientamenti dottrinali e giurisprudenziali che dovranno comunque essere verificati. In questi casi, il modello non introduce nuovi contenuti decisionali, ma agisce come un amplificatore di capacità operative già presenti, riducendo il carico cognitivo senza incidere sul nucleo del giudizio.

Esistono, tuttavia, ambiti in cui l'uso degli LLM diventa intrinsecamente critico. Il *suggerimento normativo*, la comparazione automatica tra soluzioni giuridiche, il bilanciamento di interessi contrapposti o, più in generale, qualsiasi forma di *weighing* implicano valutazioni che non possono essere ricondotte a una semplice manipolazione linguistica. In questi contesti, il rischio non è solo quello dell'errore, ma quello di una *traslazione implicita della funzione valutativa* verso un sistema che non dispone dei presupposti epistemici e istituzionali per esercitarla.

Il confine tra assistenza e sostituzione non coincide quindi con l'intenzione soggettiva dell'utilizzatore, ma con l'effetto oggettivo dell'uso dello strumento. Anche quando il giudice mantiene formalmente la decisione finale, l'adozione di suggerimenti linguistici non verificati può influenzare in modo significativo la direzione del ragionamento, orientando la scelta delle categorie, delle soluzioni e persino delle conclusioni. In questo senso, la sostituzione può avvenire in forma attenuata e progressiva, senza che vi sia una delega esplicita.

Per evitare questo slittamento, è necessario che l'uso degli LLM sia concepito come *ausilio strumentale e non cognitivo*. Il modello può aiutare a scrivere, ma non a decidere; può facilitare l'accesso al materiale, ma non selezionarne la rilevanza; può accelerare operazioni formali, ma non anticipare il giudizio. Ogni utilizzo che oltrepassi questo confine rischia di compromettere la distinzione, fondamentale per lo Stato di diritto, tra supporto tecnico e responsabilità decisionale.

6. Implicazioni per la funzione giudiziaria

Una delle convinzioni più diffuse nel dibattito pubblico sull'AI è che l'aumento della scala – più dati, modelli più grandi, maggiore potenza di calcolo – possa risolvere progressivamente i problemi di affidabilità. Nel caso degli LLM applicati al diritto, questa convinzione è fuorviante. La scala migliora la fluidità, raffina la coerenza, riduce alcuni errori superficiali, ma *non introduce criteri epistemici*. Un modello più grande resta un modello che genera linguaggio senza accesso alla verità, alla rilevanza giuridica o alla responsabilità della decisione. Il problema non è quantitativo, ma strutturale.

Per la funzione giudiziaria, questo implica che non esiste una soglia tecnologica oltre la quale l'AI diventa "affidabile" in senso decisionale. Anche sistemi estremamente sofisticati continuano a operare sul piano della plausibilità linguistica. Affidarsi alla scala come soluzione rischia quindi di spostare l'attenzione dal nodo reale: la compatibilità tra automatismi linguistici e giudizio giuridico.

In questo contesto, il ruolo della *formazione del magistrato* diventa centrale. Non si tratta di acquisire competenze informatiche avanzate, né di imparare a "programmare" o a ottimizzare l'uso degli strumenti. La competenza cruciale è di tipo critico: saper riconoscere quando un *output* è solo ben costruito, distinguere tra supporto formale e contenuto decisionale, comprendere i limiti epistemici dei sistemi utilizzati. Formare all'uso critico dell'AI significa formare alla non-delega, alla vigilanza cognitiva, alla consapevolezza dei propri processi decisionali.

Questa esigenza è coerente con le principali linee di riflessione istituzionale a livello europeo. Le indicazioni elaborate in ambito CEPEJ insistono sulla centralità del

controllo umano, sulla trasparenza e sulla non sostituzione della funzione giudiziaria. Analogamente, l'*AI Act* colloca l'uso dell'AI nei contesti ad alto rischio all'interno di un quadro di cautele che presuppone una responsabilità umana piena e non trasferibile. Tuttavia, tali cornici normative rischiano di restare astratte se non vengono tradotte in pratiche operative e culturali condivise.

La cautela istituzionale non deve essere intesa come resistenza all'innovazione, ma come *condizione della sua governabilità*. Introdurre strumenti potenti in un ambito delicato come quello giudiziario senza una chiara comprensione dei loro effetti cognitivi equivale a modificare le condizioni del giudizio senza discuterne le conseguenze. La vera sfida non è decidere se usare o meno l'AI, ma stabilire *come preservare l'autonomia del giudizio* in un ambiente informativo sempre più saturo di risposte automatiche.

7. Conclusione: una competenza nuova, non uno strumento nuovo

Il confronto tra intelligenza artificiale generativa e funzione giudiziaria non ruota attorno all'adozione di un nuovo strumento, ma alla necessità di sviluppare una *competenza nuova*. Non si tratta di imparare a utilizzare l'AI in modo più efficiente, bensì di comprenderne i limiti strutturali per evitare forme di delega implicita che possono incidere sul processo decisionale senza essere formalmente riconosciute.

Comprendere come funzionano i modelli linguistici significa, per il giudice, preservare la propria autonomia cognitiva. Significa sapere quando una risposta è solo linguisticamente plausibile, quando una ricostruzione è coerente ma infondata, quando la fluidità del testo rischia di anticipare e chiudere prematuramente il giudizio. In questo senso, la conoscenza tecnica di base diventa uno strumento di autodifesa istituzionale.

Governare l'uso dell'AI senza mitizzarlo richiede di sottrarsi tanto alla narrativa dell'entusiasmo quanto a quella del rifiuto. L'AI generativa non è un soggetto, non è un decisore, non è un interprete del diritto. È uno strumento che opera su un piano diverso da quello del giudizio giuridico e che, proprio per questo, deve essere mantenuto in una posizione di supporto rigorosamente controllata. Ogni slittamento verso la sostituzione, anche parziale, non è un progresso tecnologico, ma una perdita di responsabilità.

Il giudizio resta uno spazio non automatizzabile. Non per ragioni etiche astratte, ma per ragioni strutturali: perché implica assunzione di responsabilità, consapevolezza dei limiti, esposizione al controllo e alla critica. Nessun sistema che genera linguaggio senza criteri di verità e senza coscienza della propria ignoranza può occupare questo spazio senza alterarne la natura.

La sfida posta dall'AI generativa alla giustizia non è quindi quella di decidere se affidarsi alle macchine, ma di *rafforzare le condizioni umane e istituzionali del giudicare* in un contesto informativo sempre più denso, fluido e persuasivo. È in questa direzione che si gioca il futuro di un uso responsabile della tecnologia nella giurisdizione.